Elicitation and Aggregation of Crowd Information

THÈSE N^{O} 7242 (2016)

PRÉSENTÉE LE 26 OCTOBRE 2016

À LA FACULTÉ INFORMATIQUE ET COMMUNICATIONS
LABORATOIRE D'INTELLIGENCE ARTIFICIELLE
PROGRAMME DOCTORAL EN INFORMATIQUE ET COMMUNICATIONS

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

POUR L'OBTENTION DU GRADE DE DOCTEUR ÈS SCIENCES

PAR

Goran RADANOVIC

acceptée sur proposition du jury:

Prof. M. Grossglauser, président du jury
Prof. B. Faltings, directeur de thèse
Prof. D. Parkes, rapporteur
Prof. A. Krause, rapporteur
Prof. K. Aberer, rapporteur

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Suisse
2016

If people do not believe that mathematics is simple,
it is only because they do not realize how complicated life is.
— John von Neumann

To my family…

Acknowledgements

To my thesis director Prof. Boi Faltings for the excellent supervision, enduring support, insightful discussions, and wise guidance. To Prof. Matthias Grossglauser, Prof. David Parkes, Prof. Andreas Krause, and Prof. Karl Aberer, for agreeing to take part in my thesis jury and providing me with valuable comments. To Prof. Kate Larson for introducing me to multi-agent systems and all the discussions that followed. To Jason Li and Arnaud Jutzeler for sharing their air pollution models. To Radu Jurca for contributing to some of the results of this thesis. To Sylvie Thomet for translating the tutorial of the peer grading experiment into French and all of the other administrative services. Alphabetically, to Aleksei Triastcyn, Ana Radanovic, Anastasia Mavridou, Florent Garcin, Igor Kulev, and Ivan Radanovic, for valuable feedback on this thesis, stimulating conversations and their life-changing wisdom. To former and current members of the AI lab for helping me in various ways, such as with teaching assistantships, experiments and presentations; apart form the mentioned colleagues, these alphabetically include: Christos Dimitrakakis, Claudiu Musat, Fei Mi, Li Pu, Ludek Cigler, Marina Boia, Naman Goel, Stephane Martin, Sujit Gujar, Valentina Sintsova. To many students whom I worked with on exciting projects, including Natalija Gucevska, whose work was partially included in the thesis. To many friends, colleagues, and acquaintances, for all the joyful moments that we shared together during the past five years. Most importantly, to my family, especially my parents Stjepan Radanovic and Zorka Radanovic, for immense support and love throughout my life.

Thank you very much!

Lausanne, October 2016 G. R.

Abstract

This thesis addresses challenges in elicitation and aggregation of crowd information for settings where an information collector, called *center*, has a limited knowledge about information providers, called *agents*. Each agent is assumed to have noisy private information that brings a high information gain to the center when it is aggregated with the private information of other agents. We address two particular issues in eliciting crowd information: 1) how to incentivize agents to participate and provide accurate data; 2) how to aggregate crowd information so that the negative impact of agents who provide low quality information is bounded. We examine three different information elicitation settings.

In the first elicitation setting, agents report their observations regarding a single phenomenon that represents an abstraction of a crowdsourcing task. The center itself does not observe the phenomenon, so it rewards agents by comparing their reports. Clearly, a rational agent bases her reporting strategy on what she *believes* about other agents, called *peers*. We prove that, in general, no payment mechanism can achieve *strict properness* (i.e., adopt truthful reporting as a strict equilibrium strategy) if agents only report their observations, even if they share a common belief system. This motivates the use of payment mechanisms that are based on an additional report. We show that a general payment mechanism cannot have a *simple* structure, often adopted by prior work, and that in the limit case, when observations can take real values, agents are constrained to share a common belief system. Furthermore, we develop several payment mechanisms for the elicitation of non-binary observations.

In the second elicitation setting, a group of agents observes multiple *a priori* similar phenomena. Due to the *a priori* similarity condition, the setting represents a refinement of the former setting and enables one to achieve stronger incentive properties without requiring additional reports or constraining agents to share a common belief system. We extend the existing mechanisms to allow non-binary observations by constructing *strongly truthful* mechanisms (i.e., mechanisms in which truthful reporting is the highest-paying equilibrium) for different types of agents' population.

In the third elicitation setting, agents observe a time evolving phenomenon, and a few of them, whose identity is known, are trusted to report truthful observations. The existence of trusted agents makes this setting much more stringent than the previous ones. We show that, in the context of online information aggregation, one can not only incentivize agents to provide informative reports, but also limit the effectiveness of malicious agents who deliberately misreport. To do so, we construct a reputation system that puts a bound on the negative

impact that any misreporting strategy can have on the learned aggregate.
Finally, we experimentally verify the effectiveness of novel elicitation mechanisms in community sensing simulation testbeds and a peer grading experiment.

Key words: Game theory, Mechanism design, Incentive schemes, Peer prediction, Reputation systems, Online learning, Crowdsourcing, Community sensing, Peer grading

Résumé

Cette thèse aborde les défis posés par l'obtention et l'agrégation de l'information de groupe (*crowd information*) dans les cas où un collecteur d'information, appelé *centre*, n'a qu'une connaissance limitée des fournisseurs d'information, appelés *agents*. Nous supposons que chaque agent dispose d'une information privée impure, qui apporte un fort gain d'informations au centre lorsqu'elle est agrégée avec l'information privée d'autres agents. Nous aborderons deux cas particuliers de l'obtention de l'information de groupe : 1) comment motiver les agents à participer et à fournir des données précises ; 2) comment agréger l'information du groupe afin que l'impact négatif des agents fournissant une qualité d'information inférieure soit limité. Nous examinerons trois cas différents d'obtention de l'information.

Dans le premier cas d'obtention d'information, les agents rapportent leurs observations d'un phénomène unique, qui représente une abstraction d'une tâche de production participative. Le centre n'observe pas lui-même le phénomène, mais récompense les agents en comparant leurs rapports. À l'évidence, un agent rationnel rapportera ses observations selon une stratégie basée sur ce qu'il *pense* des autres agents, appelés *pairs*. Nous prouverons qu'en général, il n'existe pas de mécanisme de paiement satisfaisant la propriété d'*amélioration rigoureuse* (c'est-à-dire garantissant l'adoption d'un rapport véridique comme stratégie d'équilibre) si les agents rapportent seulement leurs observations, même s'ils partagent une croyance commune. Ce résultat motive l'utilisation de mécanismes de paiement qui soient basés sur un rapport additionnel. Nous montrerons qu'un mécanisme de paiement général ne peut pas avoir une structure simple, pourtant souvent utilisée dans les travaux de recherche précédents, et que, dans le cas limite, lorsque les observations peuvent prendre des valeurs réelles, les agents sont contraints de partager une croyances commune. De plus, nous développerons plusieurs mécanismes de paiement pour l'obtention d'observations non binaires.

Dans le deuxième cas d'obtention d'information, un groupe d'agents observent plusieurs phénomènes *a priori* similaires. Grâce à cette condition de similarité, ce cas représente une amélioration du cas précédent, et nous permet d'obtenir des propriétés de motivation plus fortes, sans exiger de rapports additionnels, ni forcer les agents à partager une croyance commune. Nous étendrons les mécanismes existants aux observations non binaires en construisant des mécanismes *fortement véridiques* (à savoir des mécanismes dans lesquels rapporter la vérité constitue l'équilibre offrant la plus grande récompense) pour différents types de population d'agents.

Dans le troisième cas, les agents observent un phénomène évoluant en fonction du temps, et

certains d'entre eux, dont l'identité est connue, sont chargés de rapporter des observations véridiques. L'existence d'agents dignes de confiance rend ce cas nettement plus rigoureux que les cas précédents. Nous montrerons que, dans le contexte d'un agrégateur d'information en ligne, il est non seulement possible de motiver les agents à donner des rapports informatifs, mais également de limiter la portée des faux rapports délivrés par des agents malintentionnés. À cette fin, nous construirons un système de réputation qui imposera une limite à l'impact négatif que peut avoir n'importe quelle stratégie trompeuse sur l'information agrégée.
Finalement, nous vérifierons expérimentalement l'efficacité de nouveaux mécanismes d'obtention de l'information dans une simulation de détection en communauté et dans une expérience d'évaluation par les pairs.

Mots clefs : Théorie des jeux, Théorie des mécanismes d'incitation, Mécanisme incitatif, Prédiction par les pairs, Systèmes de réputation, Apprentissage en ligne, Production participative, Détection communautaire, Évaluation par les pairs

Contents

List of Figures

List of Tables

Bibliographic Note

This thesis is based on the following publications:

- Radanovic G., Faltings B., and Jurca R. 2016. Incentives for effort in crowdsourcing using the peer truth serum. [RF16a]
- Radanovic G., and Faltings B. 2016. Limiting the influence of low quality information in community sensing. [RF16c]
- Radanovic G., and Faltings B. 2015. Incentive schemes for participatory sensing. [RF15a]
- Radanovic G., and Faltings B. 2015. Incentives for subjective evaluations with private beliefs. [RF15b]
- Radanovic G., and Faltings B. 2014. Incentives for truthful information elicitation of continuous signals. [RF14]
- Radanovic G., and Faltings B. 2013. A robust Bayesian truth serum for non-binary signals.[RF13]

1 Introduction

Involving many stakeholders in a decision making process is beneficial as it ensures that the reached decision is preferred by the majority of the individuals affected by it. Therefore, it comes as no surprise that modern systems rely on collective intelligence formed by aggregating information from multiple sources. The most notable example of such systems is the *participatory web*, where dynamic contents are created by engaging users in the design, thus enabling them to share their knowledge and experience. MTurk[1], TripAdvisor[2], or PredictWise[3], are only some of many examples on the modern web that strongly rely on crowd intelligence.

The participatory web is often guided by the *wisdom of the crowd* approach [Sur05]: correctness is obtained by aggregating contributions from many non-expert individuals, as it is in *product reviewing, opinion polling, crowdsourcing* or *community sensing*. This approach, however, fails to provide correct aggregates in certain cases, when participants provide false or incorrect information either because [Gho13, LRSH11, HPZ06]:

1. obtaining accurate information requires effort;
2. participants have ulterior motives;
3. participants are biased towards their prior information.

These sources of inefficiencies imply that one of the key challenges is to solicit accurate information owned by the crowd, while limiting the negative influence of crowd participants with ulterior motives. The problem becomes even more challenging when reported information cannot be directly verified by a party that elicits it.

To obtain quality data, an elicitation mechanism can incentivize participation using rewards that may come in different forms but have a proper structure, so that the participants receive

[1] www.mturk.com
[2] www.tripadvisor.com
[3] www.predictwise.com

the highest rewards for providing the most accurate data. In general, without a possibility of directly verifying elicited data, rewards have to be based on comparison of reported values. In other words, a participant's reward depends on what others report, which induces a *game* among participants from a game theoretic point of view. Therefore, the presented information elicitation scenario can be modeled using standard game theory tools, where individuals are represented by *self-interested agents* who reveal their private information only if that is in their best interest [MRZ05]. The goal of an elicitation mechanism is to construct a game, i.e., a *mechanism*, in which participants are incentivized to invest effort and reveal their private information.

This approach has been tried with success on crowdsourcing platforms such as MTurk. [Har11] considered the task of screening resumes for a job description. A scheme where payments depended on the agreement of answers with those of a human resources expert provided significant improvements in accuracy. [SCH11] tested a large variety of payment mechanisms using a task of classifying the type of content present on a web site, and found that the mechanisms based on consistency of reports had the best performance. Giving rewards for agreeing with another participant has also been used in the very successful ESP game [vAD04], where players were rewarded for assigning the same label as a peer to an image. [KH12] proposed to reward crowd workers based on the comparison of their answers with the aggregate obtained from the crowd. [HF13b] investigated reward based on a consistency with a peer using a task of counting nouns in a list of 30 English words. Crowd workers were rewarded with a bonus whenever their answer agreed with that of a single, randomly chosen peer. They found that this increases accuracy more than comparing against a gold standard. The same authors also showed that the social pressure can further increase accuracy [HF13a]. [FPTJ14] designed a peer consistency mechanism that allows the answer distribution to be biased, and showed that it can correct anchoring bias in a counting task on MTurk. Furthermore, [GF14] compared this mechanism to *prediction markets* [Han03, CP07] — information aggregators that perform well in practice [PLGN01, NRTV07]. They found that the peer consistency mechanism can achieve a similar performance.

Motivated by these results, we consider two important objectives in elicitation and aggregation of crowd information from a game-theoretic perspective:

1. how to incentivize participants to invest effort in acquiring accurate information and truthfully reveal it;
2. how to limit the negative influence that participants with ulterior motives might have on the aggregated information.

Incentive mechanism design

There are two main directions in the incentive mechanism design: *gold standard* mechanisms and *peer consistency* mechanisms. The former is based on the existence of gold standards to design scoring rules, typically, for elicitation of distribution properties or predictions [Han03, GR07, LPS08].

We focus on *peer consistency* mechanisms, where incentives are formed by comparing values reported by different participants — agents. Since these mechanisms do not rely on gold standards, they are applicable even in scenarios when the ground truth is not possible to define (as for *subjective information*) or when it becomes known in a distant future. There are two basic types of peer consistency mechanisms, the *peer prediction* [MRZ05] and the *Bayesian truth serum* [Pre04], that differ in the amount of knowledge they have about agents' beliefs and the amount of information they elicit from agents. The peer prediction is a *minimal* mechanism, in a sense that it elicits only targeted information, but it assumes a certain knowledge about agents' beliefs. Contrary to the peer prediction, the Bayesian truth serum is a *knowledge-free* mechanism, but it elicits additional information, in particular, it elicits agents' beliefs. Notice that both mechanisms are dependent on agents' beliefs, which are formed through agents' *belief systems* that model how agents reason about each other's private information.

Considerable amount of literature has been devoted to making these mechanisms more robust, often achieving better properties only for binary information structures [JF09, WP12b, WP12a, DG13]. For example, robustness of the Bayesian truth serum in terms of the size of agents' population has been analyzed in [WP12b], while [DG13] substantially improves the properties of the peer prediction by modifying the classical peer prediction setting. These mechanisms, however, do not completely generalize to elicitation scenarios with non-binary information structures, often requiring additional restrictions on how agents form their beliefs. This resembles a common pattern in game theory that *two significantly differs from three or many*[4]. Therefore, in this thesis, we address the following challenge:

Challenge 1: *Designing robust knowledge-free incentive mechanisms applicable to elicitation of non-binary information.*

Information aggregation

In certain cases, incentives for quality might not be sufficient to prevent low quality reports in the elicited data sets. This happens when participants do not respond to incentives, either because they have ulterior motifs, as is the case for malicious participants, or because they are *spammers* who provide random data.

A common way of addressing this issue is by using statistical inference methods for noise reduction in the elicited data sets [RYZ+10, KOS11, LPI12, KOS13, VVV13]. However, these methods are not designed for an online information fusion scenario where the reported data has to be processed on the fly, as it is in real time community sensing.[5] A particular

[4]For example, the famous Arrow's impossibility theorem [Arr70] requires a non-binary outcome space.

[5]While one could potentially apply the mentioned methods repeatedly in batch mode, we are more interested

algorithm suitable for this type of an online setting is the influence limiter algorithm [RS07], first proposed for recommender systems. The algorithm places a reputation mechanism on top of a predefined information fusion component: its reputation updating rule and influence limitation procedure make it provably resistant to a wide range of manipulation strategies. However, the algorithm does not scale very well with the number of participants, so it is impractical for many crowdsourcing scenarios. Hence, the second challenge we address in this thesis is:

Challenge 2: *Designing a robust online aggregation method suitable for large scale crowdsourcing.*

Our contributions

We develop our results systematically, by considering three different elicitation settings and providing for each of them mechanisms with provable elicitation properties.

Setting 1: Single-task elicitation

The first setting we consider is the classical peer consistency setting where participants have private information about a single object, which we refer to as a *phenomenon* or a *task*. We start by establishing the necessity of eliciting additional information in knowledge-free elicitation. In particular, we show that there does not exist a *strictly proper* knowledge-free minimal mechanism unless agents' belief systems are highly constrained. Remarkably, even for non-minimal mechanisms whose structure is *decomposable* (in a sense that they separately score targeted and additional information), a condition that agents share a common belief system[6] does not suffice to allow truthful elicitation of private information. This explains why the Bayesian truth serum [Pre04] and its robust version for small population [WP12b] do not fully extend to more general elicitation scenarios that simultaneously allow small populations and non-binary private information, without putting additional restrictions on how agents form their beliefs. In addition to the common belief condition (i.e., agents sharing a common belief system), we define a mild constraint for which we construct a strictly proper mechanism called the **multi-valued robust Bayesian truth serum**. This shows that decomposable mechanisms are strictly more general than minimal mechanisms.

To further push the limits of possibility results in terms of the type of elicited information, we investigate the realm of non-decomposable mechanisms. In cases where agents with similar private information also have similar beliefs, it is possible to construct a strictly proper multi-report mechanism. One such mechanism is the **divergence-based Bayesian truth serum** in which agents with similar private information are penalized if their beliefs are substantially dif-

in a design that also includes incentives for informed reporting.

[6]A common belief system means that the agents acquire their beliefs in the same way. In particular, if two agents have the same private information, they should also have the same beliefs.

ferent. The mechanism can be adopted for the elicitation of real-valued information. We refer to the modified mechanism as the **continuous Bayesian truth serum**. Like the prior Bayesian truth serums, it requires agents' to have a common belief system, but it does not put any additional restrictions on the agents' beliefs. Furthermore, we show that without the common belief condition, any mechanism fails to elicit truthful real-valued information, thus proving the generality of the divergence-based Bayesian truth serum in its setting. Using a community sensing simulation testbed, we quantitatively demonstrate the importance of mechanisms designed for the elicitation of information with a non-binary structure. Furthermore, we discuss the design of a Bayesian truth serum type of mechanism called the **competitive Bayesian truth serum**, which is a building block of a contest for eliciting subjective information.

Setting 2: Multi-task elicitation

To address the issue of the common belief condition, we consider a variation of the basic peer consistency setting in which agents provide their private information about multiple *phenomena*. This is a common case in micro-task crowdsourcing where, for example, a worker solves several micro-tasks and reports a set of answers, each answer being associated to a different task. With such an information structure, we design a strictly proper minimal elicitation mechanism that allows agents to have private (uncommon) belief systems; we call it the **minimal peer prediction with private priors**. The mechanism assumes that the agents are *homogeneous* in a way they acquire their private information. Furthermore, if additionally the population of agents is large, we show how to make truthful reporting the highest paying strategy profile using the **logarithmic peer truth serum**. We say that the logarithmic peer truth serum is *strongly truthful*.

We then focus on heterogeneous population of agents. As we show with our impossibility results, the transition from homogeneous to heterogenous population is not trivial, so as a compromise between the two ends of the spectrum, we developed the **robust peer truth serum**. The mechanism allows limited heterogeneity of agents' population under a *mild* restriction on the agents' beliefs. As its variant for homogeneous populations (the logarithmic peer truth serum), the robust peer truth serum makes the truthful reporting the most profitable strategy profile (i.e., it is strongly truthful), but it requires smaller number of peers (per task) to evaluate contributions, and, thus is numerically more stable. Due to its incentive properties and relatively simple structure, we advocate the use of the robust peer truth serum in peer grading and community sensing, which we also support by experiments and simulations.

Setting 3: Elicitation with trusted agents

Finally, we consider a setting that is more stringent than the previous two, as some agents are trusted to provide truthful information. However, the existence of trusted agents enables us to address Challenge 2. We consider a reputation based framework, suitable for filtering out low quality reports. We first show that a *simple* reputation based approach, often used in practice, fails to limit the negative influence of malicious agents, even when they adopt relatively simple misreporting strategies. Second, we prove that the influence limiter [RS07] is not directly applicable to a large scale crowdsourcing due to its computational complexity. To overcome the drawback of the influence limiter, we modify its reputation updating rule and its influence limitation procedure. The novel algorithm is called the **stochastic influence limiter**, and, just like the original influence limiter, it is resistant to manipulative behaviour. We evaluate the algorithm on a community sensing simulation testbed and empirically verify that it outperforms a state of the art reputation system for sensing.

Summary of the most important contributions

To summarize, the most important contributions of the thesis are:

- Two impossibility results highlighting the need for non-minimal peer consistency mechanisms and a more sophisticated design of these mechanisms. The formal results are presented in Theorem 1 and Theorem 2.
- A general positive result regarding the non-minimal *strictly proper* elicitation, knowledge-free of agents' common belief system. The result is presented through two novel mechanisms — the *divergence-based BTS* and the *continuous BTS* — and their incentive properties, stated in Theorem 4, Corollary 2, Theorem 5 and Theorem 6.
- Two positive results regarding the *strongly truthful* elicitation in the multi-task setting, for acquiring large amount of data from a homogeneous population of agents and for eliciting information from a heterogeneous population of agents whose beliefs satisfy the *self-predicting* condition. The results are respectively presented through two novel mechanisms — *logarithmic PTS* (log-PTS) and the *robust PTS* (RPTS) — and their incentive properties, stated in Theorem 9, Theorem 10, Theorem 11, and Theorem 12.
- A modified version of the influence limiter algorithm, called the *stochastic influence limiter* (SIL), suitable for large scale crowdsourcing. The stochastic influence limiter is provably resistant to myopic misreporting strategies (Theorem 14 and Theorem 15), while having low computational complexity (Theorem 13) and incentivizing strategic agents to provide informative reports (Theorem 16).
- Quantitative validations of the developed mechanisms in community sensing testbeds and an empirical study on the performance of peer consistency methods in peer grading.

Figure 1.1 provides classifications of peer consistency methods according to different criteria, thus, relating the contributions of this thesis to other peer consistency methods.[7] Note that

[7] For the definition of strict properness and strong truthfulness, and the explanation of informed truthfulness, we refer the reader to Chapter 3 and Chapter 4. Targeted information is elicited through an *information report*, while an agent's posterior belief is elicited through a *prediction report*.

not all of the peer consistency methods are included in the figure. More thorough comparison can be found in the related work sections of the following chapters. Furthermore, Table 1.1 shows quantitative studies on incentive mechanism design that are (most) related to our work.

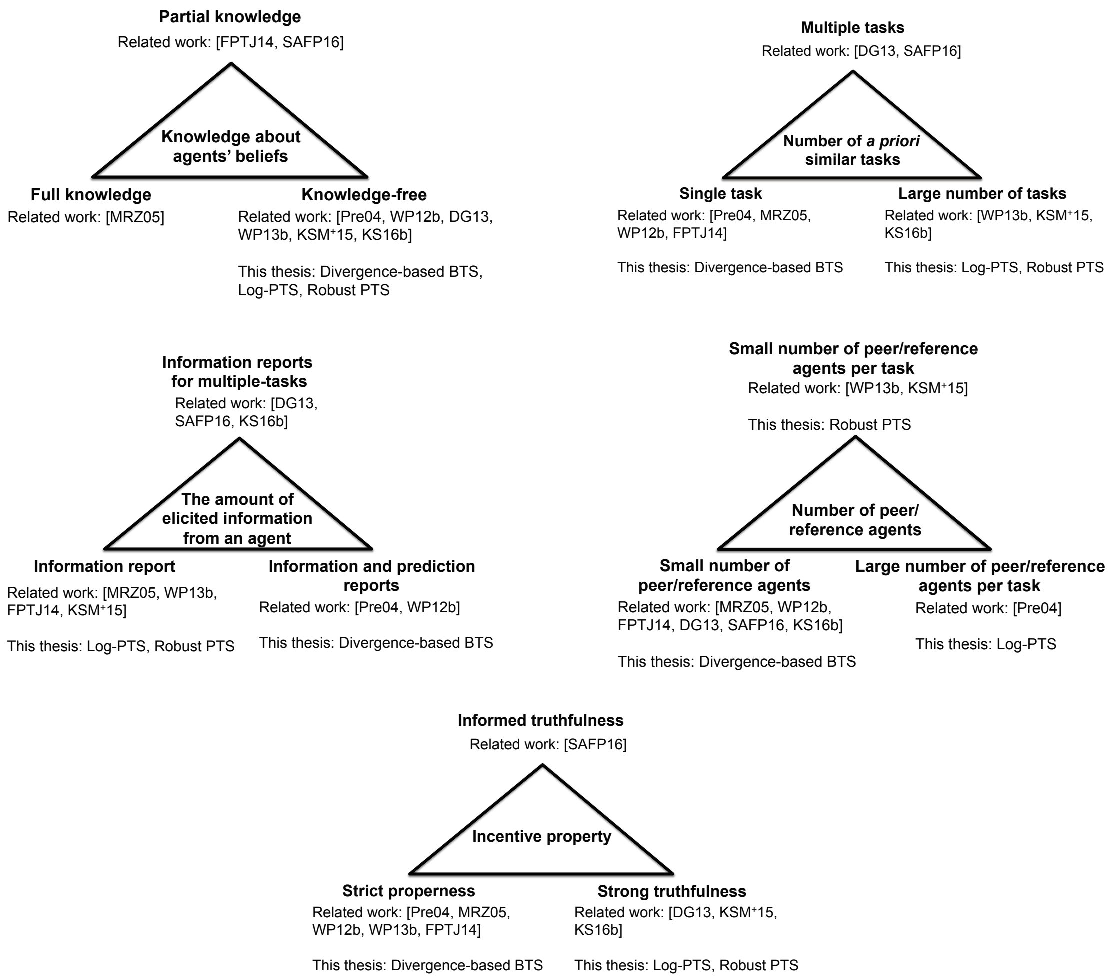


Figure 1.1 – Classifications of peer consistency mechanisms according to different criteria

Table 1.1 – Quantitative studies

Type of study	**Related work**	**This thesis**
Experimental	Crowdsourcing ([SCH11, Har11, HF13a, HF13b, FPTJ14, GMCA14]), Opinion polling ([GF14]), Surveys ([PS06, JLP12, WP13a])	Peer grading
Simulation based	Peer grading ([SP16a, SP16b]), Community sensing ([FLJ14])	Community sensing

Organization of the thesis

In Chapter 2, we start by explaining one of the basic building blocks that we use throughout the thesis: *strictly proper scoring rules*. Furthermore, Chapter 2 also contains the justification of why we need strict incentives that make truthful reporting strictly optimal strategy. Chapter 3 discusses peer consistency mechanisms designed for an elicitation scenario where a group of participants provides information about a single object. This is the classical peer consistency setting, often used for modeling acquisition of subjective information. Chapter 4 describes peer consistency mechanisms that are developed for a typical crowdsourcing scenario where a group of workers solves multiple a priori similar tasks. Naturally, these mechanisms achieve stronger incentive properties than those discussed in Chapter 3. Chapter 5 considers the issue of aggregating elicited information and discusses reputation based incentives for limiting the negative influence of participants who deliberately misreport their information. The focus is put on mechanisms designed for an online information fusion. We conclude the thesis with Chapter 6 that provides final remarks and directions for future work.

2 Preliminaries

Incentives for quality are based on functions that assign quality scores to the information content of the reported data. In this chapter, we provide a short survey of scoring functions that assign quality scores to probabilistic estimates. These are called *proper scoring rules* and they play a crucial role in the development of the results presented in the following chapters. Furthermore, we outline the importance of having strict incentives that reflect the quality of the reported data, and we show their use in motivating participants to provide accurate information, even when its acquisition is costly.

2.1 Proper scoring rules

One of the main tools for elicitation of probability distribution function is a class of mechanisms called strictly proper scoring rules. Suppose that a respondent is asked to report her probabilistic prediction F regarding an event that can be modeled as a random variable X, and whose outcome x eventually becomes known to the elicitation mechanism. For example, in a weather forecast, prediction F is a probability distribution function over possible weather conditions, while x is the realized weather condition. If the quality of F is determined by a proper scoring rule $S(F, x)$, the expected score is maximized for the optimal choice of F, i.e., the one that corresponds to the true distribution of X, denoted by $Pr(X)$. We focus on strictly proper scoring rules for which $Pr(X)$ is the unique maximizer of $\mathbb{E}_X(S(F, X))$.

There is a wide variety of strictly proper scoring rules, such as the logarithmic, quadratic, and spherical scoring rules. We describe in more details the logarithmic and quadratic scoring rules since these two are the most relevant to the elicitation mechanisms that we develop in this thesis. For an extensive overview of proper scoring rules, we refer the reader to [GR07].

2.1.1 Logarithmic scoring rule

For a random variable X that takes values in a finite discrete set and a fully mixed forecast F (i.e., $F(x) > 0$ for all values of x), the logarithmic scoring rule is defined as [Goo52]:

$$S_L(F, x) = \alpha \cdot \ln(F(x)) + \beta \tag{2.1}$$

where α and β are scaling parameters. Although the logarithmic scoring rule is not bounded, from a practical point of view this is almost never a problem. Namely, the lower bound on possible values of $F(x)$ is usually not hard to estimate, so by using scaling parameters α and β, one can easily fit scores to an arbitrary interval. For simplicity, we set $\alpha = 1$ and $\beta = 0$ in the remaining text.

To see why the logarithmic scoring rule is strictly proper, let us examine the expected value of the score for report F, when the true distribution is $Pr(X)$:

$$\begin{aligned}
\mathbb{E}_X(S_L(F, X)) &= \sum_x Pr(X = x) \cdot \ln(F(x)) = \sum_x Pr(X = x) \cdot \ln(Pr(X = x)) \\
&- \left(\sum_x Pr(X = x) \cdot \ln(Pr(X = x)) - \sum_x Pr(X = x) \cdot \ln(F(x)) \right) \\
&= \mathbb{E}_X(S_L(Pr, X)) - KL(Pr||F)
\end{aligned}$$

where $KL(Pr||F)$ is a Kullback-Leibler divergence between Pr and F. From the properties of the KL divergence (e.g., see [Bis06]), it follows that $KL(Pr||F)$ is non-negative and is equal to 0 if and only if $F = Pr$. Since the expected score depends on F only through $KL(Pr||F)$, we can conclude that it is strictly maximized for $F = Pr$.

Notice that the definition (2.1) extends to continuous domains as well. In particular, when X takes values in $\mathbb{R}$ and F is a probability density function, the logarithmic scoring rule $S_L(F, x) = \ln(F(x))$ is strictly maximized for $F = p$, where p is the true probability density of X. In that case, KL divergence between p and F is equal to $KL(p||F) = \int_{\mathbb{R}} p(x) \cdot \ln \frac{p(x)}{F(x)} dx$.

2.1.2 Quadratic scoring rule

For a random variable X that takes values in a finite discrete set, the quadratic scoring rule (or Brier score) is defined as [Bri50]:

$$S_Q(F, x) = \alpha \cdot \left(F(x) - \frac{1}{2} \cdot \sum_z F(z)^2 \right) + \beta \tag{2.2}$$

where α and β are scaling parameters, which we set to $\alpha = 1$ and $\beta = 0$ in the remaining text. In this case, the quadratic scoring rule takes values in $[-\frac{1}{2}, \frac{1}{2}]$.

As done for the logarithmic scoring rule, we can inspect the expected value of the quadratic scoring rule to obtain:

$$\begin{aligned}
\mathbb{E}_X(S_Q(F,X)) &= \sum_x Pr(X=x)\cdot\left(F(x)-\frac{1}{2}\cdot\sum_z F(z)^2\right)\\
&= \sum_x Pr(X=x)\cdot F(x)-\sum_x Pr(X=x)\cdot\frac{1}{2}\cdot\sum_z F(z)^2\\
&= \sum_x Pr(X=x)\cdot F(x)-\frac{1}{2}\cdot\sum_z F(z)^2\\
&= \sum_x Pr(X=x)\cdot F(x)-\frac{1}{2}\cdot\sum_z F(z)^2+\sum_x Pr(X=x)^2-\sum_x Pr(X=x)^2\\
&= \sum_x Pr(X=x)^2-\frac{1}{2}\cdot\sum_x Pr(X=x)^2\\
&-\frac{1}{2}\cdot\left(\sum_x Pr(X=x)^2-2\cdot\sum_x Pr(X=x)\cdot F(x)+\sum_x F(x)^2\right)\\
&= \mathbb{E}_X(S_Q(Pr,X))-\frac{1}{2}\cdot D(Pr||F)
\end{aligned}$$

where D is the squared euclidian distance between probability vectors Pr and F, and it represents the Bregman divergence associated to the quadratic scoring rule. Clearly, $D(Pr||F)$ is always positive and equal to 0 only if $F = Pr$. Therefore, the expected score is strictly maximized when $F = Pr$. The quadratic score can also be defined for a real-valued random variable X, in which case prediction F, that is represented with a probability density function, is scored with:

$$S_Q(F,x) = F(x)-\frac{1}{2}\cdot\int_{\mathbb{R}} F(z)^2 dz \tag{2.3}$$

and the associated divergence function is equal to $D(p||F) = \int_{\mathbb{R}}(p(x)-F(x))^2 dx$.

Notice that the other strictly proper scoring rules S also have the corresponding Bregman divergences. We will abuse our notation and denote a divergence of a generic strictly proper scoring rule by $D(||)$.

2.2 Strict incentive mechanisms

The simplest form of incentives assigns equal rewards to participants without inspecting the quality of reported data. The drawback of such a simple design is that it does not take into account respondents' valuation of different reports. For example, data acquisition typically requires some effort, which means that the best option for a participant is to provide uninformative reports, instead of accurate information. Another example would be when a participant has privacy concerns, so that truthful reporting is worse off for the participant as it reveals substantial amount of private information.

2.2.1 Binary participation choice

We can generally model the valuation of a participant in providing a certain information by using a cost function $c(R)$ that incorporates all aspects of revealing value R to the mechanism. We adopt the conventional model of elicitation (e.g., [DG13, SAFP16]), in which the cost function depends on the participant's binary choice on how to acquire her information or, alternatively, what type of information she will report. In particular, report R can be *informed* or *uninformed*. An informed report is based on the private information of a participant, and can be honest (the participant reveals her private information) or dishonest (the participant misreports her private information). An uninformed report is an outcome of a heuristic reporting strategy where a participant does not base her reporting decision on her private information. An example of such a report is a randomly reported value. It is reasonable to assume that the informed reports result are more costly than the uninformed reports because one does not need to acquire any data for the latter case.

2.2.2 Design goal

The goal of an elicitation mechanism can now be cast to the problem of designing a payment rule $\tau(R)$ such that the profit of a participant for reporting R, i.e., $\tau(R) - c(R)$, is maximized when the participant provides high-quality information. Notice that c and τ are functions that map reports to positive numbers that respectively represent costs and payments. Furthermore, R can also model the action of not participating, for example, by setting $R = \emptyset$, $\tau(\emptyset) = 0$ and $c(\emptyset) = 0$. This means that our design goal includes the *individual rationality* condition, which states that the participants should not engage in interaction with the system unless they expect to positively profit from it.

From the perspective of a participant, the value of $\tau(R) - c(R)$ is not known in advance, since it does not only depend on report R. In particular, payment rule $\tau(R)$ is not only dependent on R, but also on other variables whose values might not be known to the participant. For example, if a participant provides a prediction about a future event, payment rule τ could be a proper scoring rule that depends on the participant's prediction and the true outcome. A rational participant would in that case aim to maximize expectation $\mathbb{E}(\tau(R) - c(R))$ that is conditioned on her private information. Therefore, one can relax the design goals of an elicitation mechanism, and require that $\tau(R) - c(R)$ is maximized in expectation for a high-quality report.

Scaling incentives

Suppose now that there exists a certain mechanism τ_0 that rewards a high-quality report of a participant, denoted by R^{honest}, with strictly higher expected payoff than any other report R:

$$\mathbb{E}(\tau_0(R^{honest})) > \mathbb{E}(\tau_0(R)), \forall R \neq R^{honest}$$

Let us consider a scaling parameter α:

$$\alpha > \max_{R'} \frac{\max\left(0, c(R^{honest}) - c(R')\right)}{\mathbb{E}(\tau_0(R^{honest})) - \mathbb{E}(\tau_0(R'))} \tag{2.4}$$

and a meta-mechanism τ_1, such that:

$$\tau_1(R) = \alpha \cdot \tau_0(R)$$

The expected profit for reporting R^{honest} in the meta-mechanism is then equal to:

$$\begin{aligned}
&\mathbb{E}\left(\tau_1(R^{honest}) - c(R^{honest})\right) = \mathbb{E}\left(\alpha \cdot \tau_0(R^{honest}) - c(R^{honest})\right) \\
&= \mathbb{E}\left(\alpha \cdot \left(\tau_0(R^{honest}) - \tau_0(R)\right)\right) + \alpha \cdot \mathbb{E}(\tau_0(R)) - c(R^{honest}) \\
&> \max_{R'} \frac{\max\left(0, c(R^{honest}) - c(R')\right)}{\mathbb{E}(\tau_0(R^{honest})) - \mathbb{E}(\tau_0(R'))} \cdot \mathbb{E}\left(\tau_0(R^{honest}) - \tau_0(R)\right) + \alpha \cdot \mathbb{E}(\tau_0(R)) - c(R^{honest}) \\
&\geq \frac{c(R^{honest}) - c(R)}{\mathbb{E}(\tau_0(R^{honest})) - \mathbb{E}(\tau_0(R))} \cdot \left(\mathbb{E}(\tau_0(R^{honest})) - \mathbb{E}(\tau_0(R))\right) + \alpha \cdot \mathbb{E}(\tau_0(R)) - c(R^{honest}) \\
&= \alpha \cdot \mathbb{E}(\tau_0(R)) - c(R) = \mathbb{E}\left(\tau_1(R) - c(R)\right)
\end{aligned}$$

for all $R \neq R^{honest}$. The first inequality comes from (2.4) and the fact that $\mathbb{E}\left(\tau_0(R^{honest})\right) - \mathbb{E}\left(\tau_0(R)\right)$ is strictly positive, while the second inequality is due to the fact that R might not maximize the lower bound on α. Therefore, using the described scaling technique, one can convert any mechanism that satisfies the same property as τ_0 into a meta-mechanism that satisfies our objective. The technique requires an appropriate choice of α, which can be estimated empirically [DG13] or elicited from participants using an auctioning protocol [RF16a].[1] This means that we can further relax our design goal: it suffices to find a payment function $\tau(R)$ that is in expectation maximized for a good quality report.

Finally, we can also relax the condition that $\tau(R)$ should produce positive payments. Namely, if $\tau_0(R)$ is a payment rule with minimal payments equal to τ_{min}, then a payment rule $\tau_1(R) = \alpha \cdot \tau_0(R) + \beta$ satisfies the aforementioned conditions for $\beta = -\alpha \cdot \tau_{min}$ and an appropriate choice of α. τ_{min} can be determined by the theoretical lower bound of payment rule $\tau(R)$ or empirically (which is convenient if $\tau(R)$ does not have a theoretical lower bound).

We see that, in designing incentives for quality, the focus can be put on finding a payment rule τ that maps reports to real numbers and results in maximum expected payoff for high-quality reports. In the following chapters, τ is defined as a function on a set of reported values coming from different participants. This induces a game among participants, which means that a rational participant conditions her reporting strategy on what she believes about other participants. We start with a minimum requirement on a payment mechanism that a participant strictly maximizes her payoff by reporting truthfully whenever other participants are honest.

[1] [RF16b] show how to make a better separation between payments for high and low quality reports.

3 Single-task peer consistency mechanisms

In this chapter, we discuss an elicitation scenario with a single information acquisition task in which crowd participants observe a single phenomenon and report their observations. The scenario we consider represents a very general peer consistency setting where a reward mechanism does not inspect the phenomenon, but instead utilizes the correlations among reported observations to reward participants.

3.1 Formal setting

We investigate a formal setting that can be described by a group of agents that observe a certain phenomenon and report their observations to an entity called *center*. An agent a observes a signal $X_a = x_a$, updates her prior belief $Pr(X_p)$ regarding the observation of another agent p to her posterior belief $Pr(X_p|X_a = x_a)$, and reports her observation x_a to the center through an information report $Y_a = y_a$. Moreover, the center might also ask agent a to submit a prediction (forecast) F_a about the frequencies of signal values in the population, .i.e., her belief $Pr(X_p|X_a = x_a)$. In order to obtain truthful observations, the center provides agents with rewards that are calculated by comparing the reported observations (the center does not sample the phenomenon so it cannot directly verify the observations). In this chapter, we investigate mechanisms that incentivize an agent to report honestly whenever the other agents submit truthful reports. The formal setting is depicted by Figure 3.1, and we call the illustrated process *sensing* to emphasize that the agents observe (measure) the phenomenon. This abstracts an elicitation process in crowd work, in particular, how crowd participants acquire their private information and form beliefs about each other's information. Thus, a phenomenon can represent a crowdsourcing task, a real physical phenomenon, or a model of how subjective information is formed, while observations are answers to the task, measurements, or opinions, respectively.

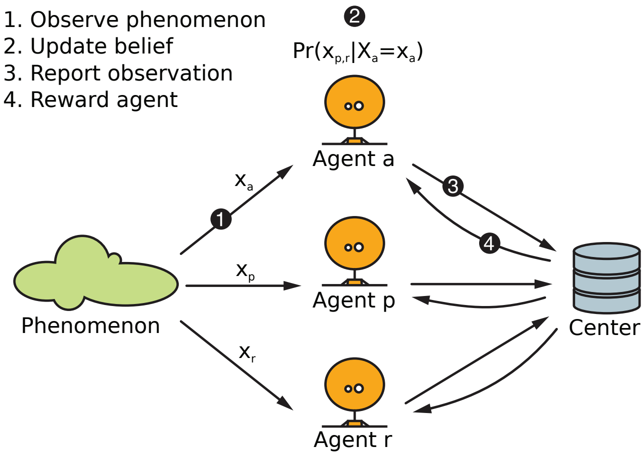


Figure 3.1 – Single-task peer consistency setting

3.1.1 Elicitation model

We consider a group of agents $\mathscr{A}$ that make observations of a certain phenomenon, and report their observations to an entity called *center*. A generic agent in $\mathscr{A}$ is denoted by a. With each agent a, we associate k peer agents from $\mathscr{A} \setminus \{a\}$ whose reports are used in assessing the quality of agent a's report. A generic peer of agent a is denoted by p. In the case when several agents a or peers p are put in the same context, we put subscripts, e.g., a_i and $p_{i,j}$. The number of agents N in group $\mathscr{A}$ is bounded from below by 2, meaning that $N = |\mathscr{A}| \geq 2$, while each agent a has $k \in \{1, ..., |\mathscr{A}| - 1\}$ peers.

The agents observe a phenomenon, and their observations are modeled as random variables X that take values in an observation set $\mathscr{X}$. The observation of each agent a, X_a, is *private* so neither the center nor the other agents know of its realization. The observation of a peer p is denoted by X_p, while the observation profile of k peers is denoted by $\mathbf{X}_p = (X_{p,1}, ..., X_{p,k})$. Generic values in $\mathscr{X}$ are represented by x, y, and z, and the set of all probability distribution functions over $\mathscr{X}$ by $\mathscr{P}$. Unless indicated differently, $\mathscr{X}$ is a finite discrete set, and, for simplicity, we describe the setting by assuming that this indeed holds.

An agent a has a probabilistic belief about how agents acquire their observations. The agent's belief system B_a is not known to the center and consists of:

- *prior* belief regarding her own observation $Pr(X_a) \in \mathscr{P}$;
- *prior* belief regarding peers' observations $Pr(\mathbf{X}_p) \in \mathscr{P}^k$, and similarly for any subset of peers, e.g., for a single peer, we denote $Pr(X_p) \in \mathscr{P}$;
- *posterior* belief regarding peers' observations $Pr(\mathbf{X}_p|X_a) \in \mathscr{P}^k$, and similarly for any subset of peers, e.g., for a single peer, we denote $Pr(X_p|X_a) \in \mathscr{P}$.

We assume that two basic conditions are satisfied in B_a:

- all the probability distribution functions in B_a are *fully mixed*, meaning that they assign strictly positive probabilities to each outcome;
- posterior $Pr(X_p|X_a)$ is *stochastically relevant*, meaning that $D(Pr(X_p|X_a = x)||Pr(X_p|X_a = y)) > 0$ for $\forall x \neq y$, where $D(||)$ is the Bregman divergence of a strictly proper scoring rule.

The second condition indicates that the posterior distribution regarding X_p conditioned on X_a is different for different realizations of X_a [MRZ05]. The belief profile of group of agents $\mathscr{A}$ is denoted by $\mathbf{B}_{\mathscr{A}} = (B_{a_1}, ..., B_{a_N})$, and when it is needed to distinguish between the beliefs of different agents a_i, we put subscript Pr_{a_i} (if there is no subscript, Pr denotes agent a's beliefs).

Once she observes the phenomenon, agent a is asked to submit a report R_a that contains:

- *information report* $Y_a \in \mathscr{X}$, which represents agent a's reported observation;
- additionally, she might be asked to submit a *prediction report* $F_a \in \mathscr{P}$, which represents agent a's prediction regarding the frequencies of signal values in the overall population. When agents are honest, this report corresponds to agent a's posterior belief $Pr(X_p|X_a)$.

Therefore, the structure of report R_a, denoted by $\mathscr{R}$, can be of the form $\mathscr{R} = \mathscr{X}$ or $\mathscr{R} = \mathscr{X} \times \mathscr{P}$, indicating that agent a's report is $R_a = Y_a$ or $R_a = (Y_a, F_a)$, respectively. For an agent a's peer, the notation is Y_p, F_p and R_p, and the report profiles of agent a's peers are denoted by $\mathbf{Y}_p = (Y_{p,1}, ..., Y_{p,k})$, $\mathbf{F}_p = (F_{p,1}, ..., F_{p,k})$ and $\mathbf{R}_p = (R_{p,1}, ..., R_{p,k})$.

The center rewards the agents based on the quality of the information they provide, and the quality is estimated by comparing their reports. That is, a payment mechanism τ does not only depend on the report of the agent that is being rewarded, but also on the reports of other agents, her peers. In this chapter, we investigate *single-task* payment mechanisms τ_S, which are formally defined as $\tau_S : \times_{i=0}^{k} \mathscr{R} \to \mathbb{R}$. Depending on the information structure that the agents report, the payoff of an agent a is either a function of information reports alone, i.e., $\tau_S(Y_a, Y_{p_1}, ..., Y_{p_k})$ for $\mathscr{R} = \mathscr{X}$, or a function of information and prediction reports, i.e., $\tau_S(Y_a, F_a, Y_{p_1}, F_{p_1}, ..., Y_{p_k}, F_{p_k})$ for $\mathscr{R} = \mathscr{X} \times \mathscr{P}$.

Finally, each agent is assumed to be rational and risk-neutral, and her utility is reflected through the reward given by the mechanism. Thus, an agent's objective is to maximize the payments provided by the center through mechanism τ_S. This further implies that an agent is inclined to strategize on what to report to the center, which means that payment mechanism τ_S should incentivize truthful revelation of private information, as further discussed in Section 3.1.6.

3.1.2 Beliefs about peers

The mechanisms we are investigating are based on comparing reports that different agents make about the same phenomenon. Clearly, there are cases in which such comparisons do not make sense. For example, if agents all interpret the phenomenon differently, or use different scales for measurement, their reports cannot be compared directly. Furthermore, what matters is not the true situation, but what agents believe about their peers: to provide the right incentives, it is sufficient that they believe their peers to be comparable to themselves, even if in reality that might not be the case.

Therefore, we introduce a notion of belief constraint set $\mathscr{C}$, whose elements are conditions C that specify:

- how belief systems of different agents relate,
- how observations are acquired,
- how priors are updated to posteriors.

The set of admissible belief profiles under constraint set $\mathscr{C}$ is denoted by $\mathscr{B}(\mathscr{C})$, and it contains belief profiles $\mathbf{B}_a$ that satisfy conditions in $\mathscr{C}$. In the following subsections, we define belief conditions C important for developing the formal results of this chapter.

3.1.3 Relational constraints

Relational constraints describe how agents' belief systems relate to each other, that is, how much agents' belief systems differ from each other. We identify two conditions: *common belief condition* and *divergence-based condition*.

Common belief condition

The simplest constraint and arguably the most stringent condition is the one that states that agents share a common belief system. We will, therefore, use this condition as a baseline in exploring impossibility results. In particular, we would like to at least be able to design a mechanism that elicits truthful observations for the case when all of the agents share the same belief system.

Definition 1. *A belief profile* $\mathbf{B}_{\mathscr{A}}$ *satisfies the* common belief *condition* C_{CB} *if* $\forall a_1, a_2 \in \mathscr{A}$: $B_{a_1} = B_{a_2}$.

Divergence-based condition

A particular property of the common belief condition C_{CB} is that agents who have the same observations should also have the same posterior belief about their peer agents. A possible relaxation of this condition would be to say that two agents should have more similar posterior beliefs when they have equal observations than when their observations are different. More formally:

Definition 2. *A belief profile* $\mathbf{B}_{\mathscr{A}}$ *satisfies the* divergence-based *condition* C_{DB} *if there exists* $\Theta \geq 0$ *such that* $\forall a_1, a_2 \in \mathscr{A} \wedge \forall x \neq y$*:*

$$D(Pr_{a_1}(X_{p_1}|X_{a_1} = x)||Pr_{a_2}(X_{p_2}|X_{a_2} = x)) \leq \Theta < D(Pr_{a_1}(X_{p_1}|X_{a_1} = x)||Pr_{a_2}(X_{p_2}|X_{a_2} = y))$$

where D(||) is the Bregman divergence of a strictly proper scoring rule.

3.1.4 Acquisitional constraints

Acquisitional constraints model belief assumptions on how agents reason about the acquisition of their private information, in particular, their observations.

State model condition

We define an acsquisitional constraint based on a *state* model that is similar to the ones introduced in [Pre04, MRZ05, WP12b].

Definition 3. *Consider a random variable* Ω *taking values in* $\mathbb{R}$. *A belief profile* $\mathbf{B}_{\mathscr{A}}$ *satisfies the* state model *condition* C_{SM} *if each belief system* B_a *is constrained with the following set of assumptions:*

- *observations* X_{a_1} *and* X_{a_2} *of any two different agents* a_1 *and* a_2 *in* $\mathscr{A}$ *are conditionally independent given* Ω*;*
- *agent a's prior belief regarding* Ω *is a probability density function* $p_a(\Omega)$ *that takes strictly positive values;*
- *for all agents* a_i *in* $\mathscr{A}$*, probabilities* $Pr_a(X_{a_i}|\Omega)$ *are strictly positive.*

Notice that $Pr_a(X_{a_i}|\Omega)$ and $Pr_{a'}(X_{a_i}|\Omega)$ (and similarly $p_a(\Omega)$ and $p_{a'}(\Omega)$) are allowed to be different for two different agents a and a'. However, we often drop the subscript a from $Pr_a(X_{a_i}|\Omega)$ (and $p_a(\Omega)$) to simplify the notation.

Gaussian state model condition

We also consider a refinement of the *state model* condition C_{SM} that specifies the probability distribution functions of agents' belief systems. We are particularly interested in Gaussian state model condition C_{GSM} that is defined for observations that take values in $\mathbb{R}$ and for which the parameters of agents' belief systems, i.e., probability density functions, are Gaussian distributions.

Definition 4. *Consider observation space* $\mathscr{X} = \mathbb{R}$. *A belief profile* $\mathbf{B}_{\mathscr{A}}$ *satisfies the* Gaussian state model *condition* C_{GSM} *if it satisfies the state model condition* C_{SM} *in which:*

- *the observation values* X_a *are generated by a Gaussian distribution* $p_a(X_a) \sim \mathcal{N}(\mu_\Omega, \sigma)$, *where* σ *is fixed (given);*
- μ_Ω *defines state* Ω *and is distributed according to a Gaussian distribution* $p_a(\mu_\Omega) \sim \mathcal{N}(\mu_0, \sigma_0)$.

As it is the case for the state model condition C_{SM}, we often drop the subscript a from $p_a(X_a)$ and $p_a(\mu_\Omega)$ when it is clear that we are referring to agent a.

3.1.5 Updating constraints

The updating constraints describe conditions about the strength of agents' beliefs. The more confident an agent is in her observation, the more likely it is (by her beliefs) that her peer observes the same value.

Self-dominant condition

The strongest condition we consider is the *self-dominant* condition, where an agent a believes that the value x she observes is also the most likely value observed by her peer p.

Definition 5. *A belief profile* $\mathbf{B}_{\mathscr{A}}$ *satisfies the* self-dominant condition C_{SD} *if for the posterior belief of each* B_a*:*

$$Pr(X_p = y | X_a = x) < Pr(X_p = x | X_a = x), \forall y \neq x \tag{3.1}$$

A general class of cases where the self-dominant condition holds is when agents believe that they observe the exact same signal only perturbed by an unbiased noise.

Self-predicting condition

As many settings do not satisfy this condition, we introduce a weaker condition, called the *self-predicting* condition. Here, an agent a believes that she is most likely to observe a certain value x when another agent p observes the same value.

Definition 6. *A belief profile* $\mathbf{B}_{\mathscr{A}}$ *satisfies the* self-predicting condition C_{SP} *if for the posterior belief of each* B_a*:*

$$Pr(X_a = x | X_p = y) < Pr(X_a = x | X_p = x), \forall y \neq x \tag{3.2}$$

By applying Bayes' rule, we obtain an alternative form of the self predicting condition:

$$\frac{Pr(X_p = y|X_a = x)}{Pr(X_p = y)} < \frac{Pr(X_p = x|X_a = x)}{Pr(X_p = x)}, \forall y \neq x \tag{3.3}$$

This form is important for the next chapter, where we redefine the self-predicting condition for a multi-task setting.

A general class of cases where the self-predicting condition holds is when agents believe that they observe different samples drawn from the same random distribution, but with the condition that these samples are *categorical* [SAFP16], so that observing value x reduces an agent's belief that her peer observes another value $y \neq x$ [JF11]. Notice, however, that unlike the described categorical case, the self-predicting condition allows (limited) correlation between different observation values x and y.

Self-correlated condition

Arguably, the weakest condition that we consider is a self-correlated condition, which states that an agent a should believe that observing a certain value x only increases the chances that her peer observes the same value. More formally:

Definition 7. *A belief profile $\mathbf{B}_{\mathscr{A}}$ satisfies the* self-correlated condition C_{SC} *if for the posterior belief of each B_a:*

$$Pr(X_p = x|X_a = x) > Pr(X_p = x), \forall x \in \mathscr{X} \tag{3.4}$$

The self-correlated condition holds whenever agents observe the same phenomenon in a similar way. In particular, it holds in the *state* models of standard peer consistency methods [Pre04, MRZ05, WP12b].

Relative self-dominant condition

Finally, we consider a condition that is a combination of relational and updating constraints. It states that the difference between $Pr_a(X_p = x|X_a = x)$ and $Pr_a(X_p = y|X_a = x)$ of an agent a, be it greater or smaller than 0, is always strictly greater than the expected difference between $Pr_{a_j}(X_{p_j} = x|X_{a_j})$ and $Pr_{a_j}(X_{p_j} = y|X_{a_j})$ of another agent a_j. Notice that this condition is similar to the self-dominant condition, but it only requires that self-dominance is satisfied relative to the beliefs of other agents. Therefore, we call it *relative self-dominance*.

Definition 8. *A belief profile $\mathbf{B}_{\mathscr{A}}$ satisfies the* relative self-dominant condition C_{RSD} *if for the posterior beliefs any two agents, a_1 and a_2, we have that:*

$$\begin{aligned} &Pr_{a_1}(X_{p_1} = x|X_{a_1} = x) - Pr_{a_1}(X_{p_1} = y|X_{a_1} = x) \\ &\qquad > \mathbb{E}_{X_{a_2}}\left(Pr_{a_2}(X_{p_2} = x|X_{a_2}) - Pr_{a_2}(X_{p_2} = y|X_{a_2})|X_{a_1} = x\right) \end{aligned}$$

for all $x \neq y$.

3.1.6 Reporting strategies

As a rational agent, agent a aims to maximize the reward obtained from the center, and in the case of uncertainties, she is assumed to maximize her *expected* reward. To decide which information to provide to the center, agent a should estimate the rewards expected for different reports. Notice that she has no knowledge about what her peers have reported, so her best reporting strategy depends crucially on the agent's beliefs about the reports of her peers.

In this chapter, we assume that agent a *believes* her peers are honest. Given that, agent a faces a choice between two basic strategies:

- *honest*: observe the phenomenon to obtain evaluation $X_a = x$, and report honestly $Y_a = x$ (and $F_a = Pr(X_p|X_a = x)$). We denote honest report by R_a^{honest}.
- *dishonest*: observe the phenomenon to obtain evaluation $X_a = x$, but report $R_a^{dishonest} \neq R_a^{honest}$.

To incentivize truthful revelation, payment mechanism τ_S should satisfy the property called *strict properness* (e.g., see [SAFP16]), which we define for the setting of this chapter as follows.

Definition 9. *We say that a* singe-task *payment mechanism* τ_S *is* proper *under the set of belief constraints* $\mathscr{C}$ *if for all* $\mathbf{B}_{\mathscr{A}} \in \mathscr{B}(\mathscr{C})$*,* $a \in \mathscr{A}$*, and* $R_a \in \mathscr{R} \setminus \{R_a^{honest}\}$*, we have that:*

$$\mathbb{E}_{\mathbf{X}_p}\left(\tau_S(R_a^{honest}, \mathbf{R}_p^{honest})|X_a\right) \geq \mathbb{E}_{\mathbf{X}_p}\left(\tau_S(R_a, \mathbf{R}_p^{honest})|X_a\right) \tag{3.5}$$

If the inequality is strict, then τ_S *is* strictly *proper.*

The definition of strict properness states that truthful reporting is a strict equilibrium of mechanism τ_S. In particular, when the set of belief constraints consists of the common belief condition and the state model condition (i.e., $\mathscr{C} = \{C_{CB}, C_{SM}\}$), mechanism τ_S has a form of a Bayesian game (e.g., see [SLB08]), and the strict properness property implies that truthful reporting is a strict *Bayesian Nash equilibrium*, irrespective of the parameters of agents' common belief system.[1] In general, agents might not have a common belief system, in which case the strict properness property implies that truthful reporting is a strict *ex-post subjective equilibrium* (see [WP12a]), with the admissible set of belief profiles defined by $\mathscr{B}(\mathscr{C})$.[2] Notice that the concept of Bayesian Nash equilibrium is a special case of the ex-post subjective equilibrium concept, obtained for the set of belief constraints equal to $\mathscr{C} = \{C_{CB}, C_{SM}\}$.

[1] As argued by [Wit14, SAFP16], one can adopt a correlated equilibrium concept instead.

[2] The (original) definition of the ex-post subjective equilibrium concept is based on admissible *belief types* that categorize agents' belief systems [WP12a], but using a reasoning similar to [FW16], one can define it via $\mathscr{B}(\mathscr{C})$.

One might wonder why the strategy space of an agent does not include a possibility of the agent not inspecting the considered phenomenon. The following proposition shows, however, that we can, without loss of generality, constrain the strategy space of the agents to *honest* and *dishonest* strategies. Namely, strategies in which an agent does not observe the phenomenon do not represent the best response to the honest behaviour of the other agents whenever the considered mechanism is strictly proper. On the other hand, to establish strict properness, one only needs *honest* and *dishonest* reporting strategies.

Proposition 1. *Consider an agent a that has not yet made an observation and suppose her peers are honest ($\mathbf{R}_p = \mathbf{R}_p^{honest}$). For any strictly proper payment mechanism τ_S, agent a's payoff is expected to strictly increase if she decides to adopt honest reporting strategy ($R_a = R_a^{honest}$) instead of a reporting strategy in which she does not make an observation.*

Proof. Since mechanism τ_S is strictly proper, we know that the expected payoff of agent a when she observes the phenomenon is maximized for report R_a^{honest}. Consider now the situation before the agent observes the phenomenon. The expected payoff for reporting R_a is equal to:

$$\begin{aligned}\mathbb{E}_{\mathbf{X}_p}\left(\tau_S(R_a, \mathbf{R}_p^{honest})\right) &= \sum_{x \in \mathcal{X}} Pr(X_a = x) \cdot \mathbb{E}_{\mathbf{X}_p}\left(\tau_S(R_a, \mathbf{R}_p^{honest}) | X_a = x\right) \\ &< \sum_{x \in \mathcal{X}} Pr(X_a = x) \cdot \mathbb{E}_{\mathbf{X}_p}\left(\tau_S(R_a^{honest}, \mathbf{R}_p^{honest}) | X_a = x\right)\end{aligned}$$

where the inequality comes from the fact that: τ_S is strictly proper, the best response (R_a^{honest}) is dependent on X_a, and $Pr(X_a)$ is fully mixed. Since the right most part in the expression is the expected payoff for honest reporting (calculated prior to the observation), we obtain that agent a is expected to strictly increase her payoff when observing the phenomenon. □

Remark 1. In this chapter, we do not analyze collusion properties of the developed mechanisms. As it turns out, even achieving strict properness is not trivial for the single-task elicitation setting. We note that there are many practical scenarios where truthfulness, if an equilibrium strategy profile, is a *focal point*, since other reporting strategy profiles might require more *unnatural* coordination among agents. These include opinion polling or human computation tasks in which workers do not frequently interact with each other. For example, the study of [FPTJ14] did not appear to have a problem with collusion. On the other hand, a susceptibility to collusion might be problematic if agents repeatedly and frequently interact with a peer consistency mechanism, as experimentally shown in [GMCA14] and further argued in [SP16a]. We refer the reader to [KS16a] on how to extend some of the results of this chapter to be more robust to collusive behaviour.

3.2 Related work

Gold standard mechanisms are the simplest design of incentives for quality. They assume that the center has access to gold standards, such as test tasks in crowdsourcing, and use these to score agents based on how accurate their reports are [OSL+11, Har11]. The more complex designs of gold standard mechanisms allow agents to also express confidence in their answers [SZP15, SZ15] or are suitable for the elicitation of accurate aggregates [FCK15, UDG15]. In this type of mechanisms, we can also include proper scoring rules [Sav71, GR07], prediction markets [Han07, CP07], and scoring rules for elicitation of averages, medians and modes of an unknown quantities [LS09].

As our setting precludes the center from having the access to the gold standard, we focus on *peer consistency* techniques, which are based on the comparison of reports. One of the most basic peer consistency mechanisms is the output agreement [vAD04, vAD08] that rewards agents if their reports agree. In a more general sense, agents can be rewarded by how close their reports are, measured by a predefined distance function, and in this case, the output agreement is shown to elicit common knowledge rather than agents' private information [WC13, WC14]. We show in this chapter a condition under which the mechanism is strictly proper.

In order to deal with a potential bias towards prior information, the peer prediction method [MRZ05] scores agents' posterior beliefs for a reported value. The main idea behind the mechanism is to extract an agent's posterior belief from her reported value and score it using a proper scoring rule and a report obtained from her peer agent. Due to the reliance on strictly proper scoring rules, the mechanism is strictly proper. However, it assumes that agents have a common belief system, known to the mechanism. Several modifications of the peer prediction method were investigated in the literature. Instead of applying proper scoring rules, [JF06a, JF06b] construct budget minimizing payment schemes using automated mechanism design. They prove that if an agent is scored on the comparison of several reports rather than just one, the minimum budget required to achieve incentive compatibility decreases. The results also indicate that small deviations of agents' beliefs from the common belief system may lead to large increases in payments. [JF09] and [KSL16] investigate how to make collusive strategies less profitable in the peer prediction framework.

The collective revelation [GRP09] elicits individual predictions and aggregate estimates. It has a setting similar to the peer prediction mechanisms, with the common belief system known to the mechanism, and agents that may make multiple observations, generated from a distribution of a particular form (e.g., Bernoulli distribution).

In the group of knowledge-dependent mechanisms, we can also include mechanisms that require knowledge regarding agents' prior belief, but instead of assuming common belief system, they put conditions on how agents update their beliefs. These include the shadowing method [Wit14][3] and the peer truth serum [JF11, FPTJ14]. A full overview of these mechanisms can be found in [FW16], where it is shown how these mechanisms relate to the partitioning of the probability simplex $\mathscr{P}$.

[3] We refer to the shadowing method as the one in which the shadowing approach transforms a known prior to a posterior using an agent's report.

Weakly truthful mechanisms do not necessarily provide strict incentives [LS08] or may not necessarily be proper [JF08, JF11]. The latter mechanism is the peer truth serum that rewards agents using an estimate of the agents' prior, as opposed to the exact one. The estimate is obtained with a public statistic. If the statistic is close to the prior, the mechanism is strictly proper, otherwise, it is helpful in a sense that it drives the statistics toward the agents' prior.

Except for the output agreement, none of the mentioned mechanisms fits in the setting of this chapter, either because they require knowledge about agents' beliefs or because they are not strictly proper. The mechanisms that have a formal setting closest to ours are called Bayesian truth serums.

The (original) Bayesian truth serum (BTS) [Pre04] assumes a setting similar to the one used in the peer prediction method, but does not require a common belief system to be known to the mechanism. Instead, agents are obliged to provide two reports: the information report (their observation) and the prediction report (the prediction of what other agents have observed). BTS is strictly proper for large populations of agents. The robust Bayesian truth serum (RBTS) of [WP12b] corrects the main drawback of BTS: its inadequacy to operate on small populations. RBTS is strictly proper for small populations, but requires additional belief constraints when observations are non-binary [Wit14]. The minimum truth serums of [Ril14] aim to minimize the number of reported values in BTS type of mechanisms by exploiting the common structure of agents' beliefs. The mechanism requires that the number of agents is at least as large as the observation space. [KS16a] builds on our work, presented in the following sections, to improve the properties of the BTS design in terms of collusion resistance. Finally, four interesting empirical results relate to the BTS mechanisms: [PS06] describe how to use the BTS mechanism in order to obtain the ground truth even when the majority is wrong, while [SCH11, WP13a, JLP12] demonstrate that the BTS mechanism rewards truthful responses and has a positive effect in quality control. We study the application of a BTS type of mechanism to a small scale community sensing setting, and provide a complementary approach to the existing knowledge-dependent elicitation mechanisms proposed for community sensing [FLJ14]. Furthermore, we provide an application based BTS result that relates to the design of contests that optimize user involvement [GR14].

For completeness, we also mention the mechanisms which operate in a setting that separates time before the observations are made from the time after the observations are made. In this case, it is possible to exploit the temporal structure to elicit binary observations even when the beliefs are private and subjective [WP12a]. The key idea is that the agents first report their private prior belief about what the other agents will observe, then observe a binary signal, and after the observation report their signal values. In this group, we can also include the knowledge-free mechanism of [ZC14], which establishes temporal separation by creating two step reporting protocol. The mechanism relies on a common belief system and asks agents to first submit their information report, after which they report their prediction report knowing the information report of another agent (this information is revealed to them).

3.3 Single-report mechanisms

We begin our analysis with the mechanisms that elicit only the observed values, meaning that reports take values in $\mathscr{R} = \mathscr{X}$. In other words, an agent a is asked to provide only her *information report* Y_a.

Since the center has no knowledge of the agents' beliefs, we can expect that there does not exist a strictly proper mechanism under arbitrary belief constraints. However, we show in this section that this is true even if the agents' beliefs are constrained to be common or satisfy the self-predicting condition, i.e., $\mathscr{C} \subseteq \{C_{SB}, C_{SP}\}$.

Before formalizing the statement, let us take a closer look at single peer payment mechanisms.

Definition 10. *A payment mechanism τ_S is 1-peer if it can be written as a function $\tau_S : \mathscr{R} \times \mathscr{R} \to \mathbb{R}$ of an agent a's report and the report of one of her peers, i.e., $\tau_S(R_a, R_p)$.*

As it is stated in the setting section, we consider scoring functions that depend on the reports of k peers, which in the case of a single report have form $\tau_S(Y_a, Y_{p,1}, ..., Y_{p,k})$. On the other hand, 1-peer payment mechanisms represent a restricted version of general scoring function as they only consider an agent having one peer. In the case of a single report, 1-peer payment mechanisms have form $\tau(Y_a, Y_p)$. Nevertheless, their structure simplifies the theoretical analysis while keeping the obtained results general enough.

The following lemma shows that for proving impossibility results, it suffices to examine 1-peer payment mechanisms: if strict properness is required for a relatively general, yet constrained enough belief profile $\mathbf{B}_a$ that satisfies the common belief condition C_{CB} or the self prediction condition C_{SP}, it is enough to consider 1-peer payment mechanisms. [4]

Lemma 1. *Suppose that agents provide only their information report, i.e., $\mathscr{R} = \mathscr{X}$. If there exists a strictly proper payment mechanism under the set of belief constraints $\mathscr{C} \subseteq \{C_{CB}, C_{SP}\}$, then there exists a 1-peer payment mechanism that is strictly proper under the same belief constraint.*

Proof. Let τ_S be a strictly proper payment mechanism. If $k = 1$ (agent a has only one peer), the statement follows immediately. Let us now consider $k > 1$.

Provided that all her peer agents are honest, the expected score of an agent a who observes x for reporting y is equal to:

$$\sum_{x_1,...,x_k} Pr(X_{p_1} = x_1, ..., X_{p_k} = x_k | X_a = x) \cdot \tau_S(y, x_1, ..., x_k)$$
$$= \sum_{x_1} (Pr(X_{p_1} = x_1 | X_a = x) \cdot$$
$$\cdot \sum_{x_2,...,x_k} Pr(X_{p_2} = x_2, ..., X_{p_k} = x_k | X_a = x, X_{p_1} = x_1) \cdot \tau_S(y, x_1, ..., x_k))$$

[4] However, Lemma 1 does not imply that it is enough to consider 1-peer payment mechanisms in order to achieve strict properness when additional restrictions are put on agents' belief systems.

$$= \sum_{x_1} Pr(X_{p_1} = x_1 | X_a = x) \cdot \tilde{\tau}(y, x_1)$$

where we put $\tilde{\tau}(y, x_1) = \sum_{x_2,...,x_k} Pr(X_{p_2} = x_2, ..., X_{p_k} = x_k | X_a = x, X_{p_1} = x_1) \cdot \tau_S(y, x_1, ..., x_k))$. Notice that $\tilde{\tau}(y, x_1)$ depends on $Pr(X_{p_2} = x_2, ..., X_{p_k} = x_k | X_a = x, X_{p_1} = x_1)$. However, the original mechanism is strictly proper under any belief profile $\mathbf{B}_a$ that satisfies $\mathscr{C} \subseteq \{C_{CB}, C_{SP}\}$, so it must be strictly proper when the updating process keeps $Pr(X_{p_2} = x_2, ..., X_{p_k} = x_k | X_a = x, X_{p_1} = x_1)$ fixed, but alters $Pr(X_{p_1} = x_1 | X_a = x)$. This implies the existence of a 1-peer payment mechanism that is strictly proper under $\mathscr{C} \subseteq \{C_{CB}, C_{SP}\}$ because $\tilde{\tau}$ is strictly proper for arbitrary beliefs $Pr(X_{p_1} = x_1 | X_a = x)$ that satisfy $\mathscr{C} \subseteq \{C_{CB}, C_{SP}\}$. □

Using the lemma, we now show the main result of this section: even under a relatively constrained beliefs, in particular, agents having a common belief system, no single-report peer consistency technique is strictly proper. Notice that the setting of this chapter assumes that the center has no knowledge about the agents' belief systems. Clearly, the knowledge regarding the structure of the agents' belief systems plays a crucial role in the elicitation process. Thus, the knowledge-free elicitation comes at a certain price, either through the structure of elicited information or the restrictions imposed on agents' belief systems.

Theorem 1. *Suppose that agents report only their information report, i.e. $\mathscr{R} = \mathscr{X}$. There exists no strictly proper payment mechanism under the set of belief constraints $\mathscr{C} \subseteq \{C_{CB}, C_{SP}\}$.*

Proof. Let us assume that there exists a strictly proper payment mechanism τ_S for $\mathscr{C} \subseteq \{C_{CB}, C_{SP}\}$. Due to Lemma 1, we restrict our attention to 1-peer payment schemes. Let agents' have a common belief system, with arbitrary beliefs denoted by $\mathbf{p}_x = Pr(X_p | X_a = x)$, $\mathbf{p}_y = Pr(X_p | X_a = y)$, $\mathbf{p}_z = Pr(X_p | X_a = z)$, etc., and let these beliefs satisfy the self-predicting condition. In particular, we set $\mathbf{p}_y(y) > \mathbf{p}_x(y) > \mathbf{p}_x(x) > \mathbf{p}_y(x)$ and $\mathbf{p}_z(z) > \mathbf{p}_{z'}(z)$ for all $z, z' \in \mathscr{X}, z \neq z'$. Due to the strict properness, we know that the expected payoff for reporting x when x is observed should be strictly greater than the expected payoff for reporting some other value. Similarly we obtain for another value y. The posterior belief when x is observed is equal to $\mathbf{p}_x$, while for observation y is equal to $\mathbf{p}_y$. Therefore, the strict properness implies:

$$\sum_{z \in \mathscr{X}} \mathbf{p}_x(z) \cdot \tau_S(x, z) > \sum_{z \in \mathscr{X}} \mathbf{p}_x(z) \cdot \tau_S(y, z)$$

$$\sum_{z \in \mathscr{X}} \mathbf{p}_y(z) \cdot \tau_S(y, z) > \sum_{z \in \mathscr{X}} \mathbf{p}_y(z) \cdot \tau_S(x, z)$$

which can be rearranged to:

$$\mathbf{p}_x(x) \cdot [\tau_S(x, x) - \tau_S(y, x)] + \mathbf{p}_x(y) \cdot [\tau_S(x, y) - \tau_S(y, y)] + \\ + \sum_{z \in \mathscr{X} \setminus \{x, y\}} \mathbf{p}_x(z) \cdot [\tau_S(x, z) - \tau_S(y, z)] > 0$$

$$\mathbf{p}_y(x)\cdot[\tau_S(y,x)-\tau_S(x,x)]+\mathbf{p}_y(y)\cdot[\tau_S(y,y)-\tau_S(x,y)]+$$
$$+\sum_{z\in\mathscr{X}\setminus\{x,y\}}\mathbf{p}_y(z)\cdot[\tau_S(y,z)-\tau_S(x,z)]>0$$

To simplify the notation, let $\Delta_z = \tau_S(x,z) - \tau_S(y,z)$. The above expressions are then equal to:

$$\mathbf{p}_x(x)\cdot\Delta_x+\mathbf{p}_x(y)\cdot\Delta_y+\sum_{z\in\mathscr{X}\setminus\{x,y\}}\mathbf{p}_x(z)\cdot\Delta_z>0$$

$$-\mathbf{p}_y(x)\cdot\Delta_x-\mathbf{p}_y(y)\cdot\Delta_y-\sum_{z\in\mathscr{X}\setminus\{x,y\}}\mathbf{p}_y(z)\cdot\Delta_z>0$$

Let us now consider a new set of beliefs $\mathbf{p}'_x$, $\mathbf{p}'_y$, $\mathbf{p}_z$, etc., where: $\mathbf{p}'_x(x)=\mathbf{p}_y(y)$, $\mathbf{p}'_x(y)=\mathbf{p}_y(x)$, $\mathbf{p}'_y(x)=\mathbf{p}_x(y)$, $\mathbf{p}'_y(y)=\mathbf{p}_x(x)$, $\mathbf{p}'_x(z)=\mathbf{p}_y(z)$ and $\mathbf{p}'_y(z)=\mathbf{p}_x(z)$ for $z\neq x,y$. Notice that the new set of beliefs satisfy the self-predicting condition. Since the incentive compatibility also has to hold for the new posterior beliefs, we have:

$$\mathbf{p}_y(y)\cdot\Delta_x+\mathbf{p}_y(x)\cdot\Delta_y+\sum_{z\in\mathscr{X}\setminus\{x,y\}}\mathbf{p}_y(z)\cdot\Delta_z>0$$
$$-\mathbf{p}_x(y)\cdot\Delta_x-\mathbf{p}_x(x)\cdot\Delta_y-\sum_{z\in\mathscr{X}\setminus\{x,y\}}\mathbf{p}_x(z)\cdot\Delta_z>0$$

The last 4 inequalities give us:

$$(\mathbf{p}_x(x)-\mathbf{p}_x(y))\cdot(\Delta_x-\Delta_y)>0 \tag{3.6}$$
$$(\mathbf{p}_y(y)-\mathbf{p}_y(x))\cdot(\Delta_x-\Delta_y)>0 \tag{3.7}$$

Because we set $\mathbf{p}_y(y)>\mathbf{p}_x(y)>\mathbf{p}_x(x)>\mathbf{p}_y(x)$, it cannot be that both (3.6) and (3.7) are satisfied. That is, we have a contradiction. □

The significance of Theorem 1 is that it motivates the use of mechanisms with an additional report. Namely, as we show in the next sections, under the same set of belief constraints, there exists a mechanism that requires an additional report but is strictly proper.

3.3.1 Output agreement

An alternative approach would be to examine stricter conditions regarding agents' belief systems. It suffices to use the self-dominant condition C_{SD} in order to allow strict properness among minimal knowledge-free mechanisms. This leads us to a very well known mechanism called the output agreement (OA) [vAD04].

Output Agreement. Consider an agent a and her peer p. The output agreement mechanism rewards an agent a with 1 if her report matches the report of her peer. Otherwise, the reward of agent a is equal to 0. That is, for reporting y, the agent obtains reward:

$$\tau_S(y, Y_p) = \mathbb{1}_{Y_p = y} \tag{3.8}$$

where $\mathbb{1}$ is an indicator variable.

Proposition 2. *The output agreement mechanism is strictly proper under the self-dominant condition, i.e.,* $\mathscr{C} = \{C_{SD}\}$.

Proof. Consider an agent a that observes $X_a = x$ and suppose her peer agent p is honest. If agent a aims to maximize her expected payoff in the OA mechanism, she will choose to report:

$$\operatorname*{argmax}_z \mathbb{E}(\tau_S(z, X_p)) = \operatorname*{argmax}_z Pr(X_p = z | X_a = x) = \{x\}$$

where the last equality is due to the self-dominant condition C_{SD}. Hence, we proved the statement. □

3.4 Multi-report mechanisms

We now turn to mechanisms that ask an agent to additionally provide her prediction report F, which represents her posterior belief about the reports of her peers. This means that the set of possible reports is equal to $\mathscr{R} = \mathscr{X} \times \mathscr{P}$.

3.4.1 Decomposable payment mechanisms

In general, a payment mechanism depends on all of the reported values, i.e., $\tau_S(X_a, F_a, ...)$. In this subsection, we consider a specific class of payment mechanisms τ_S that have a *decomposable* structure, meaning that they separately score an agent a's information report from her prediction report.

Definition 11. *Suppose that agents provide both their information reports and prediction reports, i.e.,* $\mathscr{R} = \mathscr{X} \times \mathscr{P}$. *We say that payment mechanism* τ_S *is decomposable if an agent* a*'s total payment is calculated as the sum of her information score and her prediction score, where the information score does not depend on the agent's prediction report and the prediction score does not depend on the agent's information report. More precisely:*

$$\tau_S(R_a, \mathbf{R}_p) = \underbrace{\tau_Y(Y_a, \mathbf{R}_p)}_{\textit{information score}} + \underbrace{\tau_F(F_a, \mathbf{R}_p)}_{\textit{prediction score}} \tag{3.9}$$

Having a decomposable structure, where an agent's information score is independent of her prediction report and her prediction score is independent of her information report, simplifies the analysis of the incentives as they do not influence each other. Notice that the robust Bayesian truth serum of [WP12b] is an example of a decomposable payment mechanism. For the limit case when the number of agents approaches infinity, the original Bayesian truth serum [Pre04] converges to a decomposable payment mechanism since a single agent does not have large impact on the frequencies of information reports nor on the average of prediction reports.

The structure of decomposable mechanisms allows us to analyze information and prediction scores separately. The prediction score elicits an agent's belief about what other agents have reported. Since the outcome is known to the mechanism (the mechanism knows what the other agents have reported), truthful elicitation of the prediction report can be achieved using proper scoring rules. Therefore, when proving impossibility results of decomposable payment mechanisms, one can focus on information scores.

Lemma 2. *There exists a decomposable payment mechanism that is strictly proper if and only if there exists an information score τ_Y such that the expected value $\mathbb{E}_{\mathbf{X}_p}(\tau_Y(Y_a, \mathbf{R}_p^{honest})|X_a)$ is strictly maximized for $Y_a = X_a$.*

Proof. Consider the following score:

$$\tau_F(F_a, \mathbf{R}_p^{honest}) = S(F_a, Y_p)$$

where S is a strictly proper scoring rule and Y_p is the information report of a randomly chosen peer p. Provided that the peer agents are honest ($Y_p = X_p$), the (strictly) best strategy for agent a is to report her true posterior belief regarding what her peers have observed, i.e., $F_a = Pr(X_p|X_a)$.

Since there exists a payment rule that elicits honest prediction reports regardless of the belief constraints $\mathscr{C}$, the existence of a decomposable strictly proper payment mechanism depends only on the existence of an information score that complies with the conditions of the lemma. Therefore, we proved the statement. □

We keep the notion of 1-peer payment mechanisms from the previous section, i.e., 1-peer payment mechanisms represent a restricted version of a general scoring functions and have a form $\tau_S(R_a, R_p)$. Notice that Definition 10 includes the payment functions with prediction reports. The following lemma shows the generality of 1-peer payment mechanisms in proving impossibility results.

Lemma 3. *If it is possible to construct a decomposable payment mechanism that is strictly proper under the set of belief constraints $\mathscr{C} \subseteq \{C_{CB}, C_{SC}\}$, then it is possible to construct a strictly proper 1-peer decomposable payment mechanism.*

Proof. Due to Lemma 2, we restrict our analysis to the information score. Let τ_S be a payment mechanism that satisfies the condition of the statement in Lemma 2, i.e., an agent a's best response to truthfulness of her peer agents is to report her honest information score. If $k = 1$ (agent a has only one peer), the statement follows immediately. Let us now consider $k > 1$.

The expected information score of an agent a who observes $X_a = x$ for reporting y, provided that her peers are honest, is equal to:

$$\sum_{x_1,...,x_k} Pr(X_{p_1} = x_1,...,X_{p_k} = x_k | X_a = x) \cdot \tau_Y(y, X_{p_1}, Pr_{p_1},...,X_{p_k}, Pr_{p_k})$$
$$= \sum_{x_1} (Pr(X_{p_1} = x_1 | X_a = x) \cdot$$
$$\cdot \sum_{x_2,...,x_k} Pr(X_{p_2} = x_2,...,X_{p_k} = x_k | X_a = x, X_{p_1} = x_1) \cdot \tau_Y(y, X_{p_1}, Pr_{p_1},...,X_{p_k}, Pr_{p_k}))$$
$$= \sum_{x_1} Pr(X_{p_1} = x_1 | X_a = x) \cdot \tilde{\tau}(y, X_{p_1}, Pr_{p_1})$$

where we put $\tilde{\tau}(y, X_{p_1}, Pr_{p_1}) = \sum_{x_2,...,x_k} Pr(X_{p_1} = x_1,...,X_{p_k} = x_k | X_a = x, X_{p_1} = x_1) \cdot$ $\cdot \tau_Y(y, X_{p_1}, Pr_{p_1},...,X_{p_k}, Pr_{p_k})$. Notice that $\tilde{\tau}(y, X_{p_1}, Pr_{p_1})$ depends on $Pr(X_{p_2} = x_2,...,X_{p_k} = x_k | X_a = x, X_{p_1} = x_1)$. However, the original mechanism is strictly proper under any belief profile that satisfies $\mathscr{C}$, so it must be strictly proper when the belief updating process of agent a keeps $Pr(X_{p_1} = x_1,...,X_{p_k} = x_k | X_a = x, X_{p_1} = x_1)$ fixed, but alters $Pr(X_{p_1} = x_1 | X_a = x)$. This implies the existence of a 1-peer payment mechanism that is strictly proper under constraints $\mathscr{C}$ because $\tilde{\tau}$ is strictly proper for arbitrary beliefs $Pr(X_{p_1} = x_1 | X_a = x)$. □

As in the previous section, we first provide the impossibility result: one cannot design a strictly proper decomposable payment mechanism under the common belief condition $\mathscr{C} = \{C_{CB}\}$. The proof of the formal result requires that the observation space is non-binary. If the observations were binary, then the self-predicting condition C_{SP} would hold under a relatively weak assumption that the same observation values are positively correlated, in particular, under the self-correlated condition C_{SC}. We show this in the proof of Corollary 1, while the following theorem provides the formal statement of the claim from the beginning of the paragraph presented in a more general form, that is, the one in which the set of belief constraints is any subset of $\{C_{CB}, C_{SC}\}$.

Theorem 2. *There exists no strictly proper decomposable mechanism under the set of belief constraints $\mathscr{C} \subseteq \{C_{CB}, C_{SC}\}$ when agents' observations take more than two values, i.e., $|\mathscr{X}| > 2$.*

Proof. Let us assume that there exists a strictly proper payment mechanism τ_S, and due to Lemma 3, we can assume that it is a 1-peer payment mechanism. Furthermore, because of Lemma 2, we restrict our attention to its information score.

Let $\mathbf{p}_x = Pr(X_p | X_a = x)$, $\mathbf{p}_y = Pr(X_p | X_a = y)$, $\mathbf{p}_z = Pr(X_p | X_a = z)$, etc., be some arbitrary distribution functions that define agents' common belief. Using the same arguments as in the proof of Theorem 1, we obtain that the strict properness of τ_S implies:

$$\mathbf{p}_x(x) \cdot (\tau_Y(x, x, \mathbf{p}_x) - \tau_Y(y, x, \mathbf{p}_x)) + \mathbf{p}_x(y) \cdot (\tau_Y(x, y, \mathbf{p}_y) - \tau_Y(y, y, \mathbf{p}_y))$$
$$+ \sum_{z \in \mathscr{X} \setminus \{x,y\}} \mathbf{p}_x(z) \cdot (\tau_Y(x, z, \mathbf{p}_z) - \tau_Y(y, z, \mathbf{p}_z)) > 0$$

$$\mathbf{p}_y(x)(\tau_Y(y,x,\mathbf{p}_x)-\tau_Y(x,x,\mathbf{p}_x))+\mathbf{p}_y(y)\cdot(\tau_Y(y,y,\mathbf{p}_y)-\tau_Y(x,y,\mathbf{p}_y))$$
$$+\sum_{z\in\mathscr{X}\setminus\{x,y\}}\mathbf{p}_y(z)\cdot(\tau_Y(y,z,\mathbf{p}_z)-\tau_Y(x,z,\mathbf{p}_z))>0$$

This gives us:

$$(\mathbf{p}_x(x)-\mathbf{p}_y(x))\cdot\Delta_x(\mathbf{p}_x)+(\mathbf{p}_x(y)-\mathbf{p}_y(y))\cdot\Delta_y(\mathbf{p}_y)$$
$$+\sum_{z\in\mathscr{X}\setminus\{x,y\}}(\mathbf{p}_x(z)-\mathbf{p}_y(z))\cdot\Delta_z(\mathbf{p}_z)>0 \tag{3.10}$$

where $\Delta_z(\mathbf{p}_z)=(\tau_Y(x,z,\mathbf{p}_z)-\tau_Y(y,z,\mathbf{p}_z))$. Since the mechanism should be strictly proper for arbitrary distribution functions, it should also be strictly proper for the following two cases:

1. When $\mathbf{p}_x(x)=\mathbf{p}_y(x)-\epsilon$, $\mathbf{p}_x(z)=\mathbf{p}_y(z)$ for $z\neq x,z'$, and $\mathbf{p}_x(z')=\mathbf{p}_y(z')+\epsilon$, where $\epsilon>0$ and $\epsilon<<1$. We denote this distribution by $\mathbf{p}_x^-$.

2. When $\mathbf{p}_x(x)=\mathbf{p}_y(x)+\epsilon$, $\mathbf{p}_x(z)=\mathbf{p}_y(z)$ for $z\neq x,z'$, and $\mathbf{p}_x(z')=\mathbf{p}_y(z')-\epsilon$, where $\epsilon>0$ and $\epsilon<<1$. We denote this distribution by $\mathbf{p}_x^+$.

Note that due to the *stochastic relevance* condition (see Section 3.1.1), we cannot put $\mathbf{p}_x=\mathbf{p}_y$. From (3.10) and $\mathbf{p}_x(z')-\mathbf{p}_y(z')=(1-\sum_{z\in\mathscr{X}\setminus\{z'\}}\mathbf{p}_x(z))-(1-\sum_{z\in\mathscr{X}\setminus\{z'\}}\mathbf{p}_y(z))=\sum_{z\in\mathscr{X}\setminus\{z'\}}(\mathbf{p}_y(z)-\mathbf{p}_x(z))$, we obtain:

$$-\epsilon\cdot(\Delta_x(\mathbf{p}_x^-)-\Delta_{z'}(\mathbf{p}_{z'}))>0$$

$$\epsilon\cdot(\Delta_x(\mathbf{p}_x^+)-\Delta_{z'}(\mathbf{p}_{z'}))>0$$

In other words:

$$\Delta_x(\mathbf{p}_x^-)<\Delta_{z'}(\mathbf{p}_{z'})$$

$$\Delta_x(\mathbf{p}_x^+)>\Delta_{z'}(\mathbf{p}_{z'}) \tag{3.11}$$

Let us consider a new $\mathbf{p}_y^{++}$ equal to: $\mathbf{p}_y^{++}(x)=\mathbf{p}_y(x)+2\epsilon$, $\mathbf{p}_y^{++}(z)=\mathbf{p}_y(z)$ for $z\neq x,z'$, and $\mathbf{p}_y^{++}(z')=\mathbf{p}_y(z')-2\epsilon$. By applying the previous steps on $\mathbf{p}_y^{++}$, we obtain:

$$\Delta_x(\mathbf{p}_x^+)=\Delta_x(\mathbf{p}_x^{++-})<\Delta_{z'}(\mathbf{p}_{z'}) \tag{3.12}$$

$$\Delta_x(\mathbf{p}_x^{+++})>\Delta_{z'}(\mathbf{p}_{z'})$$

Hence we have a contradiction (expressions (3.11) and (3.12)). Therefore, there exists no decomposable payment scheme that is strictly proper even if the common belief condition C_{CB} holds. All that is left to be shown is that that there exist distributions $\mathbf{p}_y=Pr(X_p|X_a=y)$,

$\mathbf{p}_z = Pr(X_p|X_a = z)$, etc., such that posteriors $\mathbf{p}_y$, $\mathbf{p}_z$, ..., $\mathbf{p}_x^-$, $\mathbf{p}_x^+$, $\mathbf{p}_y^{++}$, $\mathbf{p}_x^{++-}$, and $\mathbf{p}_x^{+++}$, satisfy the self-correlated condition C_{SC}. To do so, let us fix $\mathbf{p}_y$ and let a fully mixed prior $Pr(X_p)$ be such that $Pr(X_p = x) = \mathbf{p}_y(x) - 2 \cdot \epsilon$ and $Pr(X_p = y) = \mathbf{p}_y(y) - \epsilon$. Furthermore, let $\mathbf{p}_z(z) > Pr(X_p = z)$ for all $z \neq x, y$. Then, the beliefs in the first part of the proof do satisfy the basic properties of the setting and, moreover, the self-correlated condition C_{SC}. Hence, we proved the statement. □

The impossibility result obtained by Theorem 2 is quite surprising. Even though the center elicits agents' posterior beliefs, which are formed from a common belief system, it still requires additional constraints to achieve strict properness. Nevertheless, decomposable payment mechanisms are strictly more general than single-report mechanisms in terms of elicitability. Clearly, if we can elicit truthful observations using a single-report mechanism, then one can construct a decomposable mechanism for which agents are truthful as well: one simply applies the single-report mechanism as its information score, which by Lemma 2 implies the existence of a strictly proper decomposable mechanism. To show that decomposable payment mechanisms require strictly less restrictions, it suffices to construct a decomposable mechanism that is strictly proper under belief constraints $\mathscr{C} = \{C_{CB}, C_{SP}\}$. Namely, Theorem 1 tells us that under these constraints, a single-report payment mechanism cannot be strictly proper. This leads us to a decomposable payment mechanism called the *multi-valued robust Bayesian truth serum*.

Multi-valued Robust Bayesian Truth Serum. The multi-valued RBTS mechanism has the following steps:

- Each agent a is asked to provide two reports:
  - *information report* Y_a, which represents agent a's observed value;
  - *prediction report* F_a, which represents agent a's prediction about the frequencies of reported values in the overall population.
- Each agent a is linked with one *peer* agent p and is rewarded with a score:[5]

$$\tau_S(Y_a, F_a, Y_p, F_p) = \underbrace{\frac{\mathbb{1}_{Y_p = Y_a}}{F_p(Y_a)}}_{\text{information score}} + \underbrace{S(F_a, Y_p)}_{\text{prediction score}} \tag{3.13}$$

  where $\mathbb{1}$ is an indicator variable and S is a strictly proper scoring rule.

The multi-valued RBTS mechanism is an example of a strictly proper decomposable payment mechanism for the set of belief constraints equal to $\mathscr{C} = \{C_{CB}, C_{SP}\}$. The direct consequence is a strict generality of decomposable payment schemes over single-report payment mechanisms. More formally:

[5] To avoid large information scores for small values of F_p, one can multiply the information score by $\min_y(F_p(y))$ and set the information score equal to 1 if $F_p(Y_a) = 0$.

Theorem 3. *Suppose that there are $N \geq 2$ agents and each agent has $k \geq 1$ different peers. Then, the multi-valued RBTS mechanism is strictly proper under the set of belief constraints $\mathscr{C} = \{C_{CB}, C_{SP}\}$.*

Proof. Consider an agent a and suppose her peer agent p is honest. Since the prediction score is a strictly proper scoring rule, agent a strictly maximizes it by reporting her true posterior belief as her prediction report. Therefore, it is enough to examine the properties of the information score.

Suppose that agent a observes x and reports y. Since peer p is honest, her prediction report F_p is equal to $F_p = Pr_p(X_{p_p}|X_p = z)$, where z is peer p's observation and p_p is the peer of p. Since belief systems are common, $Pr_p(X_{p_p}|X_p = z) = Pr_a(X_p|X_a = z)$. This means that $F_p = Pr_a(X_p|X_a = z)$, so the expected value of the agent's information score is equal to:

$$\mathbb{E}(\tau_Y(y, Y_p, F_p)) = \frac{Pr_a(X_p = y|X_a = x)}{Pr_p(X_{p_p} = y|X_p = y)} = \frac{Pr_a(X_p = y|X_a = x)}{Pr_a(X_p = y|X_a = y)}$$

By taking into account the self-predicting condition, i.e., $Pr_a(X_p = z|X_a = x) < Pr_a(X_p = z|X_a = z), \forall z \neq x$, we get that the expected value of the information score is (strictly) maximized for:

$$\operatorname*{argmax}_y \mathbb{E}(\tau_Y(y, Y_p, F_p)) = \operatorname*{argmax}_y \frac{Pr_a(X_p = y|X_a = x)}{Pr_a(X_p = y|X_a = y)} = \{x\}$$

Therefore, the strict maximum of the information score is achieved when the agent reports her true observation. This completes the proof. □

An interesting implication of Theorem 3 is that the multi-valued RBTS is strictly proper provided that agents' common belief system satisfies the self-correlated condition, while their observations are binary signals. Namely, in the case of binary observations, the self-correlated condition implies the self-predicting condition, which means that the set of belief constraints corresponds to the one in Theorem 3.

Corollary 1. *Suppose that there are $N \geq 2$ agents and each agent has $k \geq 1$ different peers. Then, the multi-valued RBTS mechanism is strictly proper under the set of belief constraints $\mathscr{C} = \{C_{CB}, C_{SC}\}$ when the observation space is a binary set, i.e., $|\mathscr{X}| > 2$.*

Proof. By Theorem 3, it suffices to show that the self-correlated condition C_{SC} implies the self-predicting condition C_{SP} whenever $\mathscr{X}$ is a binary set. Suppose that the self-correlated condition C_{SC} holds. Then, for all $x \in \mathscr{X}$ there exists $\epsilon_x > 0$ such that $Pr(X_p = x|X_a = x) =$

$Pr(X_p = x) + \epsilon_x$. Therefore:

$$\frac{Pr(X_p = x|X_a = x)}{Pr(X_p = x)} = \frac{Pr(X_p = x) + \epsilon_x}{Pr(X_p = x)} > 1 > \frac{1 - Pr(X_p = x) - \epsilon_x}{1 - Pr(X_p = x)} = \frac{1 - Pr(X_p = x|X_a = x)}{1 - Pr(X_p = x)}$$
$$= \frac{Pr(X_p = y|X_a = x)}{Pr(X_p = y)}$$

which means that the self-predicting condition holds. Hence, we proved the statement. □

Notice that the robust BTS of [WP12b] represents a similar possibility results for a binary observation space: the setting it operates in satisfies the conditions of Corollary 1. Its generalization for non-binary observations requires an additional belief condition, different than and not comparable to the self predicting condition [Wit14].

3.4.2 Divergence-based Bayesian truth serum

The impossibility result of Theorem 2 motivates us to examine a broader class of scoring functions than that of the decomposable payment mechanisms. In this subsection, we investigate mechanisms which are composed of information and prediction scores, but the information score is no longer independent of the prediction report.

As it is common in all of the BTS mechanisms, the prediction score is a strictly proper scoring rule applied on the prediction report of an agent a and the information report of her peer p. However, the information score is intuitively different: it *penalizes* the agent if her information report agrees with that of her peer while their prediction reports are significantly different. Disagreement between prediction reports is characterized by the condition that the divergence between the reports is larger than a threshold Θ.

Divergence-based Bayesian Truth Serum. The divergence-based BTS has the following steps:

- Each agent a is asked to provide her information report Y_a and her prediction report F_a.
- Each agent a is linked with a *peer* agent p and is rewarded with:

$$\tau_S(Y_a, F_a, Y_p, F_p) = \underbrace{-\mathbb{1}_{Y_a = Y_p \wedge D(F_a||F_p) > \Theta}}_{\text{information score}} + \underbrace{S(F_a, Y_p)}_{\text{prediction score}}$$

  where $\mathbb{1}$ is an indicator variable, S is a strictly proper scoring rule, $D(||)$ is the divergence associated to a strictly proper scoring rule, and Θ is a parameter of the mechanism.

The intuition behind the penalty of the divergence-based BTS is that honest agents will not have inconsistent prediction reports and consistent information reports. This is certainly the case if agents have a common belief system, but the mechanism also allows deviations from this condition. The following theorem shows the condition on the belief systems and the choice of parameter Θ that make this intuition true.

Theorem 4. *Suppose that there are $N \geq 2$ agents and each agent has $k \geq 1$ different peers. Then, the divergence-based BTS mechanism is strictly proper under the divergence-based condition, i.e., $\mathscr{C} = \{C_{DB}\}$, for Θ that satisfies:*

$$D(Pr_a(X_p|X_a = x)||Pr_p(X_{p_p}|X_p = x)) \leq \Theta < D(Pr_a(X_p|X_a = x)||Pr_p(X_{p_p}|X_p = y)) \tag{3.14}$$

for all $x \neq y$, where we denoted a peer of p by p_p.

Proof. Consider an agent a who observes x and believes that her peer agent is honest. Furthermore, suppose that Θ satisfies the conditions of the theorem.

Due to the properties of the strictly proper scoring rules, agent a's prediction score is in expectation maximized when she reports $F_a = Pr_a(X_p|X_a = x)$, and because stochastic relevance holds, this is a strict optimum.

If agent a's prediction report is $F_a = Pr_a(X_p|X_a = x)$, then we conclude from condition (3.14) that the maximum of her information score is achieved when she reports x, and is equal to 0.

Since the optimal value of the information score is equal to 0 and the prediction score is maximized for $F_a = Pr_a(X_p|X_a = x)$, it follows that $Y_a = x$ and $F_a = Pr_a(X_p|X_a = x)$ is agent a's best response.

We still need to prove that this is the strictly optimal response. Since $F_a = Pr_a(X_p|X_a = x)$ is the strictly optimal response for the prediction score, and $Y_a = x$ achieves the optimal value of the information score, it is enough to show that agent a's information score is negative in expectation for $Y_a \neq x$ and $F_a = Pr_a(X_p|X_a = x)$. Due to condition (3.14) and the fully mixed posteriors, the expected score for reporting $y \neq x$ and $F_a = Pr_a(X_p|X_a = x)$ is:

$$\begin{aligned}
&Pr_a(X_p = y|X_a = x) \cdot (-\mathbb{1}_{D(Pr_a(X_p|X_a=x)||F_p)>\Theta}) \\
&= Pr_a(X_p = y|X_a = x) \cdot (-\mathbb{1}_{D(Pr_a(X_p|X_a=x)||Pr_p(X_{p_p}|X_p=y))>\Theta}) = -Pr_a(X_p = y|X_a = x) < 0
\end{aligned}$$

where we used the fact that peer p is honest, and, thus, $F_p = Pr_p(X_{p_p}|X_p = y)$. Putting it all together, the divergence-based BTS is strictly proper. □

The direct consequence of Theorem 4 is that the divergence-based BTS is strictly proper under the common belief condition C_{CB} with an appropriate choice of parameter Θ.

Corollary 2. *Suppose that there are $N \geq 2$ agents and each agent has $k \geq 1$ different peers. Then, the divergence-based BTS with $\Theta = 0$ is strictly proper under the common belief condition, i.e., $\mathscr{C} = \{C_{CB}\}$.*

Proof. The common belief condition C_{CB} is a special case of the divergence based condition C_{DB} in which $D(Pr_a(X_p|X_a = x)||Pr_p(X_{p_p}|X_p = x)) = 0$ for all $x \in \mathscr{X}$, and due to stochastic relevance $D(Pr_a(X_p|X_a = x)||Pr_p(X_{p_p}|X_p = y)) > 0$ for all $y \neq x$. Therefore, we can set $\Theta = 0$ to satisfy the conditions of Theorem 4, which then implies the claim. □

A convenient feature of the divergence-based BTS is that it allows a population of agents to have different belief systems, as long as the agents' posteriors are more similar when they observe the same value than when their observations are different. This is exactly what condition (3.14) states, and is realistic if agents indeed have similar observations. Notice that we formalized similarities between posteriors of two different agents using parameter Θ. Although we have assumed that the center knows Θ, it is possible to make the divergence-based BTS a non-parametric method.

Non-parametric Divergence-based Bayesian Truth Serum. To make the divergence-based BTS a non-parametric method, we change its second step. In addition to a peer agent p, the modified method also uses another peer agent $\hat{p}$, and the overall score for agent a becomes:

$$\tau_S(Y_a, F_a, Y_p, F_p, Y_{\hat{p}}, F_{\hat{p}}) = \underbrace{-\mathbb{1}_{Y_p = Y_a \wedge Y_{\hat{p}} \neq Y_a \wedge D(F_a||F_p) > D(F_a||F_{\hat{p}})}}_{\text{information score}} + \underbrace{S(F_a, Y_p)}_{\text{prediction score}}$$

Theorem 5. *Suppose that there are $N \geq 3$ agents and each agent has $k \geq 2$ different peers. Then, the non-parametric divergence-based BTS mechanism is strictly proper under the divergence-based condition, i.e., $\mathscr{C} = \{C_{DB}\}$.*

Proof. Consider an agent a who observes $X_a = x$, and suppose her peer agents p and $\hat{p}$ are honest. Due to the divergence-based condition, $D(F_a||F_p) < D(F_a||F_{\hat{p}})$ holds whenever agent a reports her true observation x and her true prediction $F_a = Pr_a(X_p|X_a = x)$. In that case, the information score achieves the optimal value. Because the prediction score is a strictly proper scoring rule, agent a strictly optimizes it by reporting her true prediction $F_a = Pr_a(X_p|X_a = x)$.

What is left to be shown is that the information score is also strictly optimal. This can be done using the same reasoning as in Theorem 4. If agent a reported $y \neq x$ and $F_a = Pr_a(X_p|X_a = x)$, she would in expectation receive negative information score:

$$\begin{aligned}
&Pr_a(X_p = y|X_a = x) \cdot (-\mathbb{1}_{D(Pr_a(X_p|X_a=x)||F_p) > D(Pr_a(X_p|X_a=x)||F_{\hat{p}})}) \\
&\leq Pr_a(X_p = y, X_{\hat{p}} = x|X_a = x) \cdot (-\mathbb{1}_{D(Pr_a(X_p|X_a=x)||Pr_p(X_{p_p}|X_p=y)) > D(Pr_a(X_p|X_a=x)||Pr_{\hat{p}}(X_{p_{\hat{p}}}|X_{\hat{p}}=x))}) \\
&= -Pr_a(X_p = y, X_{\hat{p}} = x|X_a = x) < 0
\end{aligned}$$

where we denoted peers of p and $\hat{p}$ by p_p and $p_{\hat{p}}$, respectively. The first inequality comes from the fact that we focus on the case when $X_{\hat{p}} = x$. In that case, $F_p = Pr_p(X_a|X_p = y)$ and $F_{\hat{p}} = Pr_{\hat{p}}(X_a|X_{\hat{p}} = x)$ because peers p and $\hat{p}$ are honest. Therefore, agent a is better off reporting x instead, which completes the proof. □

Remark 2. The divergence-based BTS asks agents to report full posterior beliefs regarding their peers' observations. Provided that the stochastic relevance condition and the divergence-based condition hold for coarser-grained observations as well, i.e., for posterior beliefs that could indicate probabilities of a peer observing a strict subset of possible observations, then one can minimize the size of the prediction report without affecting the incentive properties. Namely, in that case, one could ask agents to report a binary prediction on whether their peers observe a particular strict subset of possible observations, instead of asking them to report the full frequency profile of reports. The strict subset can be randomly chosen, but one can also use a domain specific knowledge to select a strict subset that maximizes the difference between the expected payoffs for honest reporting and inaccurate reporting.

Remark 3. The penalty of the divergence-based BTS can be controlled by how much predictions of an agent and her peer differ. For example, if the penalty condition is satisfied, then the agent could get a penalty equal to $-D(F_a||F_p)$ instead of -1. This way, more severe deviations would lead to greater penalties.

3.4.3 Continuous Bayesian truth serum

The structure of the divergence-based score can be extended to allow observations that are real numbers. Now, the elicitation mechanism has to take into account that perfect matching of agents' reports, as done in the information score of the divergence-based BTS mechanism, would not produce sensible incentives since we are dealing with continuous observations. The trick is to consider two observations similar if they are no more than d away from each other, where d is a randomly chosen parameter. Then, we can apply the principle of the divergence-based BTS mechanism: similar observations should lead to similar posterior beliefs.

Continuous Bayesian Truth Serum. Consider observations X_a taking continuous values, in particular, $\mathscr{X} = \mathbb{R}$. The continuous BTS mechanism has the following steps:

- For each agent a, the mechanism samples a number d_a from a uniform distribution on interval $(0,1)$[6], i.e., $d_a = rand((0,1))$. The continuous domain $\mathscr{X} = \mathbb{R}$ is then uniformly discretized with the discretization interval of a size d_a and the constraint that value Y_a is in the middle of the interval it belongs to. We denote the interval of a value $Y_a = y$ by I_y^a. The constraint can then be written as $y = \frac{\max I_y^a - \min I_y^a}{2}$.
- Finally, an agent a is scored using a modified version of the divergence-based BTS score:

$$\tau_S(Y_a, F_a, Y_p, F_p) = \underbrace{-\mathbb{1}_{Y_p \in I_y^a \wedge D(F_a||F_p) > d_a \cdot \Theta}}_{\text{information score}} + \underbrace{S(F_a, Y_p)}_{\text{prediction score}}$$

[6] In general, one can use interval $(0, b)$, where $b > 0$, and a non-uniform distribution that has a full support on that interval.

where $D(||)$ is the divergence of a strictly proper scoring rule, and S is a strictly proper scoring rule.

Parameter Θ of the continuous BTS reflects how close the posteriors of two similar signals are. For example, when agents are fully confident in the correctness of their objective observations, Θ should be big because posteriors of two similar signals can be significantly different. On the other hand, when agents make mistakes, the posteriors of two similar signals are close to each other, making the lower bound on Θ smaller. In certain applications, this fact can be used to set the appropriate value of Θ. For example, in crowd sensing, the center could assume that every crowd sensor is worse than some accurate sensor, so the center can adjust Θ according to the specifics of the accurate sensor. For a value of Θ parameter such that $d_a \cdot \Theta$ never underestimates the divergence $D(||)$ of agents posteriors that observe similar values, the continuous BTS is strictly proper under the common belief condition C_{CB}.

Theorem 6. *Suppose that there are $N \geq 2$ agents and each agent has $k \geq 1$ different peers. Consider $\mathscr{X} = \mathbb{R}$, and suppose $\Theta \in (0, \infty)$ is such that $\forall x \in \mathscr{X}, d_a \in (0,1), z \in I_x^a$:*

$$D(p(X_p|X_a = x)||p(X_p|X_a = z)) \leq d_a \cdot \Theta \tag{3.15}$$

where $p(X_p|X_a)$ denotes agent a's posterior belief. Then, the continuous BTS mechanism is strictly proper under the common belief condition, i.e., $\mathscr{C} = \{C_{CB}\}$.

Proof. Suppose agent a observes x and believes that her peer p is honest. Since peer p is honest, from the common belief condition C_{CB}, we conclude that $F_p = p_p(X_{p_p}|X_p = z) = p(X_p|X_a = z)$ when peer p's observation is $X_p = z$ (notice that p_p is a peer of p). Therefore, if agent a reports x and $F_a = p(X_p|X_a = x)$, her information score is equal to 0, because (3.15) holds. The prediction score is a strictly proper scoring rule, so in expectation the optimal choice for the prediction report is agent a's posterior $F_a = p(X_p|X_a = x)$ - this is a strict optimum due to stochastic relevance. Therefore, reporting $Y_a = x$ and $F_a = p(X_p|X_a = x)$ results in the maximum payoff. As it was the case with Theorem 4, we need to show that agent a's information score is in expectation strictly negative for any information report $Y_a \neq x$, provided that she reports the strictly optimal prediction report, i.e., $F_a = p(X_p|X_a = x)$.

Let $Y_a = y \neq x$. Consider d_{a1} such that $x \notin I_y^{a1}$. From stochastic relevance, we know that there exists $\epsilon > 0$ such that:

$$\forall z \in I_y^{a1} : D(p(X_p|X_a = x)||p(X_p|X_a = z)) > \epsilon \tag{3.16}$$

Now, consider $d_{a2} \leq \min(d_{a1}, \frac{\epsilon}{\Theta})$. Since $I_y^{a2} \subseteq I_y^{a1}$, inequality (3.16) implies:

$$\forall z \in I_y^{a2} : D(p(X_p|X_a = x)||p(X_p|X_a = z)) > \epsilon \geq d_{a2} \cdot \Theta$$

Notice that $Pr(d_a \leq d_{a2}) > 0$, because d_a is chosen uniformly at random from $(0,1)$. Moreover, $Pr(X_p \in I_y^{a2}|X_a = x) = \int_{z \in I_y^{a2}} p(X_p = z|X_a = x)dz > 0$ due to the fully mixed posteriors. Since

$F_p = p(X_p|X_a = z)$ when peer p's observation is $X_p = z$, we have that for information report $Y_a = y \neq x$ and prediction report $F_a = p(X_p|X_a = x)$, the expected information score of agent a is strictly negative. Therefore, the continuous BTS is strictly proper. □

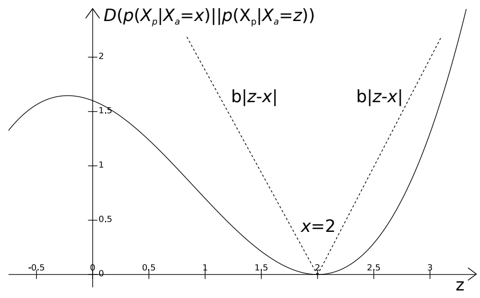


Figure 3.2 – The divergence of posteriors as a function of a peer's report.

It remains to see how to set the parameter Θ. Consider $D(p(X_p|X_a = x)||p(X_p|X_a = z))$ as a function of z for a fixed $x = 2$. Condition (3.15) simply states that one can find a coefficient b such that $b \cdot |z-2| \geq D(p(X_p|X_a = 2)||p(X_p|X_a = z))$ for $z \in (1.5, 2.5)$. As shown in Figure 3.2, this corresponds to the divergence being bounded by two lines. More formally:

Proposition 3. *Consider* $\mathscr{X} = \mathbb{R}$. *If* $\forall x \in \mathbb{R}$, $D(p(X_p|X_a = x)||p(X_p|X_a = z))$ *is a continuously differentiable and bounded function of* $z \in (x - \frac{1}{2}, x + \frac{1}{2})$, *then:*

$$\Theta > \max_x \max_{z \in (x-\frac{1}{2}, x+\frac{1}{2})} \left| \frac{\partial D(p(X_p|X_a = x)||p(X_p|X_a = z))}{\partial z} \right|$$

satisfies condition (3.15) *of Theorem 6.*

Proof. Consider a function $f(z) = \Theta \cdot |z - x| - D(p(X_p|X_a = x)||p(X_p|X_a = z))$ defined for a specific value of x on interval $z \in (x - \frac{1}{2}, x + \frac{1}{2})$. For $z = x$, the function is equal to $f(x) = 0$. Let us consider $z' \in (x, x + \frac{1}{2})$. We have:

$$\begin{aligned} f(z') &= \Theta \cdot |z' - x| - D(p(X_p|X_a = x)||p(X_p|X_a = z')) \\ &= \int_x^{z'} \left(\Theta - \frac{\partial D(p(X_p|X_a = x)||p(X_p|X_a = z))}{\partial z} \right) dz = \int_x^{z'} g(z) dz \end{aligned}$$

where $g(z) = \Theta - \frac{\partial D(p(X_p|X_a=x)||p(X_p|X_a=z))}{\partial z}$. Due to the condition of the proposition, we know that $g(z) > 0$ for all $z \in (x, x+\frac{1}{2})$. Therefore, $f(z')$ strictly increases as $|z'-x|$ increases. Similarly, we obtain that the same holds for $z' \in (x - \frac{1}{2}, x)$. By combining these results with $f(x) = 0$, we obtain the claim. □

The continuous BTS is a parametric mechanism, so the center needs to set the parameter Θ. Notice that the only restriction for strict properness is that the center sets Θ to a big enough value. However, there is a tradeoff between the value of Θ and the expected value of margin difference of the information score between truthful and non-truthful reporting. That is, the larger Θ is, the smaller the expected punishment is for an agent who deviates from truthful reporting.

If the divergence function $D(p(X_p|X_a = x)||p(X_p|X_a = z))$ increases as $|z - x|$ increases, it is possible to make the continuous BTS a non-parametric method.

Non-parametric Continuous Bayesian Truth Serum. To make the continuous BTS a parameter-free method, we introduce another peer agent $\hat{p}$ and change the score to:

$$\tau_S(Y_a, F_a, Y_p, F_p, Y_{\hat{p}}, F_{\hat{p}}) = \underbrace{-\mathbb{1}_{|Y_p - Y_a| < |Y_{\hat{p}} - Y_a| \wedge D(F_a||F_p) > D(F_a||F_{\hat{p}})}}_{\text{information score}} + \underbrace{S(F_a, Y_p)}_{\text{prediction score}}$$

Proposition 4. *Suppose that there are $N \geq 3$ agents and each agent has $k \geq 2$ different peers. Let $\mathcal{X} = \mathbb{R}$. If $\forall x, y, z \in \mathbb{R}$:*

$$|y - x| < |z - x| \implies D(p(X_p|X_a = x)||p(X_p|X_a = y)) < D(p(X_p|X_a = x)||p(X_p|X_a = z))$$

where p is agent a's posterior belief, then the non-parametric continuous BTS is strictly proper under the common belief condition, i.e., $\mathcal{C} = \{C_{CB}\}$.

Proof. The common belief condition implies that $p_p(X_{p_p}|X_p = y) = p(X_p|X_a = y)$ and $p_{\hat{p}}(X_{p_{\hat{p}}}|X_{\hat{p}} = y) = p(X_p|X_a = y)$, where p_p and $p_{\hat{p}}$ are peers of p and $\hat{p}$, respectively. Therefore, for agent a who observes x and honest agents p and $\hat{p}$, $D(F_a||F_p) < D(F_a||F_{\hat{p}})$ holds whenever $|Y_p - X_a| < |Y_{\hat{p}} - X_a|$ and agent a reports truthfully. In that case (or if $|Y_p - X_a| > |Y_{\hat{p}} - X_a|$), the information score achieves the optimal value. Furthermore, the prediction score is a strictly proper scoring rule, so agent a's best response is to report $Y_a = x$ and $F_a = p(X_p|X_a = x)$.

Similarly to how it is done in Theorem 6, we can show that reporting truthfully is agent a's best response. First of all, the prediction score achieves its strict optimum for $F_a = p(X_p|X_a = x)$ due to the stochastic relevance condition. In other words, we only need to show that for any information report $Y_a \neq x$, agent a's information score is in expectation strictly negative when she provides prediction report $F_a = p(X_p|X_a = x)$.

Suppose that the agent reports $Y_a = y \neq x$. Furthermore, suppose that $|Y_p - x| < |Y_{\hat{p}} < x|$, while $|Y_{\hat{p}} - y| < |Y_p - y|$. Due to the fully mixed beliefs, this happens with probability strictly greater than 0, which implies that the expected information score is strictly smaller than 0. Since 0 is the optimal value of the information score, agent a's best response to truthful reporting of other agents is to report truthfully. □

Remark 4. One might be wondering how efficient is the elicitation of continuous variables in terms of the number of reported values. For practical considerations, agents' beliefs can often be modeled using a parametric density distributions. For example, an agent a's posterior belief could be modeled with a Gaussian distribution $\mathcal{N}(\mu_x, \sigma_x)$, whose parameters depend on the agent's observation x. In this case, reporting the posterior belief comes down to reporting two parameters μ_x and σ_x, so the whole report consists of only three scalar values. We investigate in the following subsection a possible belief model that complies with these conditions.

Importance of the common belief condition

Unlike the divergence-based BTS mechanism, which allows agents to have limited difference in their belief systems, the continuous BTS mechanism requires agents to share the same belief system. In this subsection, we further investigate the importance of the common belief condition C_{CB} for elicitation of real-valued observations.

Let us consider a natural belief system based on the Gaussian state model condition C_{GSM} in which the parameters are Gaussian distribution functions. More precisely, the observation values X_a are generated by a Gaussian $p(X_a) \sim \mathcal{N}(\mu_\Omega, \sigma)$, where σ is fixed (given), while μ_Ω defines state Ω and is distributed according to Gaussian distribution $p(\mu_\Omega) \sim \mathcal{N}(\mu_0, \sigma_0)$.

Now, suppose that agent a observes value $X_a = x$. From the Bayesian updating of Gaussian distributions [Bis06], it follows that agent a's posterior belief regarding peer p's observations is a Gaussian $p(X_p|X_x = x) \sim \mathcal{N}(\mu_x, \sigma_x)$ with the parameters equal to:

$$\mu_x = \frac{\frac{\mu_0}{\sigma_0^2} + \frac{x}{\sigma^2}}{\frac{1}{\sigma_0^2} + \frac{1}{\sigma^2}}, \sigma_x^2 = \frac{1}{\frac{1}{\sigma_0^2} + \frac{1}{\sigma^2}} + \sigma^2 \tag{3.17}$$

The KL divergence of two normal distributions $\mathcal{N}(\mu_1, \sigma_1)$ and $\mathcal{N}(\mu_2, \sigma_2)$ is equal to [Iha93]:

$$\log \frac{\sigma_2}{\sigma_1} + \frac{\sigma_1^2}{2\sigma_2^2} + \frac{(\mu_1 - \mu_2)^2}{2\sigma_2^2} - \frac{1}{2} \tag{3.18}$$

From the expressions (3.17) and (3.18) it follows that the KL divergence between agent a's posterior belief when she observes x and her posterior when she observes z is:

$$D(p(X_p|X_a = x)||p(X_p|X_a = z)) = \hat{\sigma} \cdot \frac{(x - z)^2}{2\sigma^2}$$

where $\hat{\sigma} = \frac{\sigma_0^2}{\sigma_0^2 + \sigma^2} \frac{\sigma_0^2}{2\sigma_0^2 + \sigma^2}$. Using Proposition 3, we obtain Θ that satisfies the conditions of Theorem 6:

$$\begin{aligned} \Theta &\geq \max_x \max_{z \in (x - 1/2, x + 1/2)} \left| \hat{\sigma} \cdot \frac{2 \cdot (x - z)}{2 \cdot \sigma^2} \right| \\ &\geq \hat{\sigma} \cdot \frac{1}{2 \cdot \sigma^2} \end{aligned} \tag{3.19}$$

The center does not need to know parameters σ and σ_0^2: it is sufficient to overestimate (3.19). We often have that $\sigma_0 \gg \sigma$, and hence $\hat{\sigma} \approx \frac{1}{2}$. In that case, the center only needs to underestimate the value of σ. For example, if the agents are crowd sensors with accuracy below a certain threshold, the center can infer the minimal value of σ. Furthermore, the KL divergence between two Gaussian posteriors satisfies the conditions of Proposition 4, which means that one can also use the non-parametric continuous BTS. Considering that the divergence-based BTS allows agents to have differences in their belief systems, one might wonder if it is possible to relax the common belief condition when observations take values in $\mathbb{R}$. The following theorem shows that this is not possible under the Gaussian state model condition $\mathscr{G} = \{C_{GSM}\}$. The intuition is that one can define two belief models B_{a1} and B_{a2} that have the same posterior for two different observations x_a and y_a. The strict properness condition, however, requires that only one information report results in optimal payment.

Theorem 7. *No mechanism τ_S is strictly proper under the Gaussian state model condition $\mathscr{G} = \{C_{GSM}\}$. In particular, consider an agent a and suppose her peers have a common belief system satisfying condition C_{GSM}. If a mechanism τ_S strictly incentivizes agent a to report honestly when she adopts the common belief system, then it does not strictly incentivize her to report honestly when she adopts an alternative belief system that differs from the common one in parameter μ_0.*

Proof. Due to Lemma 3, we can, without loss of generality, focus on mechanisms in which agents report both information and prediction reports. Let us assume the opposite, i.e., there exists a mechanism τ_S that incentivizes agents to report honestly under the conditions of the theorem. Suppose that agent a observes value x_a and adopts a common belief system. Let us denote the expected payoff of agent a by $\bar{\tau}_a$. Strict properness of τ_S implies:

$$\forall z \neq x_a : \bar{\tau}_a(x_a, F_a) > \bar{\tau}_a(z, F_a) \tag{3.20}$$

where $F_a = (\mu_{x_a}, \sigma_{x_a})$. Consider an alternative belief system defined by $\mu_0 = {\mu'}_0^a \neq \mu_0^a$ and an alternative observation $y_a = \frac{\sigma^2}{\sigma_0^2}(\mu_0^a - {\mu'}_0^a) + x_a$. From expression (3.17), we know that in this case agent a's posterior is the same as for the previous case, i.e., $\mu_{x_a} = {\mu'}_{y_a}$ and $\sigma_{x_a} = {\sigma'}_{y_a}$. However, from (3.20) it follows that the best response of agent a is to report x_a, not y_a. That is, τ_S cannot incentivize both agent a, who has the same belief system as her peers p, and agent a, who has a different belief system than her peers p, to report honestly. □

3.5 Applications

We investigate two applications of the mechanisms presented in this section. First one is community sensing, where we consider a sensing scenario in which a relatively small number of sensors provide their measurements to the center. We demonstrate the importance of having robust mechanisms designed for the elicitation of non-binary information. As the second application, we study the design of a BTS type of mechanism in elicitation of subjective information through a contest and we present our preliminary work on eliciting emotions across the EPFL campus.

3.5.1 Community sensing

In a community sensing scenario, private mobile devices equipped with sensors acquire information about a spatially distributed phenomenon, such as air pollution or weather, and report it to the center [BEH⁺06, ASC⁺10, KHKZ08]. Since sensing induces a cost due to the fact that sensing modules need to be installed and maintained, the party interested in monitoring the physical phenomenon needs to incentivize the crowd to incur this cost and provide quality data.

We investigate a community sensing setting in which the center has no control over sensing devices, nor does it have a way of directly verifying the correctness of the obtained data. Therefore, to compensate for the cost of sensing, the center applies a peer consistency approach in rewarding sensors. One of the peer consistency methods proposed for information elicitation in community sensing setting is the mechanism from [FLJ14]. In this section, we provide a complementary approach that does not require the center to know participants' beliefs, but is rather based on participants reporting their beliefs. In particular, we compare the rewards of the continuous BTS with those produced by the output agreement and the RBTS for small population as presented in [WP12b][7], and demonstrate the importance of proper mechanisms that allow non-binary observations.

As an example of a community sensing scenario, we consider an air-quality monitoring over an urban area. Each sensor is assumed to be a rational agent that measures air pollution at its location and reports its measurement to the center.

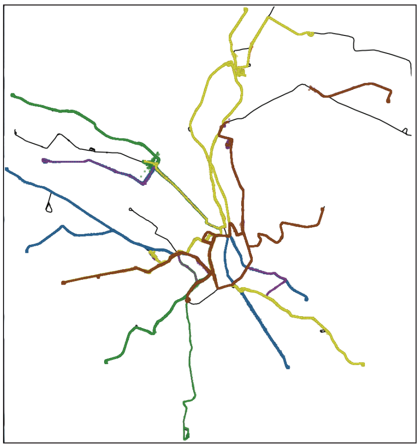
Simulation setup

Our community sensing test-bed is based on a real dataset containing levels of UFP (ultra fine particles) over Zurich urban area — the full description of the dataset can be found in [HSW⁺14] — and a region based Gaussian Processes model that incorporates the spatial features of the terrain, as described in [JLF14]. One chunk of data contains aggregated measurements over the period of two weeks for 200 locations. In total, we have 22 chunks of data, 2 for each of 11 different months. Half of the data is used for training the GP model and the other half for testing purposes.

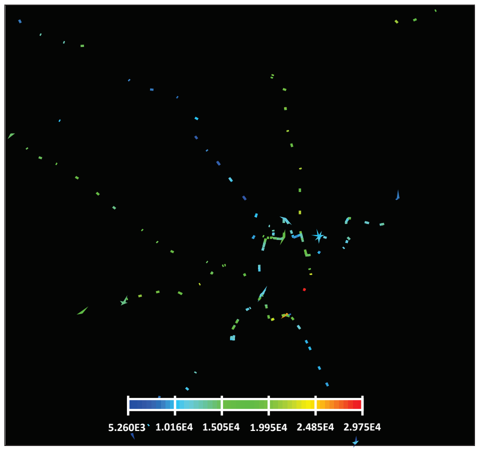
The 200 locations are unevenly distributed and concentrated on the main tram lines in the city of Zurich, as shown in Figure 3.3. Therefore, we investigate the use of peer consistency techniques designed for a single-task scenario. We show in the next chapter how to design more robust incentives provided that measurement locations are more evenly spread across an urban area, while sensors report more frequently.

We consider a group of 10 sensors whose initial locations are obtained by randomly sampling 10 out of 200 regions in the first month. Since regions differ from month to month, we take the

[7] [Wit14] shows how to extend this mechanism to a non-binary domain.



(a) The map of sensor measurements collected in one week; the colors denote different sensor nodes (nodes are placed on trams) [JLF14]



(b) Sensing locations of one of 11 months; the colors show the intensity of pollution in the number of particles per cm^2

Figure 3.3 – Sensing locations in the Zurich dataset

closest available region to the current location of a sensor as its consecutive location. Sensors are assumed to be rational agents and their beliefs are modeled with the GP model. For each sensor, we run a separate evaluation process in which the other sensors report truthfully, while the considered sensor reports according to one of the five strategies:

- *honest*: report truthfully its measurement;
- *low*: report a low level of pollution, defined by 7500;
- *high*: report a high level of pollution, defined by 22500;
- *shift*: report its measurement shifted by 2000 in the direction of 15000, for example, if the measurement is 12000, the sensor reports 14000;
- *random*: report randomly according to its prior belief, defined as a prior probability distribution of a peer's measurement obtained from the GP model.

In all of the cases, sensors report their beliefs honestly if they are asked to provide them. Furthermore, all of the numerical values express the number of particles per cm^2. We first determine the total payoff of each sensor in the period of 11 months, which then gives us the basic statistics of the sensors' payoffs: mean, minimum, 1st quartile, median, 3rd quartile and maximum. Since the payoffs depend on randomly chosen parameters, such as the initial locations of the sensors, we repeat the process 50 times to obtain the average statistics.[8]

[8]This ensures that the averages have statistical significance for best-response strategies. In particular, t-tests show that the continuous BTS payments for truthful reporting are statistically different than the continuous BTS payments for other reporting strategies, with p-values smaller than 0.01. The same holds for the best-response strategies of the output agreement and RBTS.

Simulation results

We compare the payoffs of three different reward mechanisms: the continuous BTS, the output agreement and the RBTS for small population.

In the continuous BTS mechanism, a sensor provides its real valued measurements and its prediction regarding the measurement of the closest peer sensor. The prediction report contains the parameters of a Gaussian distribution obtained from the GP model. To score a sensor a, we select another two peers, p and $\hat{p}$, closest to the peer for which the sensor reports its prediction. This way we ensure that sensors a, p and $\hat{p}$ provide the prediction about the same event: the measurement of the peer closest to sensor a. In scoring prediction reports we use the logarithmic scoring rule defined by expression (2.1) with $\alpha = 0.01$ and $\beta = 1$, while in the information scores we use KL divergence.

In the output agreement mechanism, sensors provide only their measurements. To score a sensor a, we select its closest peer and the payoff is equal to 1 if their reports do not differ by more than 2000 (particles per cm^2), and otherwise it is 0.

In the RBTS mechanism, sensors provide their real valued measurements, but in scoring sensors, these are discretized to binary values using a predefined threshold. In particular, if the provided measurement is less than 15000, the associate binary value is equal to 0, and otherwise it is equal to 1.[9] Apart from their measurements, sensors also provide their prediction regarding the discretized measurement of their closest peers. We follow the description presented in [WP12b], with two peers being selected as in the continuous BTS mechanism. The prediction score of RBTS is scaled with $\alpha = 0.01$ ($\beta = 0$).

The statistics of the average payoffs for the continuous BTS, the output agreement, and RBTS, are shown in Table 3.1, Table 3.2, and Table 3.3, respectively. As noted in the preliminaries, these payoffs can be scaled so that they take values in an arbitrary interval.

The simulation results imply that continuous BTS is strictly proper for the considered set of strategies: all of the indicators, such as mean or median, are maximized for truthful reporting. This is not the case for the output agreement nor the RBTS mechanism. For the former one, the *shift* strategy maximizes most of the indicators, including mean and median. Namely, the strategy tends to shift the result towards the mean of the data set (15179), which for normally distributed data represents the most likely value. Therefore, it is important to correct the bias towards prior information, but this cannot be achieved unless the bias is known or it is elicited from sensors (agents), as is done in the continuous BTS.

The RBTS mechanism assigns the highest payoffs to the *low* strategy, implying that the binary discretization is too coarse-grained for the considered sensing scenario. This shows the importance of robust mechanisms that allow non-binary observations — even a simple

[9] By inspecting the data set, one can calculate that the mean of all observation values is 15179, which is very close to the selected threshold.

Table 3.1 – Average payoffs — continuous BTS

Strategy	mean	min	max	median	1st quartile	3rd quartile
honest	7.869	4.352	9.659	8.042	7.112	8.778
low	7.075	2.581	9.602	7.408	5.671	8.358
high	6.145	1.183	9.464	6.242	4.346	7.673
shift	7.389	3.895	9.476	7.645	6.342	8.287
random	6.773	3.699	9.101	6.543	5.702	7.657

Table 3.2 – Average payoffs — output agreement

Strategy	mean	min	max	median	1st quartile	3rd quartile
honest	4.114	0.68	7.78	3.76	2.18	5.1
low	0.34	0.0	1.28	0.14	0.0	0.32
high	0.794	0.0	2.74	0.26	0.1	0.84
shift	4.378	1.12	7.72	4.1	2.9	5.12
random	3.176	1.04	5.68	2.86	2.06	3.76

misreporting strategy often achieves better payoffs than honest reporting when a mechanism does not take into account the complexity of reported information.

Table 3.3 – Average payoffs — RBTS [WP12b]

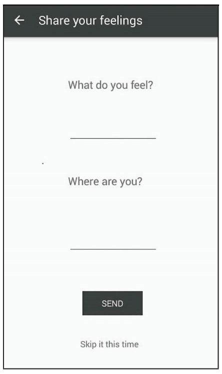
Strategy	mean	min	max	median	1st quartile	3rd quartile
honest	15.95	8.524	22.14	15.545	12.848	18.127
low	18.087	9.647	24.256	18.333	14.544	20.683
high	10.717	7.211	14.987	10.294	8.978	11.457
shift	15.13	8.104	20.958	14.863	12.22	17.149
random	15.103	8.863	20.58	14.757	12.493	16.734

3.5.2 Elicitation of subjective information[10]

One of the most basic applications of peer consistency methods is to solicit subjective judgements or answers to hypothetical questions, in which case there is no well defined criteria for evaluating the quality of the obtained report. In this section, we investigate an application of the BTS mechanisms in eliciting the reports of emotional states from crowd participants located in an area of interest, in our case, EPFL. This approach can be considered complemen-

[10]This subsection is based on the semester project 'EmoMap: Emotion Sensing of the EPFL campus', which was completed by Natalija Gucevska, under the supervision of Goran Radanovic and Boi Faltings.

tary to traditional opinion polling, which is typically performed periodically, and thus does not completely reflect the emotion profile of people at different time periods. Since emotional states carry not only information about respondents, but also their surroundings, the elicited information can be linked to different events, such as courses, and further used to provide emotion-based recommendations.



(a) *Report structure*

(b) *Selecting emotions*

(c) *Selecting location*

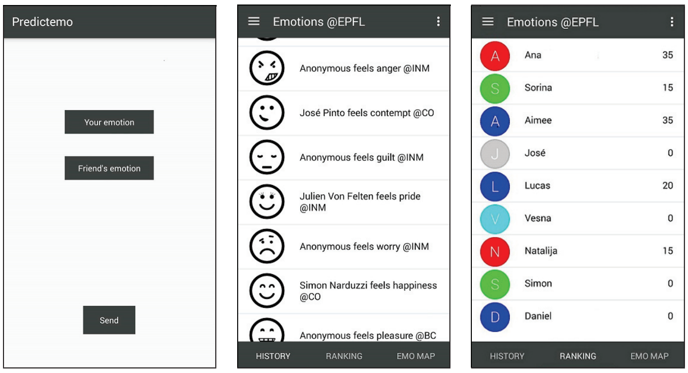
Figure 3.4 – Predictemo application

Motivated by the success of games with purposes, such as the ESP game [vAD04, vAD08], we propose an approach that has a form of a contest for subjective information and uses a BTS type of mechanism to evaluate contributions of contestants. The contest was implemented as an Android[11] application *Predictemo* that has two modes: anonymous mode, in which users anonymously provide their emotions and current locations, and game mode, in which users play the Predictemo game logged in with their Facebook[12] account. To achieve a good quality control of reported information, the underlying reward mechanism in the Predictemo game should be such that it incentivizes users to use the application in the non-anonymous (game) mode.

Figure 3.4 shows a screenshot of the Predictemo application. When a user decides to make a report, she selects one of the 20 available emotions from Geneva emotion wheel categories [Sch05, KSMP14] (Figure 3.4b) and her approximate location, i.e., one of the EPFL buildings (Figure 3.4c). Aside from reporting her emotion, a user can also see the emotions reported by other users at different locations, but this information does not contain the most recent emotions, since these are used for scoring users in the Predictemo game.

Figure 3.5 depicts the structure of the Predictemo game. When submitting a report, a player selects her emotion and location, and additionally a prediction about what her opponent reports (Figure 3.5a). The prediction constitutes of selecting the most likely emotion of the opponent and ones confidence in this prediction. Once the opponent accepts the challenge, the Predictemo game calculates the scores of players and updates the ranking list shown in

[11]www.android.com
[12]www.facebook.com



(a) *Report structure* (b) *Reported emotions* (c) *Ranking*

Figure 3.5 – Predictemo game

Figure 3.5c. The total scores in the ranking list are scaled so that they only take positive values. Players also have access to historical data, such as the list of emotions of other players at different time periods (Figure 3.5b). Notice that the data does not include the most recent emotions which are used in the scoring mechanism.

The underlying mechanism of the Predictemo game has a form of the optimal contest for simple agents, introduced in [GR14]. We focus in this section on one of its crucial components — procedure for evaluating the quality of players' contributions, i.e., the scoring mechanism. The detailed design of the Predictemo contest can be found in Section A.1 of the appendix.

Evaluating reports

The information structure that participants report in the Predictemo game consists of information and prediction reports, so an appropriate method for evaluating contributions could have a BTS structure. Instead of using the traditional BTS design, we rearrange the information and prediction scores in order to obtain a structure of the score that is easier to explain to the participants. In particular, we score an agent by how well she predicts her peer's emotion minus how well her peer predicts her emotion. However, the quality score of each prediction includes the difficulty of predicting ones emotion. Such a scoring rule induces a competition among agents, so we call the novel mechanism the competitive BTS, or simply co-BTS.

Competitive Bayesian Truth Serum. Consider two agents a_1 and a_2 that are matched with each other, and let p denote other agents, i.e., their peers. We assign the quality scores to their information and prediction reports using the following mechanism:

- The quality of agent a_1's prediction report is measured w.r.t. the information report of agent a_2, and is defined as:

$$q_{a_1}(F_{a_1}, Y_{a_2}) = \underbrace{S_Q(F_{a_1}, Y_{a_2})}_{\text{accuracy of } F_{a_1}} + \underbrace{1 - \mathbf{x}_{a_1,a_2}(Y_{a_2})}_{\text{difficulty of predicting } Y_{a_2}}$$

where S_Q is the quadratic scoring rule (2.2) with $\alpha = 1$ and $\beta = 0$, i.e.:

$$S_Q(F_{a_1}, Y_{a_2}) = F_{a_1}(Y_{a_2}) - \frac{1}{2} \cdot \sum_{z \in \mathscr{X}} F_{a_1}(z)^2$$

and $\mathbf{x}_{a_1,a_2}$ is a normalized histogram of reports from agents $\mathscr{A} \setminus \{a_1, a_2\}$:

$$\mathbf{x}_{a_1,a_2}(x) = \frac{1}{N-2} \cdot \sum_{p \in \mathscr{A} \setminus \{a_1,a_2\}} \mathbb{1}_{Y_p = x}$$

Intuitively, the quality of agent a_1's prediction depends both on how accurate the agent a_1 is and how hard it is to predict the observation of agent a_2. Similarly we define the quality of agent a_2's prediction as:

$$q_{a_2}(F_{a_2}, Y_{a_1}) = \underbrace{S_Q(F_{a_2}, Y_{a_1})}_{\text{accuracy of } F_{a_2}} + \underbrace{1 - \mathbf{x}_{a_1,a_2}(Y_{a_1})}_{\text{difficulty of predicting } Y_{a_1}}$$

- The score of agent a_1 is defined by how much better (or worse) her prediction is compared to the prediction of agent a_2:

  $$\tau_{a_1}(Y_{a_1}, F_{a_1}, Y_{a_2}, F_{a_2}, \mathbf{R}_p) = q_{a_1}(F_{a_1}, Y_{a_2}) - q_{a_2}(F_{a_2}, Y_{a_1})$$

  Similarly, for agent a_2, we have:

  $$\tau_{a_2}(Y_{a_2}, F_{a_2}, Y_{a_1}, F_{a_1}, \mathbf{R}_p) = q_{a_2}(F_{a_2}, Y_{a_1}) - q_{a_1}(F_{a_1}, Y_{a_2})$$

The competitive BTS mechanism has a decomposable score structure, quite similar to the one of the multi-valued RBTS, but with a more intuitive explanation. In particular, the quality scores of agents imply that an agent should provide her best possible prediction to maximize her score, while she cannot manipulate the quality score of another agent due to the fact that the qualities also include the difficulty of predicting ones observation. Indeed, under the relative self-dominant condition, the competitive BTS is strictly proper.

Proposition 5. *Consider $N > 2$ agents (each agent having $k = N - 1$ peers) and assume that agents do not differentiate their peers in their belief systems (e.g., $Pr(X_{a_1}|X_{a_2}) = Pr(X_{p_2}|X_{a_2})$)*[13]. *The competitive BTS is strictly proper under the relative self-dominant condition $\mathscr{C} = \{C_{RSD}\}$.*

Proof. Consider an agent a_1 whose observation is $X_{a_1} = x$, and suppose that the other agents are honest. The score assigned to an agent a_1 can be written as:

$$\tau_{a_1}(Y_{a_1}, F_{a_1}, Y_{a_2}, F_{a_2}, \mathbf{R}_p) = q_{a_1}(F_{a_1}, Y_{a_2}) - q_{a_2}(F_{a_2}, Y_{a_1})$$
$$= S_Q(F_{a_1}, Y_{a_2}) - S_Q(F_{a_2}, Y_{a_1}) + \mathbf{x}_{a_1,a_2}(Y_{a_1}) - \mathbf{x}_{a_1,a_2}(Y_{a_2})$$

[13] One can remove this requirement if condition C_{RSD} is transformed so that $Pr_{a_2}(X_{p_2}|X_{a_2})$ is changed to $Pr_{a_2}(X_{a_1}|X_{a_2})$.

$$= \underbrace{S_Q(F_{a_1}, Y_{a_2})}_{\text{prediction score}} + \underbrace{\mathbf{x}_{a_1,a_2}(Y_{a_1}) - Pr_{a_2}(X_{a_1} = Y_{a_1}|X_{a_2})}_{\text{information score}} + b$$

where b does not depend on agent a_1's report. Since the prediction score is the quadratic scoring rule we know that its optimal value is achieved when agent a_1 reports her true prediction, and due to the stochastic relevance, this is a strict optimum. Therefore, we only need to show that agent a_1 also strictly optimizes her information report by reporting honestly. By taking the expectation over other agents' observations, we obtain that the expected information score for reporting y is:

$$\mathbb{E}\left(\tau_{a_1,Y}(y, \mathbf{R}_{a_2}, \mathbf{R}_p)|X_{a_1} = x\right) = Pr_{a_1}(X_{p_1} = y|X_{a_1} = x) - \mathbb{E}_{X_{a_2}}\left(Pr_{a_2}(X_{a_1} = y|X_{a_2})|X_{a_1} = x\right)$$

Since we assumed that agent a_2 does not differentiate her peers, this gives us:

$$\mathbb{E}\left(\tau_{a_1,Y}(y, \mathbf{R}_{a_2}, \mathbf{R}_p)|X_{a_1} = x\right) = Pr_{a_1}(X_{p_1} = y|X_{a_1} = x) - \mathbb{E}_{X_{a_2}}\left(Pr_{a_2}(X_{p_2} = y|X_{a_2})|X_{a_1} = x\right)$$

The difference between the expected information scores for truthfulness, i.e., $y = x$, and misreporting, i.e., $y = z \neq x$, is equal to:

$$\begin{aligned}
&\mathbb{E}\left(\tau_{a_1,Y}(x, \mathbf{R}_{a_2}, \mathbf{R}_p)|X_{a_1} = x\right) - \mathbb{E}\left(\tau_{a_1,Y}(z, \mathbf{R}_{a_2}, \mathbf{R}_p)|X_{a_1} = x\right) \\
&= Pr_{a_1}(X_{p_1} = x|X_{a_1} = x) - Pr_{a_1}(X_{p_1} = z|X_{a_1} = x) \\
&- \mathbb{E}\left(Pr_{a_2}(X_{p_2} = x|X_{a_2}) - Pr_{a_2}(X_{p_2} = z|X_{a_2})|X_{a_1} = x\right) > 0
\end{aligned}$$

where the inequality follows from the relative self-dominant condition. This completes the proof. □

Notice that the competitive BTS does not require the common belief condition to hold. In fact, this is important because the competition should allow agents to have different proficiencies in predicting each other's emotions, which are formally modeled as observations. Furthermore, mechanism τ_a does not directly define monetary payments provided to participants, but rather scores that are used for ranking the agents.

To simplify the input for the prediction report, a player is asked to provide the most likely emotion of the other player, denoted by x_m, and ones confidence in x_m, denoted by c_{x_m}. Assuming that the player has a symmetric posterior, the mode and the confidence can be mapped into a probability distribution function over possible emotions that approximates the player's true posterior belief. To do so, we assign weights on different emotions, w_x, that approximately follow the gaussian shape:

$$w_x = \frac{1}{\sigma} \cdot e^{-\left(\frac{d(x,x_m)}{\sigma}\right)^2}$$

where σ is a decreasing function of confidence $(1/\sigma = c_{x_m}^2/(1 - c_{x_m}^2 + 0.2) + 0.001)$, and $d(x, x_m)$ represents the minimal distance of emotion x from the reported mode x_m on the Geneva emotion wheel. Once the weights are calculated, we normalize them to obtain the prediction used in the co-BTS mechanisms:

$$F_p(x) = \frac{w_x}{\sum_y w_y}$$

3.6 Conclusion

In this chapter, we explored information elicitation mechanisms that do not have access to the ground truth nor to the agents' beliefs. We showed that minimal elicitation mechanisms, which elicit only information reports, require the agents' beliefs to be highly constrained, for example, to satisfy the self-dominant condition. This led us to decomposable mechanisms that ask each agent to provide an additional report, which is separately scored from her information report. While decomposable mechanisms do provide more expressive framework, they are still not general enough to allow arbitrary beliefs constrained only with the common belief condition. Therefore, we constructed a general BTS mechanism, called the divergence-based BTS, that is strictly proper as long as the agents' posteriors are more similar when they observe the same value than when their observations are different. In the limit case, when observations are real values, the modification of the divergence-based BTS, called the continuous BTS, requires that agents have a common belief system, and we proved that this requirement is not trivial to relax. Using a community sensing testbed, we experimentally verified the importance of mechanisms that allow non-binary observations. Finally, we showed how one can use a BTS type of mechanism to evaluate the quality of reported information in a contest that follows the design principles of optimal contests.

The non-existence of a strictly proper payment mechanism when observations take real values and agents have different belief systems (Theorem 7) motivates us to further constrain the elicitation setting so that more robust properties are allowed. Therefore, in the next chapter, we turn to a different setting, more specific one, where a mechanism can extract useful statistics from statistically independent reports.

4 Multi-task peer consistency mechanisms

In this chapter, we investigate an elicitation scenario where crowd participants observe multiple *a priori* similar phenomena, with each participant observing a strict subset of them. The considered scenario models a typical crowdsourcing of micro-tasks where a bundle of tasks is assigned to a group of workers. For example, in text annotation, a requester could give 1000 sentences to a group of 100 workers, and each worker would annotate 50 sentences. By utilizing the properties of such a multi-task information elicitation, we can achieve much stronger incentive properties than the ones presented in the Chapter 3.

4.1 Formal setting

The main difference between the setting of this chapter and that of Chapter 3 is in the number of phenomena that a group of agents observe. In particular, in this chapter we assume that the center wants to elicit information about several statistically similar phenomena from the same group of agents. As depicted in Figure 4.1, each agent is assigned to a specific subset of phenomena and is asked to observe the phenomena and report her observations to the center.

By utilizing the structure of multiple phenomena, we seek to:

- allow agents to have less constrained private beliefs without necessarily reporting them;
- make uninformed reporting strategies, in which agents do not observe the phenomena, less desirable than truthful reporting.

The first property states that agents can have more diverse beliefs about each other's observations than allowed by the single-task mechanisms, while the center only requires the agents to report their private observations. The second property implies that uninformed reporting strategies are worse off for the agents. Notice that this is a stricter property than strict properness, which was discussed in Chapter 3. In particular, the property implies that truthful reporting should result in a greater expected payoff than any collusive strategy which

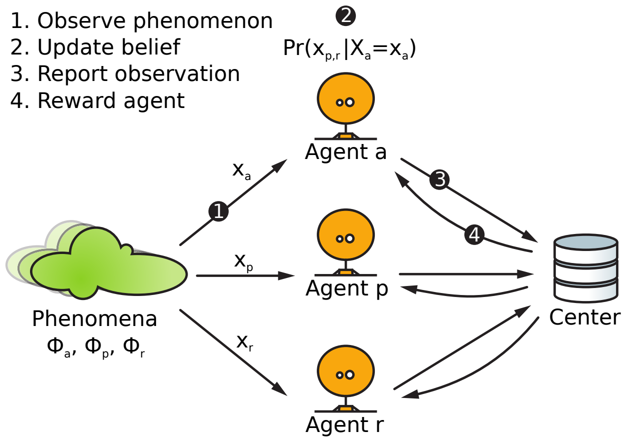


Figure 4.1 – Multi-task peer consistency setting

is not based on observations. This represents a significant improvement over single-task mechanisms, which are often susceptible to even simple collusion strategies.

We distinguish two types of agents' population: *homogeneous* and *heterogeneous*. The key difference between the two types is that the former type assumes that agents receive their private information in a similar fashion, while the latter type does not. In other words, for the former type, agents have homogeneous characteristics, although their beliefs are private, while for the latter, agents are considered to be entirely heterogeneous, both in their beliefs and in how they obtain their private information.

4.1.1 Elicitation model

To model the setting in game-theoretic terms, we follow Section 3.1.1 in Chapter 3, but with the following differences.

Instead of one, there are $M >> 1$ *a priori* similar and *statistically independent* phenomena $\Phi = \{\Phi_1, ..., \Phi_M\}$, meaning that agents distinguish them only by their observations and that the observation about phenomenon Φ_i does not contain any novel information about phenomenon Φ_j.

Each phenomenon Φ_i is observed by at least 2 different agents, randomly chosen from a large group of available agents $\mathscr{A}$ (i.e., $|\mathscr{A}| >> 1$). Unless specified differently, we will assume that an agent observes exactly one phenomenon. This assumption does not have a significant influence on the incentive properties discussed in this chapter: if it does not hold, we simply partition Φ into subsets that satisfy the assumption, and apply a reward mechanism to each subset separately. The overall reward of an agent can then be defined as an average or a sum of the obtained rewards.

Therefore, each agent a is associated with a phenomenon Φ_a. The *peers* of agent a are agents who observe the same phenomenon, while agents who observe other phenomena are called *reference* agents. A generic peer agent is denoted by p and a generic reference agent is denoted by r. We assume that agent a does not distinguish her peers nor her reference agents.[1] The number of peer agents is $k > 1$ and the number of reference agents is $K > 1$.

When an agent a observes a phenomenon Φ_i, she receives a private observation $X_a^i \in \mathcal{X}$, or simply X_a (since agent a observes a single phenomenon). The observation of a peer p is denoted by X_p, while the observation profile of k peers is denoted by $\mathbf{X}_p = (X_{p,1}, ..., X_{p,k})$. Similarly, the notation for reference agents is X_r and $\mathbf{X}_r = (X_{r,1}, ..., X_{r,K})$.

Furthermore, agent a's belief system B_a is similar to the one introduced in Section 3.1.1, and is defined by:

- *prior* belief regarding her own observation $Pr(X_a) \in \mathcal{P}$;
- *prior* belief regarding the observations of agent a's peers and references $Pr(\mathbf{X}_p, \mathbf{X}_r) \in \mathcal{P}^{k+K}$, and similarly for any subset of observations, e.g., for a single peer, we denote $Pr(X_p) \in \mathcal{P}$;
- *posterior* belief regarding the observations of agent a's peers and references $Pr(\mathbf{X}_p, \mathbf{X}_r | X_a) \in \mathcal{P}^{k+K}$, and similarly for any subset of observations, e.g., for a single peer, we denote $Pr(X_p|X_a) \in \mathcal{P}$.

The probability distributions are assumed to be *fully mixed*. Moreover, posterior belief $Pr(X_p|X_a)$ is assumed to be *stochastically relevant*. As in the previous chapter, agents a_i and a_j are allowed to have different beliefs (Pr_{a_i} and Pr_{a_j} respectively). However, it will often be clear that the beliefs are associated to a specific agent a, in which case we drop the subscript a from Pr_a (i.e., $Pr_a \to Pr$).

The statistical independence of phenomena is modeled by assuming that X_{a_1} and X_{a_2} are independent for any two agents a_1 and a_2 who observe two different phenomena. This implies that agent a's posterior belief about a reference agent $Pr(X_r|X_a)$ is equal to $Pr(X_r)$. Moreover, agent a does not distinguish agent p and agent r, implying $Pr(X_r) = Pr(X_p)$.

Finally, agents are assumed to provide only their information report Y_p, i.e., $R_a \in \mathcal{R} = \mathcal{X}$. A payment function τ_M depends on both peer reports and reference reports, i.e., $\tau_M(Y_a, \mathbf{Y}_p, \mathbf{Y}_r)$, where $\mathbf{Y}_p$ and $\mathbf{Y}_r$ are report profiles of peers and references, respectively. Formally, a *multi-task* payment mechanism is a mapping $\tau_M : \times_{i=0}^{k+K} \mathcal{R} \to \mathbb{R}$.

[1] In game-theoretic terms, this can be modeled as agents having private 'types' that are independently sampled from a common distribution. Each agent would know her type, but would not know the types of other agents.

4.1.2 Beliefs about peers

In order to establish the incentive properties of the mechanisms discussed in this chapter, we consider four conditions imposed on the agents' belief systems. The first three conditions are acquisitional constraints, while the third one is an updating constraint. One of the considered conditions is a variation of the stochastic relevance conditions defined for the acquisitional belief constraints.

Acquisitional constraints

Observation process can be directly modeled using the state model condition C_{SM} from Section 4.1.2 applied to each phenomenon Φ_i separately.

Definition 12. *Consider M random variables $\{\Omega_{\Phi_1}, ..., \Omega_{\Phi_M}\}$ taking values in $\mathbb{R}$. A belief profile $\mathbf{B}_{\mathscr{A}}$ satisfies the* multi-task state model *condition C_{MTSM} if each belief system B_a is constrained with the following set of assumptions:*

- *observation X_{a_j} of agent a_j who does not observe phenomenon Φ_i is statistically independent of Ω_{Φ_i};*
- *observations X_{a_1} and X_{a_2} of any two different agents a_1 and a_2 that observe different phenomena are statistically independent;*
- *observations X_{a_1} and X_{a_2} of any two different agents a_1 and a_2 that observe phenomenon Φ_i are conditionally independent given Ω_{Φ_i};*
- *agent a's prior belief regarding Ω_{Φ_i} is a probability density function $p_a(\Omega_{\Phi_i})$ that takes strictly positive values, and since phenomena are a priori similar, $p_a(\Omega_{\Phi_j}) = p_a(\Omega_{\Phi_i})$ for all Ω_{Φ_i} and Ω_{Φ_j};*
- *for all agents $a_i \in \mathscr{A}$, probabilities $Pr_a(X_{a_i}|\Omega_{\Phi_j})$ are strictly positive.*

To indicate that the population of agents is homogeneous, one can impose a restriction that the agents' observations are generated in a similar fashion. More precisely:

Definition 13. *Suppose that a belief profile $\mathbf{B}_{\mathscr{A}}$ satisfies the* multi-task state model *condition C_{MTSM}. Then it also satisfies the* homogeneous population *condition C_{HP} if for each belief system B_a we have that:*

$$Pr_a(X_{a_i} = x|\Omega_i = \omega) = Pr_a(X_{a_j} = x|\Omega_j = \omega)$$

for all agents a_i that observe Φ_i and a_j that observes Φ_j.

Notice that $Pr_a(X_{a_i}) = \int_{\mathbb{R}} Pr_a(X_{a_i}|\Omega_i = \omega) \cdot p_a(\Omega_i = \omega) d\omega$. Therefore, the homogenous population condition implies that agent a believes her observation to be a priori similar to the observation of any peer or reference agent: $Pr_a(X_p) = Pr_a(X_a)$. More generally, for any two agents a_i and a_j, $Pr_a(X_{a_i} = x) = Pr_a(X_{a_j} = x)$. Furthermore, for an agent a_i who observed Φ_i and her peer p_i, we have that:

$$\begin{aligned} Pr_a(X_{p_i}, X_{a_i}) &= \int_{\mathbb{R}} Pr_a(X_{p_i}, X_{a_i}|\Omega_i = \omega) \cdot p_a(\Omega_i = \omega) d\omega \\ &= \int_{\mathbb{R}} Pr_a(X_{p_i}|\Omega_i = \omega) \cdot Pr_a(X_{a_i}|\Omega_i = \omega) \cdot p_a(\Omega_i = \omega) d\omega \end{aligned}$$

where the last equality is due to the conditional independence of X_{a_i} and X_{p_i} given Ω_i. By the homogenous population condition, it follows that $Pr_a(X_{p_i}, X_{a_i}) = Pr_a(X_{p_j}, X_{a_j})$ (where a_j and p_j observe Φ_j), which gives us:

$$Pr_a(X_{p_i}|X_{a_i}) = \frac{Pr_a(X_{p_i}, X_{a_i})}{Pr_a(X_{a_i})} = \frac{Pr_a(X_{p_j}, X_{a_j})}{Pr_a(X_{a_j})} = Pr_a(X_{p_j}|X_{a_j})$$

More specifically, $Pr_a(X_{p_i}|X_{a_i}) = Pr_a(X_p|X_a)$. That is, if agent a knows the observation of agent a_i (who observed a different phenomenon), then she has the same belief about the peer of agent a_i as she has about her own peer when she observes the same value. Notice that the identities $Pr_a(X_p) = Pr_a(X_a)$ and $Pr_a(X_{p_i}|X_{a_i}) = Pr_a(X_p|X_a)$ play an important role in our analysis.

Finally, we consider a restriction that is similar to the stochastic relevance condition, but slightly more restrictive. It states that an observation x is statistically different than any linear combination of other observations.[2]

Definition 14. *Suppose that a belief profile $\mathbf{B}_{\mathscr{A}}$ satisfies the* multi-task state model *condition C_{MTSM}. Then it also satisfies the* linear separability *condition C_{LS} if each belief system B_a additionally satisfies:*

$$KL(p_a(\Omega_i|X_{a_i} = x) || \sum_{z \in \mathscr{X} \setminus \{x\}} w_z \cdot p_a(\Omega_i|X_{a_i} = z)) > 0$$

for all $w_z \geq 0$ such that $\sum_{z \in \mathscr{X} \setminus \{x\}} w_z = 1$, where an agent a_i observes Φ_i.

Notice that in the above conditions $Pr_a(X_{a_i}|\Omega_{\Phi_j})$ and $Pr_{a'}(X_{a_i}|\Omega_{\Phi_j})$ (and similarly $p_a(\Omega_{\Phi_i})$ and $p_{a'}(\Omega_{\Phi_i})$) are allowed to be different for two different agents a and a'. However, it is often clear which agent we refer to; in that case we drop the subscript a from $Pr_a(X_{a_i}|\Omega)$ and $p_a(\Omega)$.

Updating constraint

In a more general case, agents are heterogeneous in a way they observe their private information. While the multi-task state model condition C_{MTSM} allows agents to be heterogenous in a way they acquire private information, we will show that additional constraints are needed

[2] A similar condition is adopted by [CM85] in extracting the full surplus from efficient allocation as revenue.

in order to achieve desirable incentive properties. Therefore, we consider a belief updating restriction equivalent to the self-predicting condition C_{SP} introduced in Section 3.1.5, but defined for a multi-task setting.

Definition 15. *A belief profile* $\mathbf{B}_{\mathscr{A}}$ *satisfies the* multi-task self-predicting *condition* C_{MTSP} *if for the posterior belief of each* B_a*, we have that:*[3]

$$\frac{Pr(X_p = y|X_a = x)}{Pr(X_p = y)} - 1 < \frac{Pr(X_p = x|X_a = x)}{Pr(X_p = x)} - 1, \forall y \neq x \tag{4.1}$$

while $Pr(X_r = y|X_a = x) = Pr(X_r = y)$.

As noted in Section 3.1.5, the self-prediction holds in the common case where an agent believes that only the observation she endorses is more likely among her peers than was expected by her prior belief. This includes binary answer spaces (as in [DG13]), as well as a more general case when agents observe different samples drawn from the same categorical distribution, but with unknown parameters sampled from a Dirichlet distribution.

We characterize the degree of correlation that an agent a believes to be possible between different observation values x and y by the *self-predictor* δ_{SP}:

Definition 16. *For a belief profile* $\mathbf{B}_{\mathscr{A}}$ *that satisfies the multi-task self-predicting condition* C_{MTSP}*, we define* self-predictor δ_{SP} *as the smallest number in* $[0,1]$ *so that for each belief system* B_a

$$\frac{Pr(X_p = y|X_a = x)}{Pr(X_p = y)} - 1 < \left(\frac{Pr(X_p = x|X_a = x)}{Pr(X_p = x)} - 1\right) \cdot \delta_{SP}, \forall y \neq x \tag{4.2}$$

holds.

The smaller the self-predictor is, the more distinguishable different observations are. For example, $\delta_{SP} = 0$ indicates that different observation values are not correlated, in a sense that observing a value x should decrease the belief that a peer agent observes different value y. For $\delta_{SP} \approx 1$, on the other hand, agents are more likely to confuse two similar observations.

4.1.3 Reporting strategies

To examine stronger incentive properties than those analyzed in Chapter 3, we extend the agents' strategy space to include both uninformed and misreporting strategy profiles. In particular, an agent a now faces a choice between two basic strategies:

- *informed reporting*: observe a phenomenon to obtain observation $X_a = x$, but report $Y_a = y$, where y is randomly sampled from a probability distribution that depends on

[3] We keep -1 on both sides to make the proofs and the notion of the *self predictor* clear.

observation X_a. An example strategy is when the agent is honest, in which case the distribution function is of the form $\mathbb{1}_{y=x}$, where $\mathbb{1}_{cond}$ is an indicator variable equal to 1 when condition $cond$ is satisfied, and otherwise is 0. A more general informed reporting strategy would be when the distribution function is $\mathbb{1}_{y=\hat{\sigma}(x)}$, where $\hat{\sigma}$ is a bijective function $\hat{\sigma}: \mathcal{X} \to \mathcal{X}$. We refer to the strategies of this form as *permutation reporting* strategies.

- *uninformed reporting*: not observe a phenomenon and report according to a fixed probability distribution function. The *uninformed reporting* strategy includes both random reporting and collusive strategies in which agents agree in advance on reporting the same value.

Furthermore, notice that an *uninformed reporting* strategy can be modeled using an *informed reporting* strategy by equating all of the probability distributions from which Y_a is sampled. This means that we can assume agent a always makes an observation, but does not need to base her reporting strategies on it.

Therefore, a generic strategy of agent a can be expressed with a set of probability distribution functions $Q_a = \{Q_{a,x} \in \mathcal{P} | x \in \mathcal{X}\}$ that defines how Y_a is generated from agent a's observation X_a. In particular, for an observation $X_a = x$, Y_a is a sample from a probability distribution function $Q_{a,x}$. We denote a superset of all possible Q_a by $\mathcal{Q}$ and a set of probability distribution functions associated to the honest strategy by Q_a^{honest}, i.e., $Q_{a,x}^{honest}(y) = \mathbb{1}_{y=x}$.

Since agent a does not distinguish her peers and reference agents, from an agent a's perspective, they have the same (expected) behaviour. Therefore, we can assume that peers and references have a symmetric strategy profile, so that $Q_{p_1} = ... = Q_{p_k} = Q_{r_k} = ... = Q_{r_K}$.

We now turn to the elicitation properties we investigate in this chapter, and the first property we define is properness of a multi-task payment mechanism, similarly to how it is defined for the single-task payment mechanisms.

Definition 17. *We say that a multi-task payment mechanism τ_M is* proper *under the set of belief constraints $\mathcal{C}$ if for all $\mathbf{B}_a \in \mathcal{B}(\mathcal{C})$, $a \in \mathcal{A}$, and $Q_a \in \mathcal{Q}_a \backslash \{Q_a^{honest}\}$, we have that:*

$$\mathbb{E}_{\mathbf{X}_p,\mathbf{X}_r}\left(\tau_M(X_a,\mathbf{X}_p,\mathbf{X}_r)|X_a\right) \geq \mathbb{E}_{Y_a,\mathbf{X}_p,\mathbf{X}_r}\left(\tau_M(Y_a,\mathbf{X}_p,\mathbf{X}_r)|X_a\right) \tag{4.3}$$

If the inequality is strict, then τ_M is strictly *proper.*

As argued in the previous chapter, the strict properness property is closely related to the ex-post subjective equilibrium concept, in particular, it implies that truthful reporting is an ex-post subjective equilibrium under the considered belief constraints. In this chapter, we show how to achieve this property under weaker constraints on agents' belief systems.

A stricter property is to require that agents cannot profit by colluding. Clearly, without any knowledge on how the agents acquire their observations, the center cannot, in general, make permutation strategy profiles strictly worse than honest reporting. Namely, if honest reporting would result in strictly greater expected payoffs for one observation acquisition process, then, by permuting the observations, we could define another acquisition process for which the corresponding permutation reporting would result in greater expected payoffs than honest reporting. Notice, however, that all of the permutation reporting strategies require that the agents make observations and coordinate on their observations. Except for honest reporting, the latter typically induces some cost. Therefore, our goal is to make honest reporting at least as good as any other permutation reporting strategy and strictly better than any non-permutation strategy. This is captured by a property called *strong truthfulness* [DG13, SAFP16], which we define for the setting of this chapter as follows.

Definition 18. *We say that a multi-task payment mechanism* τ_M *is* strongly truthful *under belief constraints* $\mathscr{C}$ *if for all* $\mathbf{B}_a \in \mathscr{B}(\mathscr{C})$, $a \in \mathscr{A}$, $Q_a \in \mathscr{Q}$, *and* $Q_p \in \mathscr{Q}$, *we have that:*

$$\mathbb{E}_{\mathbf{X}_p,\mathbf{X}_r}\left(\tau_M(X_a,\mathbf{X}_p,\mathbf{X}_r)|X_a\right) \geq \mathbb{E}_{Y_a,\mathbf{Y}_p,\mathbf{Y}_r}\left(\tau_M(Y_a,\mathbf{Y}_p,\mathbf{Y}_r)|X_a\right) \tag{4.4}$$

where the equality holds only if Q_a *and* Q_p *represent the same* permutation reporting *strategy.*

Strong truthfulness implies strict properness in a sense that any strongly truthful mechanism is also strictly proper. The converse is not true. Therefore, we minimally require that a payment mechanism is strictly proper.

Remark 5. The strong truthfulness property implies collusive resistance in situation when agents distinguish phenomena only by the values they observe. As pointed out by [GWL16], if agents have alternative ways of distinguishing phenomena, there might exists collusive strategy profiles that are more profitable than truthful reporting. Although such a collusive behaviour has been studied on a peer grading data set [SP16b], it is not clear whether these types of misreporting strategies would be indeed adopted in practice. We leave further analysis on how to cope with this type of collusion for future work.

4.2 Related work

One of the first methods developed for a multi-task peer consistency is the peer consistency mechanism of [DG13], which we refer to as the Dasgupta&Ghosh mechanism. The setting that it operates in allows agents to be heterogeneous, both in their beliefs and in the way they acquire their private signal. The mechanism assumes that observations are binary signals related to an objective information, while the agents' proficiencies in obtaining the true value are bounded from below. Under this assumption, the Dasgupta&Ghosh mechanism is strongly truthful. The acquisitional heterogeneity is reflected through the fact that agents are allowed to have different proficiencies. However, as argued later in the chapter and independently shown by [SAFP16], the Dasgupta&Ghosh mechanism requires that the categorical property holds, i.e., that the increase from prior to posterior only happens for the observed value. This property is not trivially satisfied when observations are non-binary.

In contrast, [WP13b] analyze a minimal peer consistency setting where agents are homogeneous in the way they observe their private information, although their beliefs are subjective (private). In particular, they apply the shadowing approach on empirical frequencies sampled from independent phenomena to achieve strict properness for a homogenous population.

We develop on these mechanisms and extend them to settings with a non-binary observation space. A host of recent results on the multi-task peer consistency methods is closely related to our work. [KSM$^+$15] investigate the same two settings as we do, and derive two strongly-truthful mechanism, one for each of the settings. Both mechanisms resemble the robust peer truth serum developed in this chapter, but are analyzed only in the limit case when the number of phenomena is large. The *correlated agreement* mechanism of [SAFP16] adopts the principles of the Dasgupta&Ghosh mechanism and generalizes it to a non-binary setting. The mechanism is not necessarily strongly truthful nor strictly proper. Instead, it is *informed-truthful*, which, in the case of a binary effort model, means that the mechanism provides strict incentives for *high effort*, but agents are not *strictly* incentivized to report truthfully their observations.[4] Furthermore, the mechanism assumes a partial knowledge of agents' belief structure, which enables it to achieve its incentive properties using a small number of tasks. The authors, however, describe an alternative approach that learns this partial knowledge given a large enough set of tasks. [KS16b] provide a framework that one can use to derive a number of results on strong-truthfulness, including the proof that the Dasgupta&Ghosh mechanism is strongly-truthful. Moreover, they show how to extend the Dasgupta&Ghosh mechanism to a non-binary setting, assuming that agents solve a large number of a priori similar elicitation tasks.[5] [SP16a] use replicator dynamics to demonstrate that the equilibrium selection is important if agents are learning how to play over time. This result is supported by the empirical study of [GMCA14], which experimentally showed that people can learn to play uninformed reporting strategies if they result in higher payoffs.

Finally, we also mention the empirical work related to the peer consistency incentives designed for two applications: peer grading and community sensing [ASC$^+$10, KHKZ08]. While several references propose the peer consistency mechanisms for massive open online courses (MOOCs) [DG13, KSM$^+$15, dASP16], to our knowledge, there has been very little work done on evaluating different mechanisms in a peer grading scenario. The most systematic approach is taken in [SAFP16, SP16a], where the authors study a structure of a MOOCs dataset and using a replicator dynamics argued which mechanisms are potentially the most suitable for the peer grading in MOOCs. Furthermore, [SP16b] raise practical concerns relevant for applying peer consistency mechanisms in peer grading. By simulating several peer consistency mechanisms

[4] Informed truthfulness is a weaker notion than strong truthfulness, but as argued by [SAFP16], it is of a practical importance, because informed-truthful mechanisms make uninformed reporting strategies strictly worse of than informed ones. One of the important future steps would be to see whether a similar claim can be made for a non-binary effort model.

[5] Both the learning algorithm of [SAFP16] and the mechanism of [KS16b] assume that an agent solves a large (enough) number of a priori similar tasks. Thus, the allowed heterogeneity by these mechanisms does not contradict Proposition 6, which provides an impossibility result of a proper knowledge-free elicitation for a heterogeneous population of agents.

on a MOOCs data set, they conclude that the gain from exerting effort in peer grading might be relatively low due to frequent disagreement between peers. In contrast, we perform an on-campus experiment that compares the performance of one of the mechanisms developed in this chapter to the performance of mechanisms with a simpler structure. There is also a growing literature that studies other aspects in peer grading and takes orthogonal approaches to achieve a better quality of peer grades (e.g., [PHC$^+$13, KWL$^+$13, WTL15, WDK$^+$15]). For example, estimating the reliability of peer graders and correcting for their biases [PHC$^+$13] or improving the grading accuracy by providing the peer graders with a feedback about their grading biases [WTL15]. In regard to community sensing, there is a vast literature addressing different issues, such as the optimal sensor placement [KSG08] or privacy and trustworthiness of sensors [DBFH09, CRKH11, SK13]. One of the peer consistency methods proposed for information elicitation in community sensing setting is the *peer truth serum* (PTS) [FLJ14], and we study how to modify its design to make it more robust.

4.3 Homogeneous population

We first consider a homogeneous population of agents where the observations of different agents are formed in a similar fashion. That is, the set of belief constraints $\mathscr{C}$ contains the homogeneous population condition C_{HP}

4.3.1 Minimal peer prediction with private priors

We start by demonstrating that the elicitation of private observations can be done with a payment mechanism that asks agents to provide their information reports. To do so, we define proxy events linked to the observations of other agents whose probabilities are the same as those that would be reported in a prediction report, and use these to construct an expression that an agent expects to be the same as the quadratic scoring rule. The mechanism acts as the classical peer prediction [MRZ05], but instead of directly transforming an agent's report to her posterior belief, it appropriately samples reports of reference agents in order to obtain a term that is in expectation equal to the agent's posterior.

Mechanism

Consider an agent a_1 that observes a phenomenon Φ_1. Once agent a_1 acquires observation $X_{a_1} = x$, she updates her belief regarding the observation of her *peer* agent p_1 to $Pr(Y_{p_1}|X_{a_1} = x)$. If agent a_1 believes that the other agents are honest, $Pr(X_{p_1}|X_{a_1} = x)$ is also her belief about the report of peer p_1.

Now, consider another phenomenon $\Phi_2 \neq \Phi_1$ observed by a reference agent r_2. Agent a_1's belief about the observation of agent r_2 is $Pr(X_{r_2})$ because agent a_1 does not observe Φ_2. However, agents obtain their private observations in a similar fashion. Therefore, if agent a_1 knows that a proxy agent a_{proxy2}, associated to phenomenon Φ_2, has observed $X_{proxy2} = y$, her belief about agent r_2's observation changes from $Pr(X_{r_2})$ to $Pr(X_{p_1}|X_{a_1} = y)$. Namely, agents acquire their observation in the same way, so agent a_1 should believe that $Pr(X_{r_2}|X_{a_{proxy2}} = y)$ is equal to $Pr(X_{p_1}|X_{a_1} = y)$. This means that the indicator variable $\mathbb{1}_{X_{r_2}=z}$ is in expectation equal to $Pr(X_{p_1} = z|X_{a_1} = y)$ whenever the proxy for phenomenon Φ_2 observes y.

Suppose that r_2 is honest, i.e., $Y_{r_2} = X_{r_2}$, and that the observation of honest proxy a_{proxy2} is equal to agent a_1's report, i.e. $Y_{a_{proxy}} = X_{a_{proxy}} = Y_{a_1} = y$. Then, the indicator variable $\mathbb{1}_{Y_{r_2}=z}$ is in expectation equal to agent a_1's belief $Pr(X_{p_1} = z|X_{a_1} = y)$, which would make up her prediction report regarding her peer's observation.

The idea is to arrange indicators $\mathbb{1}_{Y_{r_j}=z}$ so that they correspond to the quadratic scoring rule (see (2.2) in Chapter 2), in which prediction $Pr(X_{p_1}|Y_{a_1})$ is scored by how well it predicts the report of peer p_1. Provided that the peer is honest, the expected score is maximized when the prediction is equal to $Pr(X_{p_1}|X_{a_1})$, which implies truthfulness of agent a_1.

Minimal Peer Prediction with Private Priors. Let agent a_1 and her peer p_1 be an arbitrary agents that report their observation Y_{a_1} and Y_{p_1} regarding the same phenomenon, here denoted by Φ_1.[6] The mechanism has the following structure:

1. Randomly sample one response for all phenomena $\Phi_2 \neq \Phi_1$ that are not observed by agent a_1. We denote this sample by Σ and we call it *double-mixed* if it contains all possible values from $\mathscr{X}$ at least twice.

2. If sample Σ is not double-mixed, agent a_1's score is equal to $\tau_M(Y_{a_1}, Y_{p_1}, \mathbf{Y}_r) = 0$, where $\mathbf{Y}_r$ are the reports of all the reference agents of agent a_1 (in this case, all agents that have not observed Φ_1).

3. Otherwise, take two different phenomena $\Phi_2 \neq \Phi_3 \neq \Phi_1$ whose Σ samples are equal to Y_{a_1}, and randomly select another sample for each of them to obtain two responses Y_{r_2} and Y_{r_3}. Finally, the score of an agent a_1 is equal to:

$$\tau_M(Y_{a_1}, Y_{p_1}, \mathbf{Y}_r) = \frac{1}{2} + \mathbb{1}_{Y_{r_2}=Y_{p_1}} - \frac{1}{2}\sum_{z\in\mathscr{X}} \mathbb{1}_{Y_{r_2}=z} \cdot \mathbb{1}_{Y_{r_3}=z}$$

Notice that the last step of the mechanism is only applied when Σ is double-mixed, and this is important to prevent a potential bias towards the more likely observations. Namely, if the fourth step is executed whenever Σ contains two reports equal to report Y_{a_1}, which is sufficient for calculating score $\tau_M(Y_{a_1}, Y_{p_1}, \mathbf{Y}_r)$ in this step, agent a_1 might report dishonestly in the hope of increasing the probability of getting non-zero payoff.

Figure 4.2 shows one possible outcome of the elicitation process for a binary evaluation space $\{x, y\}$. To score agent a_1, the minimal peer prediction first builds Σ sample based on the reports

[6] Subscripts 1, 2, and 3 are chosen for clarity; they can be replaced by any three different numbers from 1 to M.

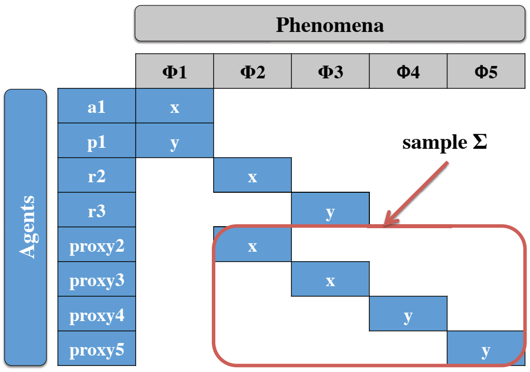


Figure 4.2 – Minimal peer prediction with private priors.

of the proxy agents. Since Σ is double-mixed, the mechanism acquires the reports of agents r_2 and r_3, that, together with the report of agent p_1, define agent a_1's score. In this case, the (ex post) score of agent a_1 is equal to $\frac{1}{2}$.

Properties

To illustrate the principle of the minimal peer prediction, consider an agent a_1 with observation equal to X_{a_1}. Her belief about the observation of her peer is $Pr(X_{p_1}|X_{a_1})$, while her belief about the observation of a reference agent is $Pr(X_r)$. The mechanism works as follows. If Σ is not double-mixed — which happens with probability strictly less than 1 for $|\Sigma| \geq 2 \cdot |\mathscr{X}|$ — agent a_1's reward is 0. Otherwise, the mechanism searches in Σ for two phenomena Φ_2 and Φ_3 whose Σ samples are equal to Y_{a_1}. Since agent a_1 knows that samples of Φ_2 and Φ_3 are equal to Y_{a_1}, and all agents make observations in a similar way, agent a_1 updates her belief about the other observations of Φ_2 and Φ_3: $Pr(X_r) \to Pr(X_{p_1}|Y_{a_1})$. This means that a_1's belief about the observations of agent r_2 (who observes Φ_2 and whose report is not in Σ) and agent r_3 (who observes Φ_3 and whose report is not in Σ) is equal to $Pr(X_{p_1}|Y_{a_1})$.

Furthermore, the indicators $\mathbb{1}_{r_2=z}$ and $\mathbb{1}_{r_3=z}$ in score $\tau_M(Y_{a_1}, Y_{p_1}, \mathbf{Y}_r)$ are in expectation equal to $Pr(Y_{p_1} = z|Y_{a_1})$. Therefore, assuming that the agents other than a_1 are honest and that sample Σ is double-mixed, the score is in expectation equivalent to the quadratic scoring rule $S_Q(Pr(X_{p_1}|Y_{a_1}), X_{p_1})$ whose expected value is maximized for $Y_{a_1} = X_{a_1}$. More formally:

Theorem 8. *The minimal peer prediction with private priors is strictly proper under the homogenous population condition, i.e., $\mathscr{C} = \{C_{HP}\}$, whenever Σ sample contains at least two times more elements than observation set $\mathscr{X}$, i.e., $|\Sigma| \geq 2 \cdot |\mathscr{X}|$.*

Proof. Consider an agent a_1 whose evaluation is equal to $X_{a_1} = x$, and suppose other agents are honest, including an agent p_1. Due to the independence of X_{r_2} and X_{r_3} and linearity of expectations, the expected score of agent a_1 for reporting $\tilde{x}$ when her peer reports y is equal to:

$$\bar{\tau}_M(\tilde{x}, y) = p_\Sigma \cdot \left[\frac{1}{2} + Pr(X_{p_1} = y | X_{a_1} = \tilde{x}) - \frac{1}{2} \cdot \sum_{z \in \mathscr{X}} Pr(X_{p_1} = z | X_{a_1} = \tilde{x})^2 \right]$$

where p_Σ is the probability that Σ is double-mixed. Fully mixed priors and $|\Sigma| \geq 2 \cdot |\mathscr{X}|$ imply that $p_\Sigma > 0$, so $\bar{\tau}_M(\tilde{x}, y)$ has the structure of the quadratic scoring rule (see (2.2) in Chapter 2), scaled by p_Σ, that rewards agent a_1's posterior beliefs $Pr(X_{p_1} | X_{a_1} = \tilde{x})$ with the realization of the outcome specified by peer p_1's report. Since the quadratic scoring rule is in expectation maximized when agent a_1 reports her true belief $Pr(X_{p_1} = y | X_{a_1} = x)$, agent a_1 is incentivized to report honestly her observation, i.e., $\tilde{x} = x$. Moreover, agent a_1 is strictly incentivized to do so because of the stochastic relevance of her posterior beliefs. □

To score an agent, the mechanism requires $2 \cdot |\mathscr{X}|$ statistically similar phenomena in addition to the phenomenon rated by the agent. Often, the number of phenomena that the mechanism wants to monitor is significantly larger than observation space $\mathscr{X}$, as in product reviewing, where the number of ratings is relatively small, e.g., 5, while many products have statistically similar features, i.e., they are a priori similar.

While the mechanism represents a possibility result for knowledge-free information elicitation, it does not satisfy strong-truthfulness. We, therefore, turn to an alternative approach, which assumes that the population of agents is large.

4.3.2 Logarithmic peer truth serum

Instead of sampling reports to design a score that acts as the classical peer prediction, one can score agents by the statistical significance of their reports. This principle comes from the (original) Bayesian truth serum (BTS) [Pre04], but we apply it in a different manner. Unlike the original BTS mechanism, the novel mechanism does not require agents to have a common prior belief and is minimal, in a sense that sensors report only their observations.

Mechanism

To determine the statistical significance, we first sample reports across different phenomena and make the normalized histogram $\mathbf{x}_\Phi$ of reported values. That is, for each possible observation value x, we evaluate the fraction of reports in the sample that are equal to x. Second, we calculate the normalized histogram $\mathbf{x}_{\Phi_i}$ of reports for each phenomena Φ_i. The statistical significance of a report equal to x associated to phenomenon Φ_i is then defined as $\log \frac{\mathbf{x}_{\Phi_i}(x)}{\mathbf{x}_\Phi(x)}$.

Logarithmic Peer Truth Serum. Consider an agent who observes Φ_i and whose report is equal to $Y_a = x$. The logarithmic peer truth serum (log-PTS) applies the following steps to reward the agent:

- Calculate two empirical frequencies:
  - Frequency of reports equal to x among agent a's peers:

$$\mathbf{x}_{\Phi_i}(x) = \frac{1}{k} \cdot \sum_{i=1}^{k} \mathbb{1}_{Y_{p_i}=x}$$

  - Frequency of reports equal to x among reference agents $(\hat{r}_1, ..., \hat{r}_{M-1})$ that are not each other's peers nor peers of agent a:

$$\mathbf{x}_{\Phi}(x) = \frac{1}{M-1} \cdot \sum_{i=1}^{M-1} \mathbb{1}_{Y_{\hat{r}_i}=x}$$

    To obtain reports from the reference agents $(\hat{r}_1, ..., \hat{r}_{M-1})$, one can randomly sample a report for each phenomenon, except the one observed by agent a.

- Finally, reward agent a for reporting $Y_a = x$ with:

$$\tau_M(Y_a, \mathbf{Y}_p, \mathbf{Y}_r) = \log \frac{\mathbf{x}_{\Phi_i}(x)}{\mathbf{x}_{\Phi}(x)}$$

To avoid potential issues with 0 values in $\mathbf{x}_{\Phi_i}$ and $\mathbf{x}_{\Phi}$ histograms, one can apply Laplace (additive) smoothing with small smoothing parameters, or simply include the report of agent a in both histograms. The latter would make the score equal to 0 when $\mathbf{x}_{\Phi_i}(Y_a) = 0$ and $\mathbf{x}_{\Phi}(Y_a) = 0$. In our analysis, we use the convention of setting the score to 0 when $\mathbf{x}_{\Phi_i}(Y_a) = 0$ and $\mathbf{x}_{\Phi}(Y_a) = 0$.

Properties

Strict-properness. Consider an agent a whose belief systems satisfies the homogenous population condition C_{HP}, and assume that the other agents are honest. It can be shown that the logarithm of ratio $\frac{\mathbf{x}_{\Phi_i}(y)}{\mathbf{x}_{\Phi}(y)}$ from agent a's perspective converges to:

$$\lim_{k,M\to\infty} \log \frac{\mathbf{x}_{\Phi_i}(y)}{\mathbf{x}_{\Phi}(y)} \underset{a.s.}{=} \log \frac{Pr(X_a = y|\Omega_{\Phi_i})}{Pr(X_a = y)}$$

where y is the agent a's report. Using Bayes' rule we obtain:

$$\lim_{k,M\to\infty} \log \frac{\mathbf{x}_{\Phi_i}(y)}{\mathbf{x}_{\Phi}(y)} \underset{a.s.}{=} \log \frac{p(\Omega_{\Phi_i}|X_a = y)}{p(\Omega_{\Phi_i})} = \log p(\Omega_{\Phi_i}|X_a = y) + b$$

where b does not depend on report y. The score has one indicative feature: agent a is scored based on how well it predicts state Ω_{Φ_i}. More precisely, $\log \frac{\mathbf{x}_{\Phi_i}(y)}{\mathbf{x}_{\Phi}(y)}$ is related to the logarithmic scoring rule (see (2.1) in Chapter 2) applied on the posterior belief of an agent whose observation is equal to y. The true belief of agent a is $p(\Omega_{\Phi_i}|X_a = x)$, where x is her observation, so in order to be scored with her true belief, the agent should report $y = x$. Since the logarithmic scoring rule incentivizes agents to report their true beliefs, we obtain that log-PTS provides proper incentives. To show this formally, we first prove the following lemma.

Lemma 4. *Consider an agent a who observes $X_a = x$. As $k \to \infty$ and $M \to \infty$, the agent's payment in the logarithmic peer truth serum is maximized under the homogenous population condition ($\mathscr{C} = \{C_{HP}\}$) if and only if she reports y such that:*

$$p(\Omega_{\Phi_i}|X_a = x) = p(\Omega_{\Phi_i}|Y_p = y)$$

Proof. Reports from histogram $\mathbf{x}_{\Phi_i}$ are conditionally independent given Ω_{Φ_i}, so we can apply the law of large numbers to obtain:

$$\lim_{k\to\infty} \mathbf{x}_{\Phi_i}(x) \underset{a.s.}{=} Pr(Y_p = x|\Omega_{\Phi_i})$$

where we used the fact that peers are homogenous ($Pr(X_{p_i} = z|\Omega_{\Phi_i}) = Pr(X_{p_j} = z|\Omega_{\Phi_i})$) and have a symmetric strategy profile.

Next, consider reference agents $(\hat{r}_1, ..., \hat{r}_{M-1})$. The reports of agents in $(\hat{r}_1, ..., \hat{r}_{M-1})$ are statistically independent, which implies that:

$$\lim_{M\to\infty} \mathbf{x}_{\Phi}(x) \underset{a.s.}{=} Pr(Y_p = x)$$

due to the properties of the homogenous population condition C_{HP} and the fact that peers and references have a symmetric strategy profile.

Therefore, the expected score of agent a who observed $X_a = x$ for reporting $Y_a = y$ is equal to:

$$\begin{aligned}
&\lim_{M\to\infty, k\to\infty} \int_{\mathbb{R}} p(\Omega_{\Phi_i} = \omega|X_a = x) \cdot \log \frac{\mathbf{x}_{\Phi_i}(y)}{\mathbf{x}_{\Phi}(y)} d\omega \\
&\quad \underset{a.s.}{=} \int_{\mathbb{R}} p(\Omega_{\Phi_i} = \omega|X_a = x) \cdot \log \frac{Pr(Y_p = z|\Omega_{\Phi_i} = \omega)}{Pr(Y_p = z)} d\omega \\
&\quad = \int_{\mathbb{R}} p(\Omega_{\Phi_i} = \omega|X_a = x) \cdot \log \frac{p(\Omega_{\Phi_i} = \omega|Y_p = y)}{p(\Omega_{\Phi_i} = \omega)} d\omega
\end{aligned}$$

where the last equality is due to Bayes' rule (the term inside the logarithm). The equation can be further reduced to:

$$\begin{aligned}
&\int_{\mathbb{R}} p(\Omega_{\Phi_i} = \omega|X_a = x) \cdot \log \frac{p(\Omega_{\Phi_i} = \omega|Y_p = y) \cdot p(\Omega_{\Phi_i} = \omega|X_a = x)}{p(\Omega_{\Phi_i} = \omega) \cdot p(\Omega_{\Phi_i} = \omega|X_a = x)} d\omega \\
&= \int_{\mathbb{R}} p(\Omega_{\Phi_i} = \omega|X_a = x) \cdot \log \frac{p(\Omega_{\Phi_i} = \omega|Y_p = y)}{p(\Omega_{\Phi_i} = \omega|X_a = x)} d\omega \\
&+ \int_{\mathbb{R}} p(\Omega_{\Phi_i} = \omega|X_a = x) \cdot \log \frac{p(\Omega_{\Phi_i} = \omega|X_a = x)}{p(\Omega_{\Phi_i} = \omega)} d\omega \\
&= -KL(p(\Omega_{\Phi_i}|X_a = x)||p(\Omega_{\Phi_i}|Y_p = y)) + KL(p(\Omega_{\Phi_i}|X_a = x)||p(\Omega_{\Phi_i}))
\end{aligned}$$

Notice that the only part that depends on the agent's report is $KL(p(\Omega_{\Phi_i}|X_a = x)||p(\Omega_{\Phi_i}|Y_p = y))$. The expected payoff is negative in the KL divergence, so the best case for the agent is when its value is equal to 0, which occurs if and only if $p(\Omega_{\Phi_i}|X_a = x) = p(\Omega_{\Phi_i}|Y_p = y)$. □

Theorem 9. *The logarithmic peer truth serum is strictly proper under the homogenous belief condition ($\mathscr{C} = \{C_{HP}\}$) as $k \to \infty$ and $M \to \infty$.*

Proof. Suppose that the peers and the reference agents are honest ($Y_p = X_p$) and let $p(\Omega_{\Phi_i}|X_a = x) = p(\Omega_{\Phi_i}|Y_p = y) = p(\Omega_{\Phi_i}|X_p = y)$. The homogeneity condition implies that $p(\Omega_{\Phi_i}|X_p = y) = p(\Omega_{\Phi_i}|X_a = y)$. Therefore, $p(\Omega_{\Phi_i}|X_a = x) = p(\Omega_{\Phi_i}|X_a = y)$, which gives us:

$$\begin{aligned} Pr(X_p|X_a = x) &= \int_{\mathbb{R}} Pr(X_p|\Omega_{\Phi_i} = \omega, X_a = x) \cdot p(\Omega_{\Phi_i} = \omega|X_a = x) d\omega \\ &= \int_{\mathbb{R}} Pr(X_p|\Omega_{\Phi_i} = \omega) \cdot p(\Omega_{\Phi_i} = \omega|X_a = x) d\omega \\ &= \int_{\mathbb{R}} Pr(X_p|\Omega_{\Phi_i} = \omega) \cdot p(\Omega_{\Phi_i} = \omega|X_a = y) d\omega \\ &= \int_{\mathbb{R}} Pr(X_p|\Omega_{\Phi_i} = \omega, X_a = y) \cdot p(\Omega_{\Phi_i} = \omega|X_a = y) d\omega \\ &= Pr(X_p|X_a = y) \end{aligned}$$

where we used the conditional independence of X_p and X_a given Ω_{Φ_i}. By the stochastic relevance condition, it follows that $y = x$, which, by Lemma 4, implies the strict properness of log-PTS. □

Strong-truthfulness. Suppose now that the agents report according to a strategy which prescribes that for an observation equal to x or y, the report is x, while for any other observation, the report is truthful. In this case, the logarithm of ratio $\frac{\mathbf{x}_{\Phi_i}(x)}{\mathbf{x}_{\Phi}(x)}$ converges to:

$$\lim_{k,M \to \infty} \log \frac{\mathbf{x}_{\Phi_i}(x)}{\mathbf{x}_{\Phi}(x)} \underset{a.s.}{=} \log \frac{Pr(X_a \in \{x, y\}|\Omega_{\Phi_i})}{Pr(X_a \in \{x, y\})} = \log \frac{p(\Omega_{\Phi_i}|X_a \in \{x, y\})}{p(\Omega_{\Phi_i})}$$

where the last equality follows from Bayes' rule. The part that is dependent on x is equal to $\log p(\Omega_{\Phi_i}|X_a = \{x, y\})$. Therefore, if agent a reports x, she will get scored by the logarithmic scoring rule applied on the belief of an agent who cannot distinguish between x and y. On the other hand, if agent a's observation is x, her true belief is $p(\Omega_{\Phi_i}|X_a = x)$. In other words, her score is expected to be suboptimal, unlike in the case when everyone is honest. More generally, log-PTS is strongly truthful for a large population of agents under the homogeneous population and linear separability constraints. To show this, we first prove a lemma which states that the agents should adopt the same reporting strategy when they use permutation reporting.

Lemma 5. *In the logarithmic peer truth serum, under the homogenous population condition ($\mathscr{C} = \{C_{HP}\}$), agent a's strictly best response to a permutation reporting strategy is to report according to it, provided that $k \to \infty$ and $M \to \infty$.*

Proof. Let $\hat{\sigma}$ be a bijective function that defines the permutation reporting strategy of an agent a's peers and references, i.e., $Q_{p,x}(z) = \mathbb{1}_{z=\hat{\sigma}(x)}$. Log-PTS rewards report y with the score of report $\hat{\sigma}(y)$ when the peers and the references are honest. Due to Theorem 9, agent a's best response to truthful reporting of the other agents is to report X_a. Therefore, when the other agents report according to $\hat{\sigma}$, agent a's best response is $\hat{\sigma}(X_a)$, which implies that the agent should use the same reporting strategy as her peers and reference agents. □

Theorem 10. *The logarithmic peer truth serum is strongly truthful under the belief constraints* $\mathscr{C} = \{C_{HP}, C_{LS}\}$ *as* $k \to \infty$ *and* $M \to \infty$.

Proof. Consider an agent a who uses strategy defined by Q_a and believes that her peers and reference agents adopt a strategy defined by Q_p. Let us rewrite the condition of Lemma 4 as:

$$p(\Omega_{\Phi_i}|X_a = x) = p(\Omega_{\Phi_i}|Y_p = y) = \frac{Pr(Y_p = y|\Omega_{\Phi_i})}{Pr(Y_p = y)} \cdot p(\Omega_{\Phi_i})$$

where we applied Bayes' rule. Due to the homogeneity condition C_{HP}, we know that:

$$\begin{aligned}
Pr(Y_p = y|\Omega_{\Phi_i}) &= \sum_{z\in\mathscr{X}} Q_{p,z}(y) \cdot Pr(X_p = z|\Omega_{\Phi_i}) = \sum_{z\in\mathscr{X}} Q_{p,z}(y) \cdot Pr(X_a = z|\Omega_{\Phi_i}) \\
Pr(Y_p = y) &= \sum_{z\in\mathscr{X}} Q_{p,z}(y) \cdot Pr(X_p = z) = \sum_{z\in\mathscr{X}} Q_{p,z}(y) \cdot Pr(X_a = z)
\end{aligned}$$

which, together with the previous expression and Bayes' rule, gives us:

$$\begin{aligned}
p(\Omega_{\Phi_i}|X_a = x) &= \sum_{z\in\mathscr{X}} \frac{Q_{p,z}(y)}{\sum_{\hat{z}\in\mathscr{X}} Q_{p,\hat{z}}(y) \cdot Pr(X_a = \hat{z})} \cdot Pr(X_a = z|\Omega_{\Phi_i}) \cdot p(\Omega_{\Phi_i}) \\
&= \sum_{z\in\mathscr{X}} \frac{Q_{p,z}(y) \cdot Pr(X_a = z)}{\sum_{\hat{z}\in\mathscr{X}} Q_{p,\hat{z}}(y) \cdot Pr(X_a = \hat{z})} \cdot p(\Omega_{\Phi_i}|X_a = z)
\end{aligned}$$

By setting $w'_z = \frac{Q_{p,z}(y) \cdot Pr(X_a = z)}{\sum_{\hat{z}\in\mathscr{X}} Q_{p,\hat{z}}(y) \cdot Pr(X_a = \hat{z})}$, we obtain:

$$p(\Omega_{\Phi_i}|X_a = x) = \sum_{z\in\mathscr{X}} w'_z \cdot p(\Omega_{\Phi_i}|X_a = z)$$

Notice that w'_z slightly differs from w_z in the linear separability condition C_{LS} as it includes also w'_x. If $w'_x < 1$, we can define $w_z = \frac{w'_z}{1-w'_x}$ for $z \neq x$, so that $w_z \geq 0$, $\sum_{z\in\mathscr{X}\setminus\{x\}} w_z = 1$ and:

$$p(\Omega_{\Phi_i}|X_a = x) = \sum_{z\in\mathscr{X}\setminus\{x\}} w_z \cdot p(\Omega_{\Phi_i}|X_a = z)$$

Due to the linear separability condition C_{LS}, this cannot be the case, which implies that $w_{x'} = 1$. Therefore:

$$\frac{Q_{p,x}(y) \cdot Pr(X_a = x)}{\sum_{\hat{z}\in\mathscr{X}} Q_{p,\hat{z}}(y) \cdot Pr(X_a = \hat{z})} = 1$$

$$\frac{Q_{p,z}(y) \cdot Pr(X_a = z)}{\sum_{\hat{z} \in \mathscr{X}} Q_{p,\hat{z}}(y) \cdot Pr(X_a = \hat{z})} = 0, \forall z \neq x$$

Since this has to hold for all possible observations x, we have that $Q_{p,x}(y) = 1$. In other words, peers and reference agents should use a permutation reporting strategy for the condition in Lemma 4 to hold. Due to Lemma 5, agent a should also use the same strategy (i.e., $Q_a = Q_p$) in order for the condition to hold, which by Lemma 4, implies the strong truthfulness of log-PTS. Hence, we proved the statement. □

The main drawback of log-PTS is that it requires a large population of agents and, in particular, a large number of peers. We therefore investigate an alternative approach which has more stable payments as the number of peers decreases, and thus is more robust in terms of the population size.

4.4 Heterogeneous population

Unlike the previous section, a group of agents is now considered to have heterogeneous characteristics, i.e., agents differ in the way they observe phenomena. In the context of the multi-task state model condition C_{MTSM}, this would mean that, for two different agents a_1 and a_2, $Pr(X_{a_1}|\Omega_{\Phi_i})$ and $Pr(X_{a_2}|\Omega_{\Phi_i})$ are allowed to be different. We show, however, that it is not possible to create strict incentives that would elicit both agent a_1's and agent a_2's private signals under the set of constraints that contains only the multi-task state model condition C_{MTSM}. The intuition behind this result is that an agent might believe she is special in the way she observes two possible observations x and y. In particular, she might believe that the other agents swap these two observations.

Proposition 6. *There exists no strictly proper mechanism τ_M under the multi-task state model condition, i.e., $\mathscr{C} = \{C_{MTSM}\}$.*

Proof. Consider two agents a_1 and a_2 that observe phenomenon Φ_i. Assume agent a_1 believes that the observation of any of her peers, X_{p_1}, is statistically similar to her observation:

$$Pr_{a_1}(X_{p_1} = x|\Omega_{\Phi_i}) = Pr_{a_1}(X_{a_1} = x|\Omega_{\Phi_i}), \forall x$$

while agent a_2 believes that the observation of any of her peers, X_{p_2}, is a sample from the distribution:

$$Pr_{a_2}(X_{p_2} = x|\Omega_{\Phi_i}) = Pr_{a_2}(X_{a_2} = y|\Omega_{\Phi_i})$$
$$Pr_{a_2}(X_{p_2} = y|\Omega_{\Phi_i}) = Pr_{a_2}(X_{a_2} = x|\Omega_{\Phi_i})$$
$$Pr_{a_2}(X_{p_2} = z|\Omega_{\Phi_i}) = Pr_{a_2}(X_{a_2} = z|\Omega_{\Phi_i}), \forall z \neq x, y$$

Furthermore, let $Pr_{a_1}(X_{a_1} = z|\Omega_{\Phi_i}) = Pr_{a_2}(X_{p_2} = z|\Omega_{\Phi_i})$ and $Pr_{a_1}(X_{r_1} = z|\Omega_{\Phi_j}) = Pr_{a_2}(X_{r_2} = z|\Omega_{\Phi_k})$ for all evaluations z, where r_1 and r_2 are reference agents of a_1 and a_2, respectively.

By applying the properties of the multi-task state model condition C_{MTSM}, we obtain from the above conditions that:

$$\begin{aligned} Pr_{a_1}(\mathbf{X}_{p_1}, \mathbf{X}_{r_1}|X_{a_1} = x) = Pr_{a_2}(\mathbf{X}_{p_2}, \mathbf{X}_{r_2}|X_{a_2} = y) \\ Pr_{a_1}(\mathbf{X}_{p_1}, \mathbf{X}_{r_1}|X_{a_1} = y) = Pr_{a_2}(\mathbf{X}_{p_2}, \mathbf{X}_{r_2}|X_{a_2} = x) \end{aligned} \tag{4.5}$$

Now, suppose that there exists a strictly proper mechanism τ_M and let $X_{a_1} = X_{a_2} = x$. Strict properness implies:

$$\sum_{\mathbf{X}_{p_1}, \mathbf{X}_{r_1}} Pr_{a_1}(\mathbf{X}_{p_1}, \mathbf{X}_{r_1}|X_{a_1} = x) \cdot \tau_M(x, \mathbf{X}_{p_1}, \mathbf{X}_{r_1}) > \sum_{\mathbf{X}_{p_1}, \mathbf{X}_{r_1}} Pr_{a_1}(\mathbf{X}_{p_1}, \mathbf{X}_{r_1}|X_{a_1} = x) \cdot \tau_M(y, \mathbf{X}_{p_1}, \mathbf{X}_{r_1})$$

Since the formal setting assumes that the agents do not differentiate their peers nor reference agents, by applying identity (4.5), we obtain:

$$\sum_{\mathbf{X}_{p_2}, \mathbf{X}_{r_2}} Pr_{a_2}(\mathbf{X}_{p_2}, \mathbf{X}_{r_2}|X_{a_2} = y) \cdot \tau_M(x, \mathbf{X}_{p_2}, \mathbf{X}_{r_2}) > \sum_{\mathbf{X}_{p_2}, \mathbf{X}_{r_2}} Pr_{a_2}(\mathbf{X}_{p_2}, \mathbf{X}_{r_2}|X_{a_2} = y) \cdot \tau_M(y, \mathbf{X}_{p_2}, \mathbf{X}_{r_2})$$

which contradicts the assumption that τ_M is strictly proper. □

Instead of imposing a restriction on how agents acquire their private information, we now take an alternative approach and put constraints on their belief updating process. More precisely, the condition we assume to hold is the multi-task self-prediction C_{MTSP}, which allows the agents to have a limited heterogeneity in observing the phenomena.

4.4.1 Robust peer truth serum

The mechanism we consider combines the ideas from [DG13, WP13b] with the peer truth serum (PTS) introduced in [JF11, FLJ14, FPTJ14], and we call it the robust peer truth serum (RPTS). The idea behind the mechanism is to use the distribution of reported values from different phenomena as the prior probability of possible observations, and scale the reward given for agreement between agents with this distribution. This solves the major issue with the PTS mechanism as presented in [FPTJ14], which is that the prior distribution had to be known.

Mechanism

The peer truth serum (PTS) [JF11, FPTJ14] rewards an agent who reports answer x only if a randomly chosen peer reported the same observation. It uses a commonly known prior probability distribution $\mathbf{x}(x)$ over possible values x, and rewards a matching report x with $\frac{1}{\mathbf{x}(x)}$.

Instead of using a predefined value for $\mathbf{x}$, we construct it from the reports of agents who observe different phenomena. Since $\mathbf{x}$ is calculated from a finite number of samples, it is possible that $\mathbf{x}(x)$ is equal to 0 for a certain report x, which would lead to an ill-defined score due to the division by 0. To overcome this problem, we distinguish values x for which statistic $\mathbf{x}(x)$ is equal to 0. When $\mathbf{x}(x) \neq 0$, an agent who reports x obtains a score proportional to $\frac{1}{\mathbf{x}(x)} - 1$ if her peer has also reported x, and a score proportional to -1 in any other case. Otherwise, if $\mathbf{x}(x) = 0$, an agent who reports x obtains 0, since there is no peer that matches the agent's report.

Robust Peer Truth Serum. Consider an agent a who observes phenomenon Φ_i and whose report is equal to $Y_a = x$. The robust peer truth serum (RPTS) rewards the agent using the following steps:

- Randomly sample M reports from M different phenomena, including the phenomenon Φ_i, but not agent a's report.
- Calculate the frequency of reported values equal to x within this sample:

  $$\mathbf{x}_a(x) = \frac{num(x)}{\sum_{y \in \mathcal{X}} num(y)}$$

  where num is the function that counts occurrences of reported values in the sample.
- Reward agent a for reporting $Y_a = x$ with:

  $$\tau_M(Y_a, \mathbf{Y}_p, \mathbf{Y}_r) = \begin{cases} \frac{\mathbb{1}_{x=x_p}}{\mathbf{x}_a(x)} - 1 & \text{if } \mathbf{x}_a(x) \neq 0 \\ 0 & \text{if } \mathbf{x}_a(x) = 0 \end{cases} \tag{4.6}$$

  where x_p is the report of agent a's peer, who observes phenomenon Φ_i and whose report is in the sample from which $\mathbf{x}_a$ was obtained.

The RPTS payment function uses multiple reference agents, but unlike log-PTS, it requires only one peer. This represents a significant improvement in the numerical stability of the score. If we can assign multiple peers p to agent a, which is the case when more than one peer agent observes Φ_i, the final score of agent a can be the average of the RPTS over all selected peers. Due to the linearity of the expectation, the transformed score has the same incentive properties. However, using multiple reports reduces the variance in payments which may often be desirable.

Properties

Although RPTS is a nonlinear scheme, the expected score can be expressed in a closed form.

Lemma 6. *The expected payment to an agent a with observation $X_a = x$ and report $Y_a = y$ in* *the RPTS mechanism is equal to:*

$$\begin{cases} \left(\frac{Pr(Y_p=y|X_a=x)}{Pr(Y_p=y)} - 1\right) \cdot (1-(1-Pr(Y_p = y))^{M-1}) & if\, Pr(Y_p = y) > 0 \\ 0 & if\, Pr(Y_p = y) = 0 \end{cases} \tag{4.7}$$

Proof. To make the proof notationally clear, let us denote $P_{p|a}(y|x) = Pr(Y_p = y|X_a = x)$ and $P_p(y) = Pr(Y_p = y)$.

It is clear that for $P_p(y) = 0$, the expected payment to agent a is 0. When $P_p(y) > 0$, the expected payment to agent a is:

$$\underbrace{\prod_{r_i|X_{r_i}=y} P_{r_i}(y) \prod_{r_i|X_{r_i}\neq y} (1-P_{r_i}(y)) \cdot P_{p|a}(y|x) \cdot \frac{1}{(num(X_{r_i} = y)+1)/M}}_{due\ to\ \frac{\mathbb{1}_{x_p=y}}{\mathbf{x}_a(y)}\ part\ when\ \mathbf{x}_a(y)\neq 0}$$
$$- \underbrace{\left(1-(1-P_{p|a}(y|x))\prod_{r_i}(1-P_{r_i}(y))^{M-1}\right)}_{due\ to\ -1\ part\ when\ \mathbf{x}_a(y)\neq 0}$$

Since agent a does not distinguish her workers, $P_p(y) = P_{r_k}(y)$. Let i be the number of reports equal to y from reference agents. We have:

$$\left[\underbrace{\sum_{i=0}^{M-1}\binom{M-1}{i} \cdot P_p(y)^i \cdot (1-P_p(y))^{M-1-i} \cdot P_{p|a}(y|x) \cdot \frac{1}{(i+1)/M}}_{due\ to\ \frac{\mathbb{1}_{x_p=y}}{\mathbf{x}_a(y)}\ part\ when\ \mathbf{x}_a(y)\neq 0}\right.$$
$$\left. - \underbrace{\left(1-(1-P_{p|a}(y|x)) \cdot (1-P_p(y))^{M-1}\right)}_{due\ to\ -1\ part\ when\ \mathbf{x}_a(y)\neq 0}\right]$$
$$= \left[P_{p|a}(y|x) \cdot M \cdot \sum_{i=0}^{M-1} \frac{(M-1)!}{i!(M-1-i)!} \cdot P_p(y)^i \cdot (1-P_p(y))^{M-1-i} \cdot \frac{1}{i+1}\right.$$
$$\left. -1 + (1-P_{p|a}(y|x)) \cdot (1-P_p(y))^{M-1}\right]$$
$$= \left[\frac{P_{p|a}(y|x)}{P_p(y)} \cdot \sum_{i=0}^{M-1} \frac{(M-1+1)!}{(i+1)!(M-1-i)!} \cdot P_p(y)^{i+1} \cdot (1-P_p(y))^{(M-1+1)-(i+1)}\right.$$
$$\left. -1 + (1-P_{p|a}(y|x)) \cdot (1-P_p(y))^{M-1}\right]$$
$$= \left[\frac{P_{p|a}(y|x)}{P_p(y)} \cdot \sum_{i=1}^{M}\binom{M-1+1}{i} \cdot P_p(y)^i \cdot (1-P_p(y))^{(M-1+1)-i}\right.$$
$$\left. -1 + (1-P_{p|a}(y|x)) \cdot (1-P_p(y))^{M-1}\right]$$

$$= \left[\frac{P_{p|a}(y|x)}{P_p(y)} \cdot \sum_{i=1}^{M} \binom{M}{i} \cdot P_p(y)^i \cdot (1 - P_p(y))^{M-i} \right.$$

$$\left. - 1 + (1 - P_{p|a}(y|x)) \cdot (1 - P_p(y))^{M-1} \right]$$

$$= \left[\frac{P_{p|a}(y|x)}{P_p(y)} \cdot \left(1 - (1 - P_p(y))^M\right) - 1 + (1 - P_{p|a}(y|x)) \cdot (1 - P_p(y))^{M-1} \right]$$

$$= \left[\frac{P_{p|a}(y|x)}{P_p(y)} \cdot \left(1 - (1 - P_p(y))^{M-1} + P_p(y) \cdot (1 - P_p(y))^{M-1}\right) \right.$$

$$\left. - (1 - (1 - P_p(y))^{M-1}) - P_{p|a}(y|x) \cdot (1 - P_p(y))^{M-1} \right]$$

$$= \left(\frac{P_{p|a}(y|x)}{P_p(y)} - 1 \right) \cdot \left(1 - (1 - P_p(y))^{M-1}\right)$$

Hence, we proved the statement. □

Strict-properness. We now give an intuitive understanding of the RPTS mechanism, assuming a scenario where there are many phenomena M in the mechanism.

To decide on her best strategy, an agent a should estimate the reward she can expect for reporting observation y. From Lemma 6, the agent's payoff converges towards $\frac{Pr(X_p=y|X_a=x)}{Pr(X_p=y)} - 1$ provided that her observation is x and that the other agents are honest. Intuitively, $\mathbf{x}_a(y)$ approximates prior $Pr(X_p = y)$ as the agent has only prior information about her reference agents, while the probability of matching the peer agent is equal to $Pr(X_p = y|X_a = x)$. If agent a's beliefs satisfy the self-predicting condition C_{MTSP}, i.e., $\frac{Pr(X_p=y|X_a=x)}{Pr(X_p=y)} < \frac{Pr(X_p=x|X_a=x)}{Pr(X_p=x)}$ for $y \neq x$, then her payoff is strictly maximized for $y = x$.

A more formal analysis, including cases with a smaller number of phenomena, is given in Theorem 11. In particular, the theorem gives a condition on the number of phenomena M that the center should use in rewarding each agent. The condition is related to the belief systems of the agents, and is expressed through self-predictor δ_{SP}.

Theorem 11. *The robust peer truth serum is strictly proper under the multi-task self-predicting condition ($\mathscr{C} = \{C_{MTSP}\}$) if the number of phenomena M in the mechanism satisfies:*

$$\frac{1 - (1 - Pr_a(X_p = x))^{M-1}}{1 - Pr_a(X_p = x)^{M-1}} \geq \delta_{SP} \tag{4.8}$$

for all $x \in \mathscr{X}$ and all agents $a \in \mathscr{A}$.

Proof. From Lemma 6, it follows that the expected payoff of agent a, with observation $X_a = x$

and report $Y_a = y$, is:

$$\begin{aligned}\left(\frac{Pr(X_p = y|X_a = x)}{Pr(X_p = y)} - 1\right)&\cdot(1-(1-Pr(X_p = y))^{M-1})\\ &\leq \left(\frac{Pr(X_p = y|X_a = x)}{Pr(X_p = y)} - 1\right)\cdot(1-Pr(X_p = x)^{M-1})\\ &< \left(\frac{Pr(X_p = x|X_a = x)}{Pr(X_p = x)} - 1\right)\cdot(1-(1-Pr(X_p = x))^{M-1})\end{aligned}$$

where the first inequality follows from $Pr(X_p = y) + Pr(X_p = x) \leq 1$ and the second inequality follows from (4.1) (the self-predicting condition) and (4.8). We see that the maximal expected payoff when the task is solved is achieved when agent a reports her true observation. □

Strong-truthfulness. As has been done for the log-PTS mechanism, suppose now that the agents report according to a strategy which prescribes that the report is x for an observation equal to x or y, while for any other observation, the report is truthful. From Lemma 6, the payoff for reporting x converges towards $\frac{Pr(X_p=x|X_a=x)+Pr(X_p=y|X_a=x)}{Pr(X_p=x)+Pr(X_a=y)} - 1$. Namely, in this case, $\mathbf{x}_a(y)$ approximates prior $Pr(X_p = x) + Pr(X_p = y)$ as both agents who observe x and y report x. Similarly, the probability of matching is $Pr(X_p = x|X_a = x) + Pr(X_p = y|X_a = x)$. Let us rearrange the expected payoff:

$$\begin{aligned}&\frac{Pr(X_p = x|X_a = x) + Pr(X_p = y|X_a = x)}{Pr(X_p = x) + Pr(X_a = y)} - 1\\ &= \frac{Pr(X_p = x|X_a = x) - Pr(X_p = x) + Pr(X_p = y|X_a = x) - Pr(X_a = y)}{Pr(X_p = x) + Pr(X_a = y)}\\ &= \frac{Pr(X_p = x)\cdot\left(\frac{Pr(X_p=x|X_a=x)}{Pr(X_p=x)} - 1\right) + Pr(X_p = y)\cdot\left(\frac{Pr(X_p=y|X_a=x)}{Pr(X_p=y)} - 1\right)}{Pr(X_p = x) + Pr(X_a = y)}\\ &= \left(\frac{Pr(X_p = x|X_a = x)}{Pr(X_p = x)} - 1\right)\cdot\frac{Pr(X_p = x) + Pr(X_p = y)\cdot b}{Pr(X_p = x) + Pr(X_a = y)}\end{aligned} \tag{4.9}$$

where we put:

$$b = \frac{\frac{Pr(X_p=y|X_a=x)}{Pr(X_p=y)} - 1}{\frac{Pr(X_p=x|X_a=x)}{Pr(X_p=x)} - 1}$$

If agent a's beliefs satisfy the self-predicting condition C_{MTSP}, i.e., $\frac{Pr(X_p=z|X_a=x)}{Pr(X_p=z)} < \frac{Pr(X_p=x|X_a=x)}{Pr(X_p=x)}$ for $z \neq x$, then it has to be that $b < 1$, which implies that expression (4.9) is strictly smaller than $\frac{Pr(X_p=x|X_a=x)}{Pr(X_p=x)} - 1$. This, on the other hand, is the expected payoff of agent a when everybody is honest. More generally, RPTS is strongly truthful if the agents' belief systems are compatible with the self-predicting condition C_{MTSP} and the number of tasks exceeds a threshold dependent on the self-predictor δ_{SP}. To show this formally, we first prove an equivalent result to that of Lemma 5 for log-PTS: agents should adopt the same reporting strategy when they use permutation reporting.

Lemma 7. *Suppose that agent a's belief system B_a satisfies the multi-task self-predicting condition ($C = \{C_{MTSP}\}$) and condition* (4.8)*. Then, in the robust peer truth serum, agent a's strictly best response to a permutation reporting strategy is to report according to it.*

Proof. The proof is equivalent to the one for Lemma 5. Let $\hat{\sigma}$ be a bijective function that defines the permutation reporting strategy of an agent a's peers and references, i.e., $Q_{p,x}(z) = \mathbb{1}_{z=\hat{\sigma}(x)}$. RPTS assings to report y a score equal to the score of report $\hat{\sigma}(y)$ when the peers and the references are honest. Due to Theorem 11, agent a's best response to truthful reporting of the other agents is to report X_a. Therefore, when the other agents report according to $\hat{\sigma}$, agent a's best response is $\hat{\sigma}(X_a)$, which implies that the agent should use the same reporting strategy as her peers and reference agents. □

Theorem 12. *The robust peer truth serum is strongly truthful under the multi-task self-predicting condition ($C = \{C_{MTSP}\}$) if the number of phenomena M in the mechanism satisfies:*

$$\left(1 - (M-1) \cdot Pr_a(X_p = x) \cdot \frac{(1 - Pr_a(X_p = x))^{M-2}}{1 - (1 - Pr_a(X_p = x))^{M-1}}\right) \geq \delta_{SP} \tag{4.10}$$

for all $x \in \mathcal{X}$ and all agents a.

Proof. Consider an agent a that observes $X_a = x$ and suppose her peers and reference agents use a strategy profile defined by distribution Q_p. The agent's payoff for reporting y is equal to:

$$\left(\frac{\sum_{z \in \mathcal{X}} Q_{p,z}(y) \cdot Pr(X_p = z | X = x)}{\sum_{z \in \mathcal{X}} Q_{p,z}(y) \cdot Pr(X_p = z)} - 1\right) \cdot \left(1 - \left(1 - \sum_{z \in \mathcal{X}} Q_{p,z}(y) \cdot Pr(X_p = z)\right)^{M-1}\right)$$

$$= \left(\frac{Q_{p,x}(y) \cdot Pr(X_p = x | X_a = x) + \sum_{z \in \mathcal{X} \setminus \{x\}} Q_{p,z}(y) \cdot Pr(X_p = z | X_a = x)}{Q_{p,x}(y) \cdot Pr(X_p = x) + \sum_{z \in \mathcal{X} \setminus \{x\}} Q_{p,z}(y) \cdot Pr(X_p = z)} - 1\right)$$
$$\cdot \left(1 - \left(1 - Q_{p,x}(y) \cdot Pr(X_p = x) - \sum_{z \in \mathcal{X} \setminus \{x\}} Q_{p,z}(y) \cdot Pr(X_p = z)\right)^{M-1}\right) \tag{4.11}$$

where we used Lemma 6 to calculate the expected payoff.

Let us simplify our notation with the following substitutions: $p = Pr(X_p = x)$, $\Delta p = Pr(X_p = x | X_a = x) - Pr(X_p = x)$, $q = \sum_{z \in \mathcal{X} \setminus \{x\}} Q_{p,z}(y) \cdot Pr(Y_p = z)$, $\Delta q = \sum_{z \in \mathcal{X} \setminus \{x\}} Q_{p,z}(y) \cdot (Pr(X_p = z | X_a = x) - Pr(X_p = z))$. Expression (4.11) can be written in the new notation as:

$$\frac{Q_{p,x}(y) \cdot \Delta p + \Delta q}{Q_{p,x}(y) \cdot p + q} \cdot \left(1 - (1 - Q_{p,x}(y) \cdot p - q)^{M-1}\right)$$

We will prove the statement by showing that under the conditions of the theorem:

1. An optimum of the expected payoff is achieved when $Q_{p,x}(y) = 1$, regardless of q.

2. When $q > 0$ the payoff is lower than when $q = 0$.

Part 1:

Notice that the self prediction implies the existence of $\epsilon > 0$ such that $\frac{\Delta p}{p} \cdot (1-\epsilon) = \frac{\Delta q}{q}$. Therefore, we have:

$$\begin{aligned}
&\frac{Q_{p,x}(y) \cdot \Delta p + \Delta q}{Q_{p,x}(y) \cdot p + q} \cdot \left(1 - (1 - Q_{p,x}(y) \cdot p - q)^{M-1}\right) \\
&= \frac{\frac{\Delta p}{p} \cdot Q_{p,x}(y) \cdot p + \frac{\Delta q}{q} \cdot q}{Q_{p,x}(y) \cdot p + q} \cdot \left(1 - (1 - Q_{p,x}(y) \cdot p - q)^{M-1}\right) \\
&= \frac{\Delta p}{p} \cdot \left(1 - \frac{\epsilon \cdot q}{Q_{p,x}(y) \cdot p + q}\right) \cdot \left(1 - (1 - Q_{p,x}(y) \cdot p - q)^{M-1}\right) \qquad (4.12)
\end{aligned}$$

The optimum of (4.12) is achieved for $Q_{p,x}(y) = 1$, regardless of q.

Part 2:

We also need to show that the agents whose evaluations are $z \neq x$ lower the value of expression (4.11) when $Q_{p,z}(y) > 0$. Consider a function of $\lambda \in [0,1]$:

$$\begin{aligned}
f(\lambda) &= \frac{\Delta p + \lambda \cdot \Delta q}{p + \lambda \cdot q} \cdot (1 - (1 - p - \lambda \cdot q)^{M-1}) = (\Delta p + \lambda \cdot \Delta q) \cdot (\sum_{i=0}^{M-2} (1 - p - \lambda \cdot q)^i) \\
&= \Delta p \cdot (1 + \frac{\Delta q}{q \cdot \Delta p} \cdot \lambda \cdot q)(\sum_{i=0}^{M-2} (1 - p - \lambda \cdot q)^i) \qquad (4.13)
\end{aligned}$$

where the second equality is due to (A.1) (see the appendix). For $\lambda = 1$, function f corresponds to expression (4.11) with $Q_{p,x}(y) = 1$. It suffices to show that function f is strictly decreasing, meaning that the optimal value is obtained when $\lambda = 0$. Since this trivially follows when $\Delta q \leq 0$, in the remaining part of the proof we only consider the case when $\Delta q > 0$. Due to the fully mixed beliefs, $p + q = 1$ implies $Q_{p,z}(y) = 1$ for all $z \in \mathscr{X}$, which further implies $\Delta q = \sum_{z \in \mathscr{X} \setminus \{x\}} (Pr(X_p = z | X_a = x) - Pr(X_p = z)) = (1 - Pr(X_p = x | X_a = x)) - (1 - Pr(X_p = x)) = -\Delta p < 0$. This means that for $\Delta q > 0$, we have $p + \lambda \cdot q \leq p + q < 1$. The partial derivative of f w.r.t. λ is equal to:

$$\frac{\partial f}{\partial \lambda}(\lambda) = \Delta q \cdot (\sum_{i=0}^{M-2} (1 - p - \lambda \cdot q)^i) - (\Delta p + \lambda \cdot \Delta q) \cdot (\sum_{i=1}^{M-2} i \cdot q \cdot (1 - p - \lambda \cdot q)^{i-1})$$

Due to (4.13), condition (4.10) (which implies $\frac{\Delta q}{q \cdot \Delta p} < \frac{1}{p}$), $p + \lambda \cdot q < 1$ and Lemma 9 (in Appendix), a sufficient condition for f to be strictly decreasing is that $\frac{\partial f}{\partial \lambda}(\lambda) < 0$ for $\lambda = 0$. We have:

$$\frac{\partial f}{\partial \lambda}(0) = \Delta q \cdot (\sum_{i=0}^{M-2}(1-p)^i) - q \cdot \Delta p \cdot (\sum_{i=1}^{M-2} i(1-p)^{i-1})$$

Expressions (A.1) and (A.2) imply:

$$\frac{\partial f}{\partial \lambda}(0) = \Delta q \cdot \frac{1-(1-p)^{M-1}}{p} - q \cdot \Delta p \cdot \frac{(1-(1-p)^{M-1}) - p \cdot (M-1) \cdot (1-p)^{M-2}}{p^2}$$

Therefore, $\frac{\partial f}{\partial \lambda}(0) < 0$ whenever:

$$\frac{\Delta q}{q} < \frac{\Delta p}{p} \cdot \left(1 - (M-1) \cdot p \cdot \frac{(1-p)^{M-2}}{1-(1-p)^{M-1}}\right)$$

Since $\frac{\Delta q}{q} \leq \max_{z \neq x} \frac{Pr(X_p = z | X_a = x)}{Pr(X_p = z)}$ and condition (4.10) holds, we conclude that $\frac{\partial f}{\partial \lambda}(0) < 0$. This means that, from agent a's perspective, the optimal value of $Q_{p,z}(y)$ for $z \neq x$ is 0.

Conclusion:

From Part 1 and Part 2 of the proof, we conclude that agent a strictly maximizes her payoff when the other agents adopt a permutation reporting strategy. Now, notice that condition (4.10) is stricter than condition (4.8) (see below Lemma 8). Therefore, by Lemma 7, the strictly optimal choice for agent a is to adopt the same permutation reporting strategy, which proves that the RPTS mechanism is strongly truthful. □

Condition (4.10) is stricter than condition (4.8) in a sense that any self-predictor δ_{SP} that satisfies (4.10) necessarily satisfies (4.8).

Lemma 8. *If self-predictor δ_{SP} satisfies condition* (4.10)*, then it also satisfies condition* (4.8).

Proof. Suppose δ_{SP} satisfies condition (4.10). We have that:

$$\begin{aligned}
\delta_{SP} &\leq \left(1-(M-1)\cdot Pr(X_p = x)\cdot \frac{(1-Pr(X_p=x))^{M-2}}{1-(1-Pr(X_p=x))^{M-1}}\right) \\
&= \left(1-(M-1)\cdot \frac{(1-Pr(X_p=x))^{M-2}}{\frac{1-(1-Pr(X_p=x))^{M-1}}{1-(1-Pr(X_p=x))}}\right) = \left(1-(M-1)\cdot \frac{(1-Pr(X_p=x))^{M-2}}{\sum_{i=0}^{M-2}(1-Pr(X_p=x))^i}\right) \\
&< \left(1-(M-1)\cdot \frac{(1-Pr(X_p=x))^{M-2}}{M-1}\right) < 1-(1-Pr(X_p=x))^{M-1} \\
&< \frac{1-(1-Pr(X_p=x))^{M-1}}{1-Pr(X_p=x)^{M-1}}
\end{aligned}$$

which means that δ_{SP} satisfies the condition (4.8). Hence, we proved the statement. □

Both conditions (4.8) and (4.10), as well as the expected payoff (4.7), depend on the number of tasks M and self-predictor δ_{SP}. The bounds on δ_{SP} in(4.8) and (4.10) are always greater than or equal to 0. This means that for categorical cases, where increase from prior to posterior only happens for the observed value, the conditions (4.8) and (4.10) are satisfied regardless of M and δ_{SP}. In the next subsection, we show how the number of tasks influences the amount of positive correlation allowed between different observation values for a more general case.

4.4.2 Limiting cases with the number of tasks $M = 2$ and $M \to \infty$

We first examine the case when the mechanism uses only 2 phenomena to reward an agent and one of them is observed by the agent (i.e., $M = 2$). The expected payoff of an agent a with observation x for reporting y is in that case equal to:

$$\left(\frac{Pr(Y_p = y|X_a = x)}{Pr(Y_p = y)} - 1\right) \cdot (1 - (1 - Pr(Y_p = y))) = Pr(Y_p = y|X_a = x) - Pr(Y_p = y)$$

when $Pr(Y_p = y) > 0$, while it is 0 otherwise. This effectively means that for $M = 2$, the RPTS score is in expectation equivalent to a score that rewards $Y_a = x$ with:

$$\tau_M(x, \mathbf{Y}_p, \mathbf{Y}_r) = \mathbb{1}_{x=x_p} - \mathbf{x}'_a(x_w) \tag{4.14}$$

where $\mathbf{x}'_a(x) = 2 \cdot \left(\mathbf{x}_a(x) - \frac{\mathbb{1}_{x=x_p}}{2}\right) = \mathbb{1}_{x=x_q}$, i.e., $\mathbf{x}'_a(x)$ is constructed by sampling one report from the phenomenon not observed by agent a. x_p and $\mathbf{x}_a$ are as defined in the RPTS mechanism.

The requirement for strict properness of this score is that each belief system B_a satisfies:

$$Pr(X_p = y|X_a = x) - Pr(X_p = y) < Pr(X_p = x|X_a = x) - Pr(X_p = x), \forall y \neq x$$

That is, an agent's belief change from prior to posterior should be the largest for the peer's observation equal to the agent's observation. However, the condition for strong truthfulness (4.10) imposes restriction that observation values are categorical, i.e., the increase from prior to posterior should only occur for the observed value:

$$Pr(X_p = y|X_a = x) - Pr(X_p = y) < 0, \forall y \neq x$$

Although condition (4.10) is only a sufficient condition of Theorem 12, it is actually tight for $M = 2$. Namely, if the condition did not hold, agents with observations x and y, and beliefs $Pr(X_p = y|X_a = x) - Pr(X_p = y) > 0$ and $Pr(X_p = x|X_a = y) - Pr(X_p = x) > 0$, would be better off reporting the same value (e.g., all of them report x or y) than reporting honestly. This comes from the fact that their expected payoff with such a collusive behavior would be $Pr(X_p = x|X_a = x) + Pr(X_p = y|X_a = x) - Pr(X_p = x) - Pr(X_p = y) > Pr(X_p = x|X_a = x) - Pr(X_p = x)$ and $Pr(X_p = x|X_a = y) + Pr(X_p = y|X_a = y) - Pr(X_p = x) - Pr(X_p = y) > Pr(X_p = y|X_a = y) - Pr(X_p = y)$, respectively for agents with observations x and y.

Table 4.1 – Desirable lower bound on δ_{SP} w.r.t. M and $\min_{a,z} Pr_{a,z}(X_p = z)$

M	$\min_{a,z} Pr_{a,z}(X_p = z) = 0.05$	$\min_{a,z} Pr_{a,z}(X_p = z) = 0.1$	$\min_{a,z} Pr_{a,z}(X_p = z) = 0.2$
10	$\delta_{SP} \le 0.19$	$\delta_{SP} \le 0.36$	$\delta_{SP} \le 0.65$
30	$\delta_{SP} \le 0.55$	$\delta_{SP} \le 0.84$	$\delta_{SP} \le 0.98$
60	$\delta_{SP} \le 0.84$	$\delta_{SP} \le 0.98$	$\delta_{SP} \le 1$
100	$\delta_{SP} \le 0.96$	$\delta_{SP} \le 1$	$\delta_{SP} \le 1$

With a larger number of tasks, the RPTS mechanism is equivalent to a payment rule that rewards $Y_a = x$ with:

$$\tau_M(x, \mathbf{Y}_p, \mathbf{Y}_r) = \frac{\mathbb{1}_{x=x_p}}{\mathbf{x}'_a(x)} \tag{4.15}$$

for $\mathbf{x}'_a(x) > 0$, and with 0 otherwise. Since $\mathbf{x}'_a$ includes report x_p, the mechanism can be simply described by saying that agent a gets payment equal to $\frac{1}{\mathbf{x}'_a(x)}$ when her report matches the report of her peer, and 0 otherwise. For a large number M, the RPTS requirements for strict properness and strong truthfulness coincide and are equal to the self-predicting condition with an unconstrained self-predictor $\delta_{SP} \in [0,1]$. In other words, RPTS allows, to some extent, observation spaces that are not necessarily categorical.

We see that the center has to decide on an appropriate number of tasks to allow correlations between two different observation values. To do this, it does not need a knowledge about agents' belief systems, only an upper bound on the minimal value of priors $\min_{a,z} Pr_a(X_p = z)$ and a lower bound on the value of self-predictor δ_{SP}. That is, $\min_{a,z} Pr_a(X_p = z)$ should not be overestimated, while δ_{SP} should not be underestimated. For example, one could incrementally take reports of different phenomena into account - one by one - until the agents' responses clearly indicate a bound on $\min_{a,z} Pr_a(X_p = z)$, determined by the frequency of the least frequent report, and a bound on δ_{SP}, determined by the correlation among different reported values.

Conditions (4.8) and (4.10) specify the upper bound on correlations among different observation values, expressed by self-predictor δ_{SP}. In the table below we show how quickly the upper bound of (4.10) approaches 1 as the number of tasks grows. Since, by Lemma 8, condition (4.10) is stricter than condition (4.8), the upper bound applies for both conditions. Clearly, for a reasonable number of tasks n, the bound allows significant deviations of δ_{SP} from the categorical case ($\delta_{SP} = 0$), even for the prior with values as small as 0.05.

We have seen that RPTS reduces to a simple score when statistic $\mathbf{x}_a$ is calculated based on only one phenomenon in addition to the phenomenon being observed by an agent a. The form of the score (4.14) is similar to the Dasgupta&Ghosh mechanism introduced in [DG13]. In fact, they are equivalent (see Section A.3 in the Appendix), which means that the Dasgupta&Ghosh mechanism is a special case of RPTS obtained in the limit case when $\mathbf{x}_a$ is calculated from only two phenomena. Moreover, the equivalence implies that the Dasgupta&Ghosh mechanism requires non-correlated (categorical) observation values for the honest reporting strategy profile to result in a maximum payoff.

4.5 Applications

The scenario depicted by Figure 4.1 captures many interesting crowdsourcing tasks. These include objective tasks which have correct answers and subjective tasks where workers (agents) are asked to provide their opinions. We present two examples of such crowdsourcing tasks, peer grading and community (participatory) sensing.

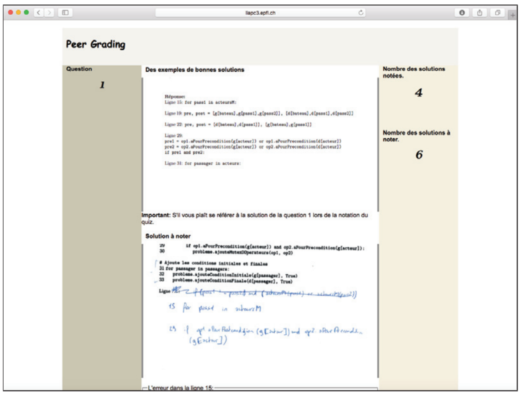
4.5.1 Peer grading

One of the main challenges in massive open online courses (MOOCs) represents evaluation of student assignments. This is especially true if assignments are essay questions that cannot be graded automatically. In such cases, peer grading techniques can be applied: a participant (student) grades assignments of their colleagues, and the grade of each student is obtained by aggregating the peer-grades.

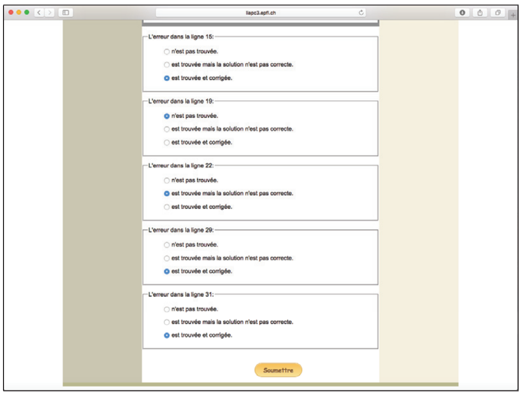
Peer grading in MOOCs represents an example of crowdsourcing where workers (agents) are students who are assigned to grade their own assignments. A proper monitoring of such a grading system is often infeasible due to a large number of participants, so the quality control has to be designed in the form of incentives. Moreover, the incentives have to take into account that participants have different grading abilities and are inclined to manipulate the reward system.

Often, the quality control in subjective crowdsourcing tasks is achieved by using the output agreement mechanism that rewards workers when their reports agree [vAD04, vAD08, HF13b]. This type of mechanism, however, does not take into account that workers may have a potential bias towards more likely evaluations. That is, workers who believe that their opinion is not the most common one, are incentivized to misreport. Moreover, colluding strategies where workers report the same value result in higher payoffs, and such behaviour is likely to occur [GMCA14].

We propose the RPTS mechanism for incentivizing peer-graders to invest their effort in grading students. We consider a simplified version of the mechanism, similar to the one described by expression (4.15), except that $\mathbf{x}_a$ is calculated from all peer and reference reports. This does not significantly effect the incentive properties, as the number of peer graders per grading task is relatively low compared to the total number of peer graders, so we can expect that $\mathbf{x}_a$ converges towards prior $Pr(X_p)$. On the other hand, the simplified version of RPTS is much easier to explain than the one which samples reference reports to construct $\mathbf{x}_a$.



(a) The correct solution to a quiz question and a student's solution



(b) Input form for corrections of a student's solution

Figure 4.3 – Peer grading task

Experimental setup

In order to test the impact that RPTS has on the quality of grades, we designed a peer grading experiments within "Artificial Intelligence" course at EPFL. In particular, as a part of the evaluation process, the course contained three quizzes, each consisting of two parts: in one part, students were asked to add a missing code; in the other, they were asked to find mistakes in a given code. The three quizzes took place at different time periods during the semester, assessing the knowledge about different topics of the course. Each problem in the quizzes had a correct solution and these solutions were used to assign points to the students, which were a part of the final grade. The official corrections of the quizzes were done by the teaching assistants of the course. Before the official points were announced, the students were asked to correct the solutions of their colleagues based on the correct solutions.

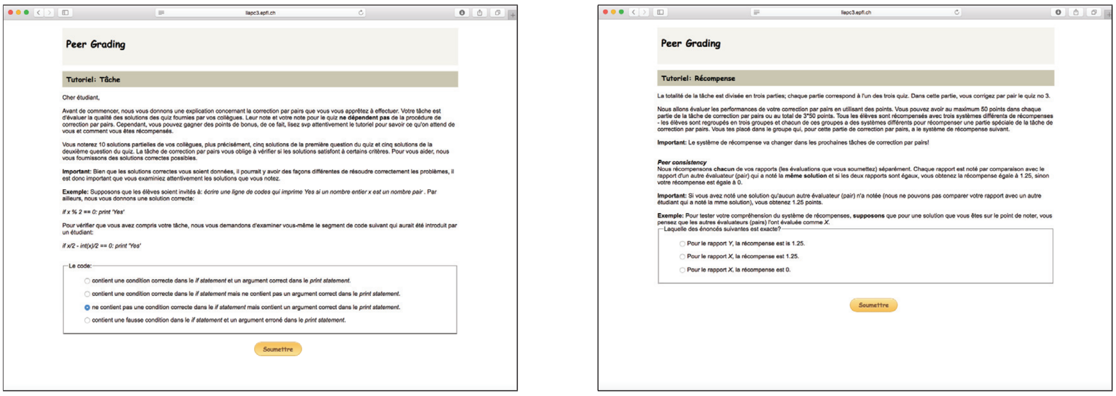
A criterion to determine the quality of a solution for a part of a quiz in which students were supposed to add a missing code was described by three to four different cases that defined potential mistakes or shortcomings of a student's solution. These cases were designed so that each of them covered combination of possibilities that could occur in the students' solutions, keeping in mind that the combinations are mutually exclusive between the cases. Naturally, a peer grader was selecting only one of these cases, and reporting only one value in total for the whole part. For the other part of the quiz, where students were supposed to find mistakes in a given code and correct them, a grading criterion was much easier to define. For each mistake in a given code, a student could either: not find the mistake; find a mistake, but not correct it; find a mistake and correct it. Therefore, a peer grader was presented with these three possibilities. Notice, however, that a peer grader made such reports for all mistakes that were in a given code (four to five), effectively reporting several values. Each reported value was treated separately in a peer rewarding mechanism. Figure 4.3a depicts the web interface of the peer grading task.

To incentivize participation we rewarded the peer graders with bonus points (additional points that could improve their grades), that were obtained using one of the three different reward schemes: the *constant* reward, the *output agreement* and RPTS. For the constant reward regime, a peer grader who participated in the peer grading obtained the maximum number of bonus points $MaxTotalReward$. For the output agreement, reward for reporting an answer was equal to $\frac{MaxTotalReward}{NumTasks}$ if a chosen peer reports the same answer, and is 0 otherwise. $NumTasks$ denotes the number of sub-parts to grade, which was equal to the number of reports that a peer grader made. The RPTS mechanism was also applied for each report separately. Furthermore, RPTS was scaled with the scaling parameters equal to $\alpha = \beta = \frac{1}{2} \cdot \frac{MaxTotalReward}{NumTasks}$ — this ensured that the bonus points remained positive (see Chapter 2). If a total number of the RPTS points exceeded $MaxTotalReward$, it was set to $MaxTotalReward$. Finally, statistic $\mathbf{x}_a(x)$ in RPTS was designed for each sub-part of a quiz separately, and it was defined as an empirical frequency of grades equal to x among all reports that were rewarded with RPTS for that sub-part of the quiz.

To test the quality of the reward schemes, we split the students into three groups of approximately the same number of students. Since participation in the peer grading experiment was not obligatory, the sizes of these groups varied. Each group was rewarded using all three reward schemes, but different mechanisms were applied for different quizzes in a round robin fashion. That is: if RPTS was used to assign rewards to a group for peer grading the first quiz, the same group was rewarded with the constant reward for peer grading the second quiz; if the output agreement was used to assign rewards to a group for peer grading the first quiz, the same group was rewarded by RPTS for peer grading the second quiz, etc.

In order to do a peer grading for a quiz, students needed to go through a tutorial that explained the peer grading task and a reward scheme that was used to assign bonus points - these two were separately explained in two different sections. The tutorial also contained two examples, one for the task explanation and one for the mechanism explanation. Each example contained a simple test questions for improving students' understanding. Different schemes had a different example question, showing the most basic features of the mechanisms. For the constant reward, students were asked to answer how many points they would obtain upon fulfilling the peer grading task, with three possible answers: *$MaxTotalReward$ per task*, *$MaxTotalReward$ in total*, or *it depends on how other raters grade*. For the output agreement, the question asked to pick the correct claim, provided that the peer reported *correct*. The claims were: *for reporting* correct *the reward is* 0, *for reporting* incorrect *the reward is* $\frac{MaxTotalReward}{NumTasks}$, or *for reporting* correct *the reward is* $\frac{MaxTotalReward}{NumTasks}$. Finally, for RPTS, the question asked what the reward was for reporting *correct* provided that everybody else reported *correct*, and the options were: $5 \cdot \frac{MaxTotalReward}{2 \cdot NumTasks}$, $3 \cdot \frac{MaxTotalReward}{2 \cdot NumTasks}$, or $1 \cdot \frac{MaxTotalReward}{2 \cdot NumTasks}$.[7] The options for each question were presented in a different order for different groups. Figure 4.4 depicts the tutorial web interface shown to users. A reward mechanism in subfigure is the output agreement; similar interfaces were designed for the constant reward mechanism and RPTS, and can be found in Section A.5 of the appendix.

[7] We used numerical values in all of the three test questions.



(a) Explanation of the peer grading task

(b) Explanation of a reward mechanism (output agreement)

Figure 4.4 – Tutorial about the peer grading task

Experimental results

We measured the quality of raw data (non-aggregated responses from students) with respect to the corrections made by the teaching assistants. For each student, we calculated the number of correct reports, and then, for each mechanism, we determined the average error rate, i.e. the percentage of incorrect grades. To measure the statistical significance, we performed two tailed student's t-test, with the significance level of 0.05. The null-hypothesis was that the students' error rates for two groups rewarded by different mechanisms follow the same distribution.

For the first two quizzes, each peer grader graded 4 partial solutions of her colleagues; more precisely, 2 solutions to the first part of the quiz, and 2 solutions to the second part of the quiz. Since our analysis did not reveal any statistical significance of the accuracy of the raw data across different schemes, we increased the number of solutions to grade for the third peer grading task. That is, for the third quiz, each peer grader graded 10 partial solutions of her colleagues; more precisely, 5 solutions to the first part of the quiz, and 5 solutions to the second part of the quiz.

Table 4.2 – Average error rate for different mechanisms

Mechanism	Num. students	Error rate (%)
RPTS	16	6.88
output agreement	16	10.48
constant	14	11.98

The results of the third quiz are shown in Table 4.2; for each group, they contain the number of students and the average error rate. As we can see, RPTS outperforms the baseline algorithms by 3-5%. Furthermore, t-tests (in Table 4.3) show that there is a statistically significant difference between the error rates for the RPTS mechanism and the error rates for the constant

Table 4.3 – T-tests: p-values for different mechanisms

Mechanism	RPTS	output agreement	constant
RPTS	-	0.0255	0.0497
output agreement	0.0255	-	0.5566
constant	0.0497	0.5566	-

reward or the output agreement, with p-values equal to 0.0497 and 0.0255, respectively.

4.5.2 Community sensing

In contrast to the previous chapter, we investigate now a community sensing scenario where the network of community sensors is evenly distributed across an urban area and each sensor reports frequently its measurements. This type of setting has been analyzed in [FLJ14], however, the proposed peer consistency method, called the peer truth serum, does not fully utilize the properties of the setting. In particular, its major drawback is that uninformed reporting strategies (strategies where sensors do not make measurements) can result in significantly higher expected payoffs than honest reporting, which we verify in this section.

We formalize the considered community sensing scenario using the multi-task peer consistency model, describe the application of the RPTS mechanism in the considered setting, and experimentally confirm that it effectively discourages a wide range of collusive strategies — those which are not based on sensors' locations. Furthermore, we also compare its experimental performance to that of the log-PTS mechanism, which also provides strong incentive properties, but requires a denser sensor network.

Air pollution monitoring

As in the previous chapter, we consider an air quality monitoring over an urban area, where each sensor is assigned a task of measuring air pollution at its location and reporting the observed value to the center.

Air pollution is a localized phenomenon, meaning that its value significantly varies with distance. Therefore, we approximate the correlation between two distant measurements by assuming that they are conditionally independent given a global state Γ, which is modeled as a random variable that takes values in a finite discrete set $\{\gamma_1, \gamma_2, ...\}$.

In particular, we model sensors' belief systems as in the formal setting of this chapter, but assuming that they depend on a specific value of Γ. In other words, the parameters of an agent a's belief system, such as probability distribution functions $Pr(\mathbf{X}_p, \mathbf{X}_r | X_a)$ and $Pr(\mathbf{X}_p, \mathbf{X}_r)$, depend on Γ, but satisfy the conditions imposed by the formal model and the set of constraints $\mathscr{C}$ for each value of Γ. For example, X_p and X_r are independent (given Γ), and if the self-predicting condition C_{MTSP} is imposed, the belief system satisfies:

$$\frac{Pr(X_p = y|X_a = x, \Gamma = \gamma)}{Pr(X_p = y|\Gamma = \gamma)} < \frac{Pr(X_p = x|X_a = x, \Gamma = \gamma)}{Pr(X_p = x|\Gamma = \gamma)}, y \neq x$$

for all $\gamma \in \{\gamma_1, \gamma_2, ...\}$. The set of peers of sensor a is now defined as the sensors located in the vicinity of sensor a, while the reference sensors are those located relatively far away from sensor a. With this structure of sensors' beliefs, the expected payoffs of sensors in the mechanisms discussed in this chapter preserve the properties of strict properness and strong truthfulness.[8]

Simulation setup

We examine the characteristics of incentives using realistic data of Nitrogen Dioxide (NO_2) concentrations over the city of Strasbourg. The data consists of both real measurements collected by ASPA [ASP13] and estimations of pollution from the physical model ADMS Urban V2.3 [CWC+02]. In total, the data set contains concentrations of NO_2 for each hour, expressed in parts per million (ppm), at 116 different locations over a period of four weeks. Each of the 116 locations represents a sensor that reports measurements on hourly basis and gets rewarded for each report separately. Figure 4.5 shows the sensor locations of the Strasbourg dataset.

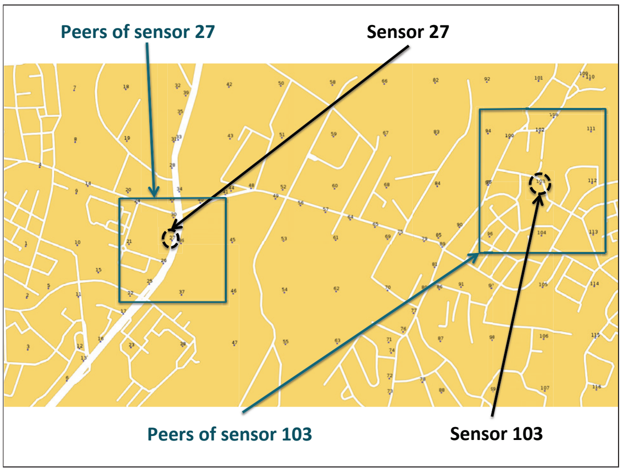


Figure 4.5 – Sensor placement in Strasbourg urban area

Although the initial measurements take values in $\mathbb{R}$, we discretize it using four levels of pollution defined as:

- *low*: concentrations 0 – 20 ppb;
- *medium*: concentrations 20 – 40 ppb;

[8] In log-PTS and RPTS, one can choose reference sensors that are not each other's peers when calculating $\mathbf{x}_\Phi$ and $\mathbf{x}_a$, respectively.

- *high*: concentrations 40 – 60 ppb;
- *extra-high*: concentrations 60 – ∞ ppb.

Each hour, sensors report the measured level of pollution to the center and are rewarded for providing their measurements. As a criterion for peer selection, we consider distance and define peers of a certain sensor as k closest sensors. For example, Figure 4.5 shows where peers of 2 sensors, sensor 27 and sensor 103, are located on the map. The peers of a sensor are in this case defined as the ten sensors closest to it.

In the RPTS mechanism, we select one of $k = 10$ peers to score a sensor using the simplified version of the RPTS mechanism, in which statistic $\mathbf{x}_a$ is calculated from all of the reports except the report of a sensor being scored. In the peer selection process of RPTS mechanism, we effectively simulate the prior knowledge of the center by identifying for each location a neighboring location at which the true measurements are the most correlated to the true measurements at the considered location.[9] The sensor located in this neighboring location is considered to be a peer.[10]

In the peer truth serum (e.g., [JF11, FPTJ14]), we use the same peer selection process as in the RPTS mechanisms. The PTS mechanism requires a knowledge about sensors' prior belief, so we construct the prior by calculating the frequency of different pollution levels in the whole dataset. Notice that the frequencies are calculated from the true data, not sensors' reports that are not necessarily truthful. The obtained prior $\mathbf{x}$ is equal to: $\mathbf{x}(low) = 0.402$, $\mathbf{x}(medium) = 0.384$, $\mathbf{x}(high) = 0.16$, $\mathbf{x}(extra-high) = 0.054$. The PTS mechanisms uses the same scaling parameters as RPTS. Thus, it rewards sensor a with $1/\mathbf{x}(Y_a) - 1$ if its report matches the report of her peer, and otherwise the payoff of the sensor is equal to -1.

In the log-PTS mechanism, we use all $k = 15$ peers to calculate the frequency of peer reports $\mathbf{x}_{\Phi_i}$, while we simplify the calculation of the frequency of reference reports $\mathbf{x}_\Phi$ by including in it the reports of all sensors, except for the report of the sensor that is being scored. The latter resembles the simplification that we adopted for the RPTS mechanism. Furthermore, $\mathbf{x}_{\Phi_i}$ and $\mathbf{x}_\Phi$ are smoothed using the Laplace (additive) smoothing operator with parameters $\alpha_{\Phi_i} = 10^{-4}$ and $\alpha_{\Phi_i} = 10^{-3}$ (parameters reflect that $\mathbf{x}_\Phi$ is calculated based on approximately 8 times more reports than $\mathbf{x}_\Phi$). No specific scaling was used in the log-PTS mechanism.

To demonstrate the correctness of our results, we examine six different reporting strategies and evaluate their performance by analyzing the average scores of sensors. The six strategies are defined as follows:

- *honest*: All sensors are honest.

[9] Notice that we examine the correlations using the true data, not sensors' reports, which are not necessarily truthful.

[10] On average, the best response to truthfulness is to report honestly, indicating that, in the considered data set, the self-predicting condition holds in an average case.

- *collude*: Sensors collude so that those who observe low or $medium$ report low, while those who observe $high$ or $extra-high$ report $high$.
- *colludeLow*: All sensors collude and report low.
- *colludeExtraHigh*: All sensors collude and report $extra-high$.
- *random*: A sensor whose score is being calculated reports uniformly at random, while others sensors are honest.
- *randomAll*: All sensors report uniformly at random.

For each sensor, we run a separate process in which the sensors report according to one of these strategy profiles and we calculate the average payoff of the considered sensor.

Simulation results

The statistics of the average RPTS payoffs are shown in Table 4.4.[11] These payoffs can be further scaled in different ways, so that, for example, the incentives take positive values and cover the cost of sensing.

Table 4.4 – Average payoffs — RPTS

Strategy	mean	min	max	median	1st quartile	3rd quartile
honest	0.678	-0.003	5.997	0.366	0.27	0.781
collude	0.232	-0.0146	2.177	0.105	0.07	0.281
colludeLow	0	0	0	0	0	0
colludeExtraHigh	0	0	0	0	0	0
random	0.022	-0.1974	2.678	-0.108	-0.143	0.017
randomAll	0.007	-0.216	0.214	0.018	-0.044	0.061

As expected, random reporting strategies lead to scores that are concentrated around 0, which is clearly seen from the median of *random* and *randomAll* strategies. Colluding on a single value results in a payoff equal to 0, and this trivially follows from the structure of the score. Collusion strategy *collude* has lower mean of the average payoffs than honest reporting. Moreover, a careful inspection of medians and quartiles shows that the collusive strategies are worse than honest reporting for the majority of sensors: the median, the 1st quartile, the 3rd quartile and the maximum of average payoffs are greater for honest reporting than for the collusive strategies.

The described scenario involves stationary sensors, which means that the sensors are solving approximately the same task over a longer period of time. This means that some sensors might

[11] T-tests show that the RPTS payments for truthful reporting are statistically different than the RPTS payments for the other strategy profiles, with p-values smaller than 0.01. The same holds for the log-PTS (Table 4.6). For the standard PTS (Table 4.5), the highest paying strategy profile is *colludeExtraHigh*, with payments that are statistically different than the PTS payments for the other strategy profiles (p-values are smaller than 0.01).

Table 4.5 – Average payoffs — PTS

Strategy	mean	min	max	median	1st quartile	3rd quartile
honest	1.992	1.449	3.094	1.908	1.72	2.183
collude	1.982	1.724	2.394	1.966	1.882	2.054
colludeLow	1.485	1.485	1.485	1.485	1.485	1.485
colludeExtraHigh	17.618	17.618	17.618	17.618	17.618	17.618
random	-0.01	-0.333	0.387	-0.022	-0.116	0.083
randomAll	0.87	0.478	1.282	0.861	0.761	0.941

be favored in terms of their average payoffs. For example, a sensor that reports randomly might obtain a relatively high average payoff over a longer sensing period when the histogram of its reports is more correlated to the reports of its peer than statistic $\mathbf{x}_a$ is. Although the sensor reports randomly, its reports carry some information about its peer w.r.t. $\mathbf{x}_a$, hence it is not surprising that such a sensor might obtain positive rewards. Notice, however, that honest reporting leads to significantly higher payoffs, as shown in Figure 4.6.

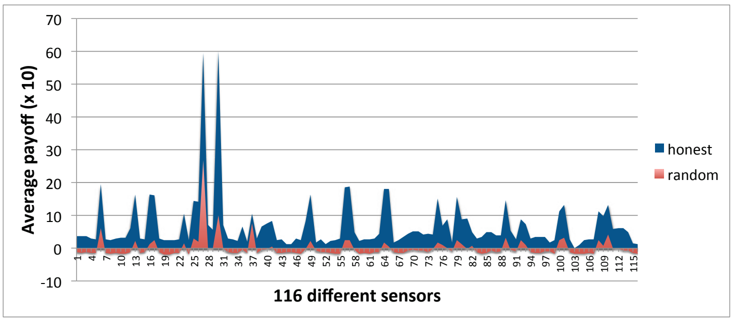


Figure 4.6 – Average payoffs (times 10) of *honest* and *random* strategies for each sensor, arranged in no particular order along the x-axis.

Unlike the RPTS mechanisms, the standard version of PTS is not resistant to collusive strategies. As shown in Table 4.5, payoffs of the PTS mechanisms are significantly higher for collusive strategies, in particular, when sensors report the least likely value ($extra-high$), which leads to the order of magnitude greater payoffs than truthful reporting. This shows us the importance of the robust design that can be achieved in the multi-task model.

Finally, we examine the payoffs of log-PTS and compare its qualitative performance to RPTS. The statistic of the average log-PTS payoffs is shown in Table 4.6. As for RPTS, these payoffs can be scaled so that they take positive values.

Qualitative performance of log-PTS is the same as for RTPS: honest reporting results in the highest payoff. However, to achieve these properties, log-PTS requires a relatively dense sensor network, with about 15 peers. We further investigate how robust log-PTS is when the density of the sensor network decreases. To do so, we randomly sample subsets of sensors of different sizes (100, 80, 60 and 40 sensors) on daily basis (i.e., each day a different subset is chosen), and we calculate the median of average payoffs. Namely, the median of average payoffs reflects

Table 4.6 – Average payoffs — log-PTS

Strategy	mean	min	max	median	1st quartile	3rd quartile
honest	0.037	-1.153	0.291	0.047	-0.017	0.102
collude	0.014	-0.27	0.106	0.019	-0.009	0.039
colludeLow	0	0	0	0	0	0
colludeExtraHigh	0	0	0	0	0	0
random	-0.876	-1.631	-0.36	-0.823	-1.075	-0.673
randomAll	-0.228	-0.362	-0.123	-0.228	-0.258	-0.19

how good a reporting strategy is for the majority of the sensors. In addition to reducing the number of sensors, we also reduce the number of peers for each sensor. For example, in a random subset of 80 sensors, the set of peers of a certain sensor contains 11 closest sensors. Since the average payoffs in Table 4.6 for honest reporting were significantly higher than for random reporting strategies, we only examine *honest*, *collude* and *colludeLow* strategy profiles (*colludeHigh* results in the same payoff as *colludeLow*). The detailed results can be found in Section A.4 of the appendix, and they include also other indicators, such as the mean of the average payoffs.

Figure 4.7 shows the median of the average payoffs for the three strategies. Scores are scaled so that the maximum score is equal to 50, while the score in *colludeLow* strategy profile is equal to 10. We can see that truthful reporting remains the optimal strategy until the number of sensors decreases to 40, which represents a critical value where the collusive strategies *colludeLow* and *collude* become more profitable than truthfulness. This can be explained by the low amount of information used in generating $\mathbf{x}_{\Phi_i}$ and $\mathbf{x}_{\Phi}$ ($\mathbf{x}_{\Phi_i}$ is constructed from only 7 reports). In *colludeLow* and *collude* strategies, sensors report one and two levels of pollution respectively, so these strategies are less susceptible to random variations in measurements than truthful reporting, where all four levels of pollution are reported.

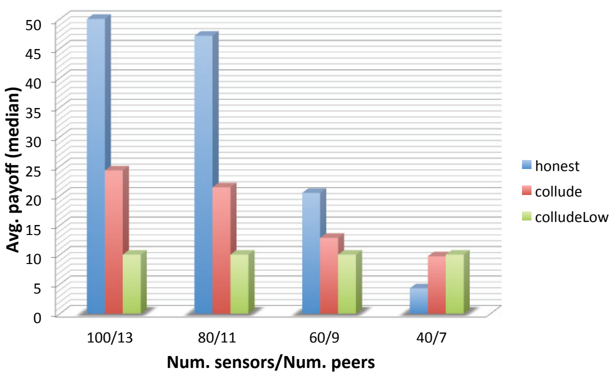


Figure 4.7 – Median of average payoffs for different number of sensors and peers (log-PTS).

Unlike log-PTS, RPTS produces relatively stable scores across different configurations in the number of sensors and peers. To avoid a potential bias in the peer selection process, we now use all peers in rewarding a sensor with the RPTS mechanism: the reward of a sensor is obtained by averaging RPTS rewards across all of the peers. As shown in Figure 4.8, RPTS scales down quite well, preserving the strong incentive properties even for a relatively small population of sensors. This implies a greater practicality of the RPTS mechanism when compared to the log-PTS mechanism.

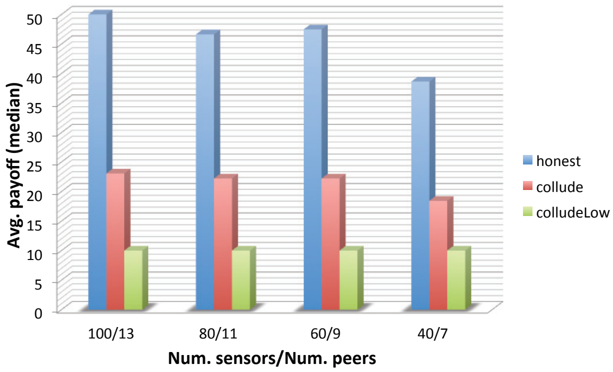


Figure 4.8 – Median of average payoffs for different number of sensors and peers (RPTS).

4.6 Conclusion

In this chapter, we investigated an elicitation setting where agents observe multiple phenomena, which models a typical multi-task crowdsourcing scenario. We showed that, when agents' characteristics are homogeneous, there exists a simple and intuitive mechanism for truthful elicitation of agents' private information. The mechanism implements the peer prediction with the quadratic scoring rule by appropriately sampling reports in the multi-task setting. Unlike the (original) peer prediction, it does not require the knowledge of agents' beliefs, that are also allowed to be different. When the population of agents is large, we showed how to adopt the principles of the (original) Bayesian truth serum in order to construct a mechanism that not only allows agents to have different private beliefs, but is also strongly truthful, meaning that truthful reporting results in the highest payoff among all strategy profiles.

On the other hand, for agents with heterogeneous characteristics, it is not possible to produce strict incentives for truthfulness in the general case. Therefore, we analyzed the case when agents have limited heterogenous characteristics, described by the self-predicting condition. We presented a robust version of the peer truth serum mechanism, and showed that it has strong incentive properties. Due to its simplicity and robustness, the mechanism is applicable to a wide variety of crowdsourcing settings, such as community sensing or peer grading, which we supported with experiments and simulations.

The mechanisms explained in this section rely on rewarding agents when their reports match. Therefore, an interesting direction would be to explore whether one can extend the principles of the continuous BTS to allow strong truthfulness in the multi-task setting when observations are real-valued. Such a result is likely to provide a great insight into how much heterogeneity one can allow in multi-tasks settings. We conjecture that, under the multi-task state model C_{MTSM}, there is no strictly proper mechanism that allows deviations from the homogenous population condition C_{HP} when observations take real values. We expect that this result could be proven in a similar fashion as Theorem 7 from the previous chapter. Notice that the claim in this case is much stronger than that of Proposition 6.

Since the strong truthfulness implies strict properness, the results related to the elicitation of real-valued observations could also explain why the existing strongly truthful mechanisms for the homogenous population condition C_{HP}, such as log-PTS or the mechanism of [KSM$^+$15], require a large number of phenomena. Namely, by examining the structure of the existing strictly proper mechanisms for the homogenous population condition C_{HP}, we expect that any strictly proper mechanism designed for the same belief constraint, but real-valued observations, might need to sample any number of phenomena with strictly positive probability.

Finally, notice that the incentive properties of this and the previous chapter, i.e., strong truthfulness and strict properness, rely on agents being rational and responding to incentives. In contrast, if there is a large enough coalition of malicious agents who deliberately misreport, these properties do not suffice to incentivize other agents to report honestly, nor do they prevent the center from learning a wrong aggregate. Therefore, in the next chapter we investigate a setting where a few agents are trusted to provide truthful observations, so that the center can use their reports to construct incentives for informed reporting and limit the negative impact of malicious participants.

5 Reputation-based incentives for online information aggregation

In the previous chapters, crowd participants are assumed to be rational agents who respond to incentives provided by the center. This approach, however, reaches its limit of effectiveness when a participant intends to be malicious and intentionally misreports values. Therefore, we investigate a more rigorous approach in order to identify faulty or malicious agents — one that is based on the reputation system framework. Notice that the agents are now assumed to interact with the center over a longer period of time. This fact enables us to track the quality of the information that agents provide via reputations, and hence discard information that comes from low quality agents.

5.1 Formal setting

Figure 5.1 depicts the particularities of our setting. The center plays a role of an aggregator that aims to estimate the state of a time evolving phenomenon based on the information provided by a group of agents. At the beginning, the center has only prior information about the observed phenomenon. After some time, the center receives a report and merges it with the current estimate of the phenomenon using a pre-specified aggregation procedure $\mathscr{M}$, thus producing a new estimate. This process repeats until a trusted agent reports her observation, after which the center can evaluate the reports of the crowd participants. We consider this to be one period of *sensing* and we denote it by t. The sensing process then continues in the same manner until the period $t = T$; we refer to T as *sensing time*.

Notice that we consider the case where the private information is noisy and trusted agents are a sparse resource. This means that the center cannot only use the reports of trusted agents to properly monitor the phenomenon, but rather it needs to support their observation with that of other agents. As an example scenario, one can consider sensing of an environmental phenomena, such as air pollution, where it is reasonable to assume that the center can place a few of its accurate sensors on, for example, public transportation, but to obtain a finer grained sensing resolution, it uses crowd participants.

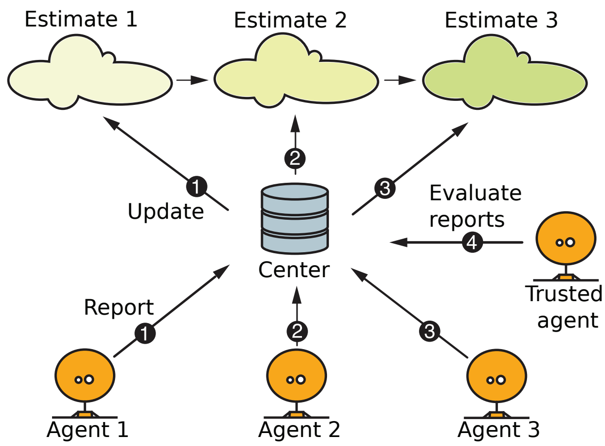


Figure 5.1 – Elicitation with online information aggregation

5.1.1 Aggregation model

In the considered setting, the center's goal is to construct and publish an estimate E about the current state of a time-evolving phenomenon, using the currently available set of observations. We are particularly interested in a real-time updating where estimate E is updated after receiving each observation using an aggregation model $\mathcal{M}$. We keep a general form of aggregation model $\mathcal{M}$, where the input is defined by a finite set of reported observations $\{Y_1, Y_2, ...\}$, while the output is an estimate E.[1] Estimate E can contain different information about the phenomenon. For example, in case of pollution sensing, which is a spatially distributed phenomenon, the estimate could contain the probability distribution functions over possible observations at different points of interest.

Since we want to keep a possibility of having a very general aggregation model $\mathcal{M}$, we consider it as a *black box*. This implies that after receiving a report from agent a, the center should decide whether to publish a new estimate E_a^{new} obtained by incorporating the report Y_a into the existing estimate E_a^{old} or to keep the existing estimate as its output. The rationale behind this is that the estimate updating should be computationally efficient. If, for example, a new output would be a linear combination of E_a^{new} and E_a^{old}, a proper updating procedure for obtaining P_a^{new} would have an exponential time complexity in the number of agents, as we argue in the following sections.

[1] Model $\mathcal{M}$ can also use other information in calculating estimates. For example, in community sensing, sensors can report both their locations and measurements, which is useful for aggregation models based on spatial correlations. Furthermore, model $\mathcal{M}$ can incorporate agent specific information in its aggregation procedure, which can be prior or elicited knowledge. For example, agents can report how confident they are in their observations, in which case, model $\mathcal{M}$ can weight reported observations according to agents' confidences.

5.1.2 Population of agents

In the considered setting, the center trusts a certain number of agents that we refer to as the *trusted* agents; these agents are assumed to report truthful observations. The center does not know the character of other agents, i.e., whether they are malicious or not. *Non-malicious* agents are considered to be strategic — rational agents that aim to maximize their payoffs — or honest, while *malicious* agents do not respond to incentives and their goal is to lower the quality of produced estimates. In the group of malicious agents, we can also put *faulty* agents that are not intentionally malicious, but do provide inaccurate data. Furthermore, notice that malicious agents might report accurately in some sensing periods in order to deceive the center. This means that the decision on how to use the agents' reports in the information fusion process should be done by monitoring the behaviour of the agents over the whole sensing time T.

As for the previous chapters, we denote agent a's observation by X_a, and we consider it to be a random variable that takes values in $\mathscr{X}$. The agents are assumed to provide only their information report Y_a, i.e., $R_a \in \mathscr{R} = \mathscr{X}$.[2] We focus on the payment mechanisms that reflect agent a's contributions to the quality of provided estimates, measured by scoring them against a trusted report. Therefore, a payment function τ_G, which we define in the next subsection, depends on agent a's report and the report of a trusted agent. Moreover, it also depends on the estimate published prior to the agent's observation and the center's aggregation model $\mathscr{M}$. To simplify the description of our algorithm, we impose three conditions for agents: an agent reports one measurement per time period, observations between two time periods t_1 and t_2 are statistically independent, and reports from different agents arrive stochastically one at a time (i.e., without a specific order).

Strategy space. We make restrictions to the strategic space of malicious agents by assuming that their reports do not have a significant impact on the quality of the information provided by non-malicious agents. As noted by [RS07], the restriction to *myopic strategies* is not a trivial assumption, but still allows a large scope of possible misreporting strategies, including strategies where malicious agents change their reporting behaviour over time.[3] Furthermore, it is likely that non-myopic strategies require complex implementation. For example, an effective malicious strategy that is based on the report sequence would require information about the start and end time of the sensing periods. Since each sensing period ends when a trusted report is submitted, the center can easily obscure the starting point of a sensing period by, for example, not immediately notifying agents of their reputation change. This also

[2] While in our formal model we assume that agents only report their observations, the main results of this chapter are not dependent on the structure of the report. As already noted, more complex report structure can be useful in making aggregation more accurate.

[3] Notice that such a strategy space is not trivially describable with *heterogeneous reporting types* (where each type defines how to transform an observation to a report), because the reporting type of an agent might be dependent on both time and the reporting types of other agents. For example, [RYZ+10] measure the performance of agents (annotators) in binary classification in terms of the *sensitivity* and *specificity* with respect to the unknown gold standard. In our setting, a malicious agent can change her *sensitivity* and *specificity* depending on, for example, the current estimate E or the *sensitivity* and *specificity* of other (malicious) agents.

provides a justification for the assumption of stochastic arrival of reports.

5.1.3 Quality score

We evaluate agents by their marginal contributions to the quality of produced estimates. More precisely, consider an estimate E_a^{new} obtained by fusing agent a's report with an estimate E_a^{old} that preceded the report of agent a. Furthermore, let $S_E(E, X_{trust})$ be a general scoring function that evaluates the quality of an estimate with respect to the report $Y_{trust} = X_{trust}$ of a trusted agent, and let it be scaled so that it takes values in interval $[-1/2, 1/2]$ (see Chapter 2). The score of agent a is then defined by the gain G_a of the center when it fully incorporates the agents's report into the existing estimate E_a^{old}:

$$score_a = G_a = S_E(E_a^{new}, X_{trust}) - S_E(E_a^{old}, X_{trust})$$

It is easy to see that the score takes values in $score_a \in [-1, 1]$. The score can further be used to calculate (monetary) incentives given to the agent. In particular, we define payment mechanism $\tau_G : \mathscr{X} \times \mathscr{X} \to \mathbb{R}$ as $\tau_G(Y_a, X_{trust}) \stackrel{\text{def}}{=} score_a$. Notice that we deliberately abuse our notation by having three equivalent quantities, $score_a$, G_a and τ_G, in order to clearly specify the meaning of the properties we investigate in this chapter.

5.1.4 Myopic impact

Following the approach from [RS07], we use the notion of agent a's *myopic impact*. Since our main method probabilistically decides whether to accept or discard agent a's report, we adopt a notion of *expected myopic impact*.

Definition 19. *The* expected myopic impact *of agent a at time period t is defined as:*

$$\bar{\Delta}_{a,t} = \pi_{update} \cdot G_{a,t} + (1 - \pi_{update}) \cdot 0 = \pi_{update} \cdot G_{a,t}$$

where π_{update} is the probability of incorporating agent a's report into the existing output. Furthermore, we define the total myopic impact as $\bar{\Delta}_a = \sum_{t=1}^{T} \bar{\Delta}_{a,t}$.

The intuition behind the definition is straightforward. Whenever the center *accepts* to fuse agent a's report into the existing estimate, the agent's impact is equal to the center's information gain: $G_{a,t} = S_E(E_a^{new}, X_{trust}) - S_E(E_a^{old}, X_{trust})$. Otherwise, when the center decides to *discard* agent a's report, the agent's impact is 0 because it does not change the center's output E_a^{old}. Notice that the myopic impacts, $\bar{\Delta}_{a,t}$ and $\bar{\Delta}_a$, are functions of $G_{a,t}$. Since $G_{a,t}$ is a random variable, we can associate expected values over $G_{a,t}$ for both $\bar{\Delta}_{a,t}$ and $\bar{\Delta}_a$, which we denote by $\mathbb{E}(\bar{\Delta}_{a,t})$ and $\mathbb{E}(\bar{\Delta}_a)$, respectively.[4]

The property we want to achieve with a reputation framework is a bounded negative impact

[4] $\mathbb{E}(\bar{\Delta}_a)$ is the expectation over gains from all time periods.

of any agent. That is, the total myopic impact of agent a should be bounded from below by a value independent of sensing time T, i.e. $\bar{\Delta}_a > -O(1)$.

5.1.5 Information loss

Bounding the negative value of a myopic impact does not entirely guarantee that a reputation system has a good performance. For example, a simple reputation system that discards all the reports completely limits the negative influence of malicious agents, but in doing so, it discards all the valuable information coming from non-malicious agents as well. Therefore, it is also necessary to measure an *information loss* for potentially discarding reports of an informed agent a.

Definition 20. *Consider an agent a whose expected scores are strictly greater than a predefined parameter* $score_{min} > 0$*, i.e.,* $\mathbb{E}(score_{a,t}) > score_{min} > 0$*. The expected* information loss IL_a *for an agent a is defined as:*

$$IL_a = \sum_{t=1}^{T} [\mathbb{E}(score_{a,t}) - \mathbb{E}(\bar{\Delta}_{a,t})]$$

The rationale behind this definition is that an agent's scores reflect her contributions — information gains — that the agent would have made had it not been limited, while her myopic impact reflects the agent's real contribution. We define information loss only for agents that in expectation provide positive contributions better than some predefined threshold. The information coming from other agents is not considered to be reliable, so we want to discard it in the first place.

5.2 Related work

A common approach to filter low quality information in crowd work is to batch process the elicited data and apply machine learning or statistical methods to infer the true labels (e.g.,[RYZ+10, KOS11, LPI12, KOS13, JSV14]). This approach, however, implies that the underlying phenomenon, whose state is being estimated, is not monitored in real-time but rather periodically.

In contrast, we investigate an online information aggregation setting, where estimates should be updated after receiving each input from the crowd participants. The setting relates to the vast literature on regret-minimization algorithms (e.g., [LR85, ACBF02, CBMS07, LW94, NRTV07, CBL06]), and is similar to the model of prediction with expert advice [CBL06], in particular, to the one where expert advice is sequential [KS10, KS11]. More precisely, in our setting, agents, who arrive sequentially, are experts that provide either adversarial (malicious) advice or advice that can be considered to be stochastic. Notice that this differs from the traditional expert algorithms (e.g., [LW94]) where experts arrive simultaneously. Furthermore, we evaluate agents by the influence they have on the aggregate, which contrasts a traditional way of evaluating each expert independently and is closely related to prediction markets [Han03, CP07, CV10]. Unlike prediction markets, we do not ask an agent for prediction but rather her private information, which is then aggregated explicitly with the existing reports. The closest learning algorithm to our work is the influence limiter [RS07]. The algorithm uses a reputation based framework and is provably resistant to misreporting. It was primarily designed for recommender systems, and, as we show later in this chapter, its direct transformation to our setting has several drawbacks.

Apart from the online learning algorithms, our work relates to the extensive literature on trust and reputation systems (e.g., [JI02, Del05, ZVdS12, XS15]), out of which we emphasize those for sensing, since the main application of our technique is community sensing.

The standard approach of dealing with untrustworthy information in sensing is by using reputation systems [MM02, BB03, GS04, BLB02, YS10, Che09], with the Beta reputation system [JI02] being the most common way of assigning reputation scores. While in the literature one can find other ways of assigning reputation scores, such as using the Gompertz function [HKH14], the classification of whether a sensor misbehaves is typically based on a simple thresholding principle: if the reputation of a sensor is lower than a certain threshold, the sensor is denoted as misbehaving, otherwise, it is considered to be trustworthy. A thresholding approach is common even among the techniques that do not necessarily use reputation systems (e.g., [WLSH09]). While such a thresholding principle can cope with simple attacks where malicious sensors report consistently wrong values, it fails to protect the center against deceiving attacks, as we describe it later in the paper.

[VRJ13] and [RRCN09] take a different approach to fuse information from multiple sensors that are not a priori assumed to be trustworthy. [VRJ13] tries to learn the parameters related to the trustworthiness using a maximum likelihood method over the assumed (Gaussian) model with unknown parameters. [RRCN09] proposes a two stage Bayesian multi-sensor fusion algorithm that incorporates model of sensors' trustworthiness. Neither of the two multi-sensor fusion methods have provable guarantees on the loss of the system experienced when the majority of sensors is untrustworthy and potentially malicious. As alternatives to reputation systems, we also mention hardware solutions, such as trusted platform modules (e.g., [SW10, GJL$^+$11]). These approaches, however, require additional hardware on each sensing module, which limits their applicability.

5.3 Traditional approach

Let us now describe the thresholding approach traditionally used in reputation systems. When the center receives a report $Y_{a,t}$ of agent a, it fuses the report with the existing information if agent a's reputation is greater than a certain classification threshold Θ, and otherwise discards it. The approach is depicted by Algorithm 1. Function $Update(E, Y_{a,t})$ uses the existing set of *included* reports (the set of reports that produced estimate E), adds to it report $Y_{a,t}$, and

Data: Initial reputation ρ_0, threshold Θ
begin
  for *Agent a* **do**
    $\rho_{a,1} \longleftarrow \rho_0$;
  end
  for $t = 1$ *to* $t = T$ **do**
    Compute prior estimate E;
    Publish E;
    for *Agent a* **do**
      Receive a's report $Y_{a,t}$;
      $E_a^{old} \longleftarrow E$;
      $E_a^{new} \longleftarrow Update(E, Y_{a,t})$;
      if $\rho_{a,t} \geq \Theta$ **then**
        $E \longleftarrow E_{new}^a$;
        Publish E;
      end
    end
    Receive report $Y_{trust,t} = X_{trust,t}$;
    for *Agent a* **do**
      $score_{a,t} \longleftarrow S_E(E_a^{new}, X_{trust,t}) - S_E(E_a^{old}, X_{trust,t})$;
      $\rho_{a,t+1} \longleftarrow RepUpdate(\rho_{a,t}, score_{a,t})$;
    end
  end
end

Algorithm 1: Thresholding

applies model $\mathscr{M}$ to obtain a new estimate. $RepUpdate$ updates the reputation of agent a using $score_{a,t}$, and has two conditions:

- if $score_{a,t}$ has a strictly positive constant value, the reputation converges over time towards its maximum value;
- if $score_{a,t}$ has a strictly negative constant value, the reputation converges over time towards its minimum value.

This simple reputation system can be considered to be a part of a large family of reputation systems that use fix thresholds to classify whether a certain agent misbehaves or not. These reputation systems can cope with simple attacks where malicious agents report consistently wrong values. For example, in case of pollution monitoring, they can limit the effectiveness of the malicious strategy that consists of reporting low pollution values. However, they fail to protect the system against deceiving attacks.

One particular deceiving strategy of a malicious agent could be to report informative values when her reputation is below threshold Θ, while report low quality information when her reputation is above the threshold. The intuition behind this attack is that an agent reports useful information only when the center does not use it, and when the center uses her information, it deliberately misreports.

Proposition 7. *Consider an aggregation model $\mathscr{M}$ that allows arbitrary generation of gains $G_{a,t}$ related to agent a. Then there exists a sequence of gains such that the total myopic impact $\bar{\Delta}_a$ of agent a in Algorithm 1 is negative and monotonically decreases with T, i.e.,* $\lim_{T\to\infty} \bar{\Delta}_a = -\infty$.

Proof. Consider a sequence of gains such that whenever $\rho_{a,t} < \Theta$, gain $G_{a,t}$ is equal to $G_{a,t} = g > 0$, while $\rho_{a,t} \geq \Theta$ implies negative gain $G_{a,t} = -g < 0$. In other words, $\pi_{update} = 1$ for $G_{a,t} < 0$ and $\pi_{update} = 0$ for $G_{a,t} \geq 0$. Since reputations converge to the maximum possible reputation if $score_{a,t}$ (i.e., $G_{a,t}$) is fixed to $g > 0$, we know that $\rho_{a,t}$ will infinitely often be greater than Θ for $T \to \infty$. Therefore, $\bar{\Delta}_a$ is negative (because $\pi_{update} = 0$ for $G_{a,t} \geq 0$) and $\lim_{T\to\infty} \bar{\Delta}_a = -\infty$ (because $\rho_{a,t} \geq \Theta$ infinitely often). □

5.4 Influence limiter

The influence limiter, when transformed to our setting, has the same skeleton structure as the thresholding algorithm with the main differences in three components, which we point out in this section. These components enable it to be provably resistant to any myopic-based manipulation strategy (Theorem 4 and Theorem 7 in [RS07]). We show, however, that all of the three components should be modified in order to obtain a practical algorithm. The structure of the influence limiter is depicted in Algorithm 2.

Data: Initial reputation ρ_0
begin
  for *Agent a* **do**
    $\rho_{a,1} \longleftarrow \rho_0$;
  end
  for $t = 1$ *to* $t = T$ **do**
    Compute prior estimate E;
    Publish E;
    for *Agent a* **do**
      Receive a's report $Y_{a,t}$;
      $E_a^{old} \longleftarrow E$;
      $E_a^{new} \longleftarrow Update(E, Y_{a,t})$;
      $w_{a,t} = \min(\rho_{a,t}, 1)$;
      $E \longleftarrow (1 - w_{a,t}) \cdot E_a^{old} + w_{a,t} \cdot E_a^{new}$;
    end
    Receive report $Y_{trust,t} = X_{trust,t}$;
    for *Agent a* **do**
      $score_{a,t} \longleftarrow MC_{QSR}(E_a^{new}, E_a^{old}, X_{trust,t})$;
      $\rho_{a,t+1} \longleftarrow \rho_{a,t} + w_{a,t} \cdot score_{a,t}$;
    end
  end
end

Algorithm 2: Influence limiter

Information aggregation. The standard version of the influence limiter has a deterministic information fusion component. In particular, the influence limiter incorporates all of the reports, but assigns different weights to different reports. In our scenario, this would mean that when a report from an agent a is received, the new estimate E_a^{new} is calculated and the published estimate E is updated to:

$$E \longleftarrow (1 - w_{a,t}) \cdot E_a^{old} + w_{a,t} \cdot E_a^{new} \tag{5.1}$$

Here, the weight is equal to $w_{a,t} = \min(\rho_{a,t}, 1)$. The crucial part of the algorithm is how E_a^{new} should be calculated, i.e., the structure of the $Update$ function.

In the influence limiter, a sensible updating function has to include the fact that all reports are fused, but with different weights. Since aggregation model $\mathscr{M}$ is assumed to be a *black box*, one has to additionally ensure that the reports are properly weighted (limited) when updating estimate E. For example, consider two reports Y_{a_1} and Y_{a_2} that arrive sequentially. Initially, E should be set to $\mathscr{M}(\emptyset)$. Once Y_{a_1} is reported, the update of E, denoted by E_1, is easy to calculate: we simply make a linear combination of E and $\mathscr{M}(\{Y_{a_1}\})$, with weights $1 - w_1$ and w_1 (see (5.1)).

The problem, however, arises when we update the current estimate E_1 for report Y_{a_2}. Namely, the new update should be a linear combination of the current estimate E_1 and the estimate E_2^{new} that does not limit Y_{a_2}, but does appropriately limit the reports that had arrived before Y_{a_2}. In our case, the limited report in E_2^{new} would be Y_{a_1}. Since Y_{a_1} should in E_2^{new} be limited in the same way as in E_1 (otherwise report Y_{a_2} has influence on the limiting process of prior information), we obtain that E_2^{new} is equal to $E_2^{new} \longleftarrow (1 - w_1) \cdot \mathscr{M}(\{Y_{a_2}\}) + w_1 \cdot \mathscr{M}(\{Y_{a_1}, Y_{a_2}\})$. Now, notice that for report Y_{a_1} we only needed to query model $\mathscr{M}$ once because there were no prior reports. For report Y_{a_2}, we needed to query model $\mathscr{M}$ twice. This can be easily generalized; for example, for the third report Y_{a_3}, we would need to query model $\mathscr{M}$ four times to obtain estimates: $\mathscr{M}(\{Y_{a_3}\})$, $\mathscr{M}(\{Y_{a_1}, Y_{a_3}\})$, $\mathscr{M}(\{Y_{a_2}, Y_{a_3}\})$ and $\mathscr{M}(\{Y_{a_1}, Y_{a_2}, Y_{a_3}\})$. By induction, it follows that:

Proposition 8. *The number of queries to a black box model $\mathscr{M}$ of the influence limiter algorithm in one time period t is $\Omega(2^n)$, where n is the number of the reported values.*

Scoring rule. The properties of the influence limiter are proven only for the quadratic scoring rule (see Lemma 5 in [RS07]). In particular, the score $score_{a,t}$ is calculated by using a function MC_{QSR} that evaluates the marginal contribution of an agent using the quadratic scoring rule. For example, estimates E might contain the likelihood of possible reports of a trusted agent, i.e., $Pr_E(X_{trust})$, which means that scoring function S_E can be defined as $S_E(E, X_{trust}) = S_Q(Pr_E, X_{trust})$, where S_Q is the quadratic scoring rule (see (2.2) in Chapter 2). Since our goal is not to make restrictions on the form of model $\mathscr{M}$, allowing general scoring techniques is crucial in our design. For example, if a model $\mathscr{M}$ is non-probabilistic, a quadratic scoring rule is not applicable.

Furthermore, the influence limiter uses a binary outcome in its scoring rule (this is a requirement of Lemma 5 in [RS07]). In our scenario, the report $Y_{trust} = X_{trust}$ of a trusted agent is not necessarily a binary observation, so one needs to transform it into a binary variable in order to apply it to the influence limiter. For example, if X_{trust} takes values in $\mathbb{R}$, the transformation can be done by defining a threshold and a binary variable equal to 0 if X_{trust} is smaller than the threshold, and 1 otherwise. An issue with this approach is that the evaluation process is much less accurate. For example, if the threshold is equal to 30, then this scoring technique would assign the same quality evaluations for both $X_{trust} = 35$ and $X_{trust} = 50$.

Reputation update. The reputation updating rule of the influence limiter is defined by $\rho_{a,t+1} \longleftarrow \rho_{a,t} + w_{a,t} \cdot score_{a,t}$, and resembles the information fusion updating. This is not a coincidence: a reputation change should reflect how much an agent influences the aggregate.

To lower the query complexity, we investigate a *non-deterministic* information aggregation approach that allows general scoring rules based on non-binary outcomes. These changes also imply a different reputation updating rule. All these structural differences point out that the influence limiter is not trivially transformable to our setting.

5.5 Stochastic influence limiter

The stochastic influence limiter (SIL) is a version of the influence limiter reputation system with an exponential reputation boosting. More precisely, its decision making rule is non-deterministic and uses reputations (weights) that have a multiplicative updating rule.[5]

The exact description of SIL can be found in Algorithm 3, and it has the following steps. Initially, agents' reputations are set to $\rho_0 > 0$. At time period t, upon the arrival of an agent a's report, the reputation system calculates estimate E_a^{new} using function $Update(E, Y_{a,t})$, which adds report $Y_{a,t}$ to the existing set of *included* reports (the set of reports that estimate E) and applies model $\mathcal{M}$ to obtain a new estimate. In the next step, the algorithm decides whether the current estimate should be replaced with the update or not. The decision is probabilistic — with probability equal to $\frac{\rho_{a,t}}{\rho_{a,t}+1}$, the center sets estimate E to E_a^{new}, while otherwise, it discards agent a's report. The final step of the repetitive algorithm is to update the reputation of agent a when the report $Y_{trust} = X_{trust}$ of a trusted agent is received. The reputation updating rule assigns a new reputation to agent a by adding to the current reputation $\rho_{a,t}$ the score of agent a modulated by $\eta \cdot \rho_{a,t}$, where η is a learning parameter. Parameter η should not exceed $\frac{1}{2}$, but its proper value depends on threshold $score_{min}$ that defines the minimum expected score of an informed agent (see Definition 20). As it is shown in the following subsections, a good value for η would be $\eta = \min(\frac{1}{2}, score_{min})$. However, often the expected score $\mathbb{E}(score_{a,t})$ is for high quality reports greater than variance $Var(score_{a,t})$, in which case one can set $\eta = \frac{1}{2}$.

[5] This is similar to the approach of the randomized weighted majority algorithm [LW94, NRTV07], designed for the standard expert setting (see Section 5.2).

Data: Initial reputation $\rho_0 > 0$, learning parameter $\eta \in (0, \frac{1}{2}]$

```
begin
    for Agent a do
        ρ_{a,1} ⟵ ρ_0;
    end
    for t = 1 to t = T do
        Compute prior estimate E;
        Publish E;
        for Agent a do
            Receive a's report Y_{a,t};
            E_a^{old} ⟵ E ;
            E_a^{new} ⟵ Update(E, Y_{a,t}) ;
            if rand(0,1) < ρ_{a,t} / (ρ_{a,t} + 1) then
                E ⟵ E_a^{new};
                Publish E;
            end
        end
        Receive report Y_{trust,t} = X_{trust,t};
        for Agent a do
            score_{a,t} ⟵ S_E(E_a^{new}, X_{trust,t}) − S_E(E_a^{old}, X_{trust,t});
            ρ_{a,t+1} ⟵ ρ_{a,t} · (1 + η · score_{a,t});
        end
    end
end
```

Algorithm 3: Stochastic Influence Limiter

5.5.1 Query complexity

Since the deterministic information fusion rule of the standard influence limiter has an exponential query complexity, we have applied a stochastic information fusion rule in the SIL algorithm. Because of that, SIL has a significantly lower query complexity, in particular, it makes only a constant number of queries per report.

Theorem 13. *The number of queries to a black box model $\mathcal{M}$ of the SIL algorithm in one time period t is $O(n)$, where n is the number of reported values.*

Proof. The SIL's function $Update$ is simple: it uses the set of reports that produced E, say $\{Y_1, ..., Y_k\}$ where $E \longleftarrow \mathcal{M}(\{Y_1, ..., Y_k\})$, adds to it the report $Y_{a,t}$ of agent a and calculates $E_a^{new} \longleftarrow \mathcal{M}(\{Y_1, ..., Y_k\} \cup \{Y_{a,t}\})$. Therefore, SIL makes $O(1)$ queries to $\mathcal{M}$ for $i-th$ agent, thus, for n agents in one time period t we have $O(n)$ queries. □

5.5.2 Bounded negative impact

An important characteristic of SIL is that the probabilistic decision making rule allows a possibility of incorporating reports of agents that are not necessarily considered to be reliable. To make the procedure sound, the probability of fusing a report of an agent with low reputation is low. For example, an agent with reputation 0.1 can affect the current estimate, but only with probability $\frac{0.1}{0.1+1}$. This way, one makes deceiving malicious strategies less effective. In particular, their overall impact cannot be highly negative, meaning that the sum of an agent's contributions, which can be positive and negative, is bounded from below.

Theorem 14. *The total myopic impact $\bar{\Delta}_a = \sum_{t=1}^{T} \bar{\Delta}_{a,t}$ of agent a is in the SIL algorithm bounded from below by:*

$$\bar{\Delta}_a > -\frac{1}{\eta} \cdot \rho_0$$

where ρ_0 is the initial reputation of agent a.

Proof. The expected myopic impact $\bar{\Delta}_{a,t}$ is equal to $\frac{\rho_{a,t}}{\rho_{a,t}+1} \cdot G_{a,t} = \frac{\rho_{a,t}}{\rho_{a,t}+1} \cdot score_{a,t}$. On the other hand, for reputation $\rho_{a,T+1}$ we have:

$$\begin{aligned}
\ln(\rho_{a,T+1}+1) &= \ln(\rho_{a,T} \cdot (1+\eta \cdot score_{a,T}) + 1) = \ln((\rho_{a,T}+1) \cdot (1 + \frac{\rho_{a,T}}{\rho_{a,T}+1} \cdot \eta \cdot score_{a,T})) \\
&= \ln(\rho_{a,T}+1) + \ln(1+\eta \cdot \bar{\Delta}_{a,T}) = ... = \ln(\rho_0+1) + \sum_{t=1}^{T} \ln(1+\eta \cdot \bar{\Delta}_{a,t}) \\
&\leq \ln(\rho_0+1) + \eta \sum_{t=1}^{T} \bar{\Delta}_{a,t} = \ln(\rho_0+1) + \eta \cdot \bar{\Delta}_a
\end{aligned}$$

where we used the fact that $\ln(1+x) \leq x$ for $x > -1$. By noting that the updating rule for reputations keeps the reputations positive, i.e., $\rho_{a,t} > 0$, we have $\ln(\rho_{a,T+1}+1) > 0$, so $\bar{\Delta}_a$ is lower bounded by:

$$\bar{\Delta}_a > -\frac{1}{\eta} \cdot \ln(\rho_0 + 1) \geq -\frac{\rho_0}{\eta}$$

where we again applied $\ln(1+x) \leq x$ for $x > -1$. □

The consequence of Theorem 14 is that the direct damage of a group of m malicious agents can be controlled by setting the agents' initial reputation to a low value. Namely, the impact $\bar{\Delta}_{a,t}$ of agent a at time period t is measured by her marginal contribution, so the total myopic impact of all malicious agents over sensing period T is by Theorem 14 at least $-\frac{1}{\eta} \cdot m \cdot \rho_0$ (i.e., the absolute value of the negative impact is at most $\frac{1}{\eta} \cdot m \cdot \rho_0$). By choosing a small value of ρ_0, one can make the (negative) impact of malicious agents close to 0, regardless of the reporting strategies they use and their reporting time frame. This also implies that, when averaged over a longer sensing period, their negative impact is negligible.[6]

5.5.3 Bounded information loss

The SIL decision making procedure also induces a certain information loss due to the fact that valuable information might be discarded. This is especially true for the initial sensing periods where all agents have relatively low reputations, including the ones that are not malicious. For example, if the reputations are set to $\rho_0 = 0.1$, the probability of including a report from an honest and informed agent is initially equal to $\frac{0.1}{0.1+1}$. Since only information that comes from agents with large reputation scores has a good chance of being considered, informed agents should build up their reputation quickly, which is indeed the case for the SIL algorithm because the reputation increase is exponential. Namely, the increase in the reputation is equal to $\eta \cdot \rho_{a,t} \cdot score_{a,t}$, which for non-malicious agents with predominantly positive scores implies an exponential reputation growth. Therefore, by using the exponential reputation boosting, SIL is capable of limiting the negative influence of malicious agents, while not discarding too many reports of non-malicious agents.

The following theorem formally shows that if an agent reports informed observations, i.e., her scores are positive in expectation and greater than $score_{min}$, then there is a bound to the amount of agent a's information discarded by SIL.

Theorem 15. *Consider an agent a whose reporting strategy does not depend on her reputation $\rho_{a,t}$ and that has expected scores strictly greater than threshold $score_{min}$, i.e., $\mathbb{E}(score_{a,t}) > score_{min} > 0$. Let parameter η be strictly greater than 0 and less than:*

$$\begin{cases} \frac{1}{2} & if\, Var(score_{a,t})) < \mathbb{E}(score_{a,t}) \\ \min(\frac{1}{2}, score_{min}) & if\, Var(score_{a,t})) \geq \mathbb{E}(score_{w,t}) \end{cases}$$

[6] One can reach the same conclusion for a malicious agent with m distinct identities.

Furthermore, let us denote: $g_{a,t} = \ln(1+\eta \cdot score_{a,t}) \in [g_{min,t}, g_{max,t}]$ *and* $h_{a,t} = \mathbb{E}(g_{a,t})$. *Then the expected information loss* IL_a *of the SIL algorithm is bounded from above by:*

$$IL_s = \sum_{t=1}^{T} (\mathbb{E}(score_{a,t}) - \mathbb{E}(\bar{\Delta}_{a,t})) < z \cdot \left[\frac{e^{-\frac{1}{2}\cdot d}}{1-e^{-\frac{1}{2}\cdot d}} + \frac{2 \cdot \ln \frac{\rho_0+1}{\rho_0}}{h} \right]$$

where $z = \max_{1\le t\le T} \mathbb{E}(score_{a,t}) \le 1$, $h = \min_{1\le t\le T}(\frac{1}{t}\sum_{\tau=1}^{t} h_{a,\tau}) \ge \min_{1\le t\le T} h_{a,t} > 0$ *and* $d = \min_{1\le t\le T} \frac{1}{t} \frac{(\sum_{\tau=1}^{t} h_{a,\tau})^2}{\sum_{\tau=1}^{t}[g_{max,\tau} - g_{min,\tau}]^2} > \frac{h^2}{2}$.

Proof. The proof requires two important inequalities from the probability theory.
Markov's inequality states that for a random variable ρ, $b \ge 0$ and monotonically increasing function $f(\cdot) > 0$, we have:

$$Pr(|\rho| \ge b) \le \frac{\mathbb{E}(f(|\rho|))}{f(b)}$$

Hoeffding's inequality states that for independent random variables $\rho_1, \rho_2, ..., \rho_n$ that take values in $\rho_i \in [l_i, u_i]$ and have total expectation $\mathbb{E}(\sum_i \rho_i) = \bar{\rho}$, we have:

$$Pr(\sum_i \rho_i - \bar{\rho} \ge t) \le e^{-2\cdot \frac{t^2}{\sum_{i=1}^{n}(u_i - l_i)^2}}$$

$$Pr(\sum_i \rho_i - \bar{\rho} \le -t) \le e^{-2\cdot \frac{t^2}{\sum_{i=1}^{n}(u_i - l_i)^2}}$$

Now we are ready to prove the statement.

The expected value of the myopic impact is:

$$\mathbb{E}(\bar{\Delta}_{a,t}) = \mathbb{E}\left(\frac{\rho_{a,t}}{\rho_{a,t}+1} \cdot score_{a,t}\right)$$

Since scores are stochastically generated (they are independent of reputation $\rho_{a,t}$), we obtain that:

$$\mathbb{E}(\bar{\Delta}_{a,t}) = \mathbb{E}\left(\frac{\rho_{a,t}}{\rho_{a,t}+1}\right) \cdot \mathbb{E}(score_{a,t})$$

Furthermore, Markov's inequality gives us:

$$\mathbb{E}\left(\frac{\rho_{a,t}}{\rho_{a,t}+1}\right) \ge Pr(\rho_{a,t} \ge \rho_0 \cdot b_t) \cdot \frac{\rho_0 \cdot b_t}{\rho_0 \cdot b_t + 1}$$

where we used: $b_t = e^{\frac{1}{2}\cdot \sum_{\tau=1}^{t} h_{a,\tau}}$, $h_{a,\tau} = \mathbb{E}(\ln(1+\eta \cdot score_{a,\tau}))$. Let us also denote: $h = \min_{1\le t\le T} \frac{1}{t}\sum_{\tau=1}^{t} h_{a,\tau}$. Using $\ln(1+x) \ge x - x^2$ for $x \ge -\frac{1}{2}$, it follows that $h_{a,t} \ge \eta \cdot$

$\mathbb{E}(score_{a,t}) - \eta^2 \cdot \mathbb{E}((score_{a,t})^2)$. Due to the conditions of the theorem, we know that $\eta < \mathbb{E}(score_{a,t})$ or $\mathbb{E}(score_{a,t}) > \frac{1}{2} \cdot \mathbb{E}((score_{a,t})^2)$ (when $Var(score_{a,t}) < \mathbb{E}(score_{a,t})$), which by $score_{a,t} \in [-1,1]$, implies that $h > 0$.

Now, notice that:

$$\begin{aligned}
Pr(\rho_{a,t} \geq \rho_0 \cdot b_t) &= Pr(\ln \rho_{a,t} \geq \ln(\rho_0 \cdot b_t)) = Pr\left(\ln \rho_{a,t} \geq \ln \rho_0 + \frac{1}{2} \cdot \sum_{\tau=1}^{t} h_{a,\tau}\right) \\
&= Pr\left(\ln \rho_{a,t} - \sum_{\tau=1}^{t} h_{a,\tau} - \ln \rho_0 \geq -\frac{1}{2} \cdot \sum_{\tau=1}^{t} h_{a,\tau}\right) \\
&\geq 1 - Pr\left(\ln \rho_{a,t} - \sum_{\tau=1}^{t} h_{a,\tau} - \ln \rho_0 \leq -\frac{1}{2} \cdot \sum_{\tau=1}^{t} h_{a,\tau}\right) = 1 - p_t
\end{aligned}$$

where we denoted the last term $Pr(.)$ by p_t. Since $\ln \rho_{a,t} - \ln \rho_0$ is a sum of t independent random variables $g_{a,\tau} = \ln(1 + \eta \cdot score_{a,\tau})$ (with $1 \leq \tau \leq t$) that are in expectation equal to $h_{s,\tau} = \mathbb{E}(g_{a,\tau})$, using Hoeffding's inequality, we obtain:

$$p_t \leq e^{-\frac{2 \cdot (\sum_{\tau=1}^{t} h_{a,\tau})^2}{4 \cdot \sum_{\tau=1}^{t} [g_{max,\tau} - g_{min,\tau}]^2}} \leq e^{-\frac{2 \cdot \left(\frac{\sum_{\tau=1}^{t} h_{a,\tau}}{t}\right)^2 \cdot t}{4 \cdot \frac{\sum_{\tau=1}^{t} [g_{max,\tau} - g_{min,\tau}]^2}{t}}} \leq e^{-\frac{1}{2} \cdot d \cdot t}$$

where we put $d = \min_{1 \leq t \leq T} \frac{\left(\frac{\sum_{\tau=1}^{t} h_{a,\tau}}{t}\right)^2}{\frac{\sum_{\tau=1}^{t} [g_{max,\tau} - g_{min,\tau}]^2}{t}}$, which is greater than $d > \frac{h^2}{2}$ because $\eta \cdot score_{s,\tau} \in [-0.5, 0.5]$ (and, hence, $[g_{max,\tau} - g_{min,\tau}]^2 < 2$). The expected information loss (the difference between the agent's score and its impact) in round t is bounded by:

$$\begin{aligned}
&\mathbb{E}(score_{a,t}) - \mathbb{E}(\bar{\Delta}_{a,t}) = \mathbb{E}(score_{a,t}) \cdot \left(1 - \mathbb{E}\left(\frac{\rho_{a,t}}{\rho_{a,t} + 1}\right)\right) \\
&\leq \mathbb{E}(score_{a,t}) \cdot \left[1 - (1 - e^{-\frac{1}{2} \cdot d \cdot t}) \cdot \frac{\rho_0 \cdot b_t}{\rho_0 \cdot b_t + 1}\right] \\
&= \mathbb{E}(score_{a,t}) \cdot \left[\frac{1}{\rho_0 \cdot b_t + 1} + e^{-\frac{1}{2} \cdot d \cdot t} \cdot \frac{\rho_0 \cdot b_t}{\rho_0 \cdot b_t + 1}\right]
\end{aligned}$$

Therefore, over time period T, the information loss is in expectation upper bounded by:

$$z \cdot \left[\sum_{t=1}^{T} \frac{1}{\rho_0 \cdot b_t + 1} + \sum_{t=1}^{T} e^{-\frac{1}{2} \cdot d \cdot t} \cdot \frac{\rho_0 \cdot b_t}{\rho_0 \cdot b_t + 1}\right]$$

where $z = \max_{1 \leq t \leq T} \mathbb{E}(score_{a,t})$. We examine bounds for each of the terms in the bracket. We have:

$$\begin{aligned}
\sum_{t=1}^{T} e^{-\frac{1}{2} \cdot d \cdot t} \cdot \frac{\rho_0 \cdot b_t}{\rho_0 \cdot b_t + 1} &\leq \sum_{t=1}^{T} e^{-\frac{1}{2} \cdot d \cdot t} = e^{-\frac{1}{2} \cdot d} \cdot \sum_{t=0}^{T-1} e^{-\frac{1}{2} \cdot d \cdot t} \\
&< e^{-\frac{1}{2} \cdot d} \cdot \sum_{t=0}^{\infty} e^{-\frac{1}{2} \cdot d \cdot t} = \frac{e^{-\frac{1}{2} \cdot d}}{1 - e^{-\frac{1}{2} \cdot d}}
\end{aligned}$$

where we applied $\sum_{t=0}^{\infty} x^t = \frac{1}{1-x}$ for $x \in (0,1)$. Furthermore, using the fact that $b_t = e^{\frac{1}{2}\sum_{\tau=1}^{t} h_{s,\tau}} \geq e^{\frac{1}{2}\cdot t\cdot h}$ we obtain:

$$\sum_{t=1}^{T} \frac{1}{\rho_0 \cdot b_t + 1} \leq \sum_{t=1}^{T} \frac{1}{\rho_0 \cdot e^{\frac{1}{2}\cdot t\cdot h} + 1} \leq \int_{t=0}^{T} \frac{1}{\rho_0 \cdot e^{\frac{1}{2}\cdot t\cdot h} + 1} dt < \int_{t=0}^{\infty} \frac{1}{\rho_0 \cdot e^{\frac{1}{2}\cdot t\cdot h} + 1} dt$$
$$= \frac{2}{h} \cdot \ln \frac{\rho_0 + 1}{\rho_0}$$

which completes the proof. □

The intuition behind this result is fairly simple. If an agent has mostly positive scores greater than $score_{min}$, it will boost up her reputation rather quickly to the values where her reports are practically no longer limited. Notice that the bound on the total information loss does not (directly) depend on time (i.e., does not monotonically increase with time), which means that the information loss averaged over a long sensing period T becomes negligible. Furthermore, the bound multiplicatively depends on parameter z that represents an agent's expected score: the better the agent is, the more quality information the center looses when it discards the agent's reports. The second multiplicand in the bound describes how quickly an agent can boost up its reputation, which depends on how *informative* the agent is: the more useful the agent's reports are, the higher its score is, and thus the greater its reputation increase is. This is captured by parameters h and d, which are related to the performance of an agent through random variable $g_{a,t} = \ln(1 + \eta \cdot score_{a,t})$. Notice that by Theorem 15, we can set $z = 1$ and $d = \frac{h^2}{2}$ in order to obtain a looser upper bound that does not require estimates of z and d.

Theorem 14 and Theorem 15 provide guarantees on the performance of the SIL algorithm that depend on initial reputation ρ_0 and learning parameter η. The bounds of the theorems indicate that the value of the initial reputation ρ_0 should be such that it limits the negative impact of malicious agents, while not discarding too much information from non-malicious agents. Since for a longer sensing period accurate agents have enough time to build up their reputations, the initial reputation ρ_0 can be set to a relatively small value so that the SIL algorithm is more robust against malicious reporting strategies.

A proper value for the learning parameter, on the other hand, depends on how informed good agents should be and whether the qualities of their reports are allowed to have a high variance. If the center considers that informed agents are only those that consistently report good information, i.e., those who have positive expected scores with low variance, then η can be set to $\frac{1}{2}$. On the other hand, if the center considers that informed agents are all those that are expected to provide good information, i.e., those that have expected scores strictly greater than $score_{min}$, then a good value of η is $\min(score_{min}, \frac{1}{2})$. Notice that the bound on negative impact is inversely proportional to η. Therefore, low values of η should be avoided.

5.5.4 Helpful reporting

Finally, we analyze the incentive component of the SIL algorithm. The important property of agents' scores, which define payments τ_G, is that they incentivize non-malicious agents to provide reports that maximize the information gain of the center. Notice that the most *useful* information is not necessarily the true observations. This is due to the presence of malicious agents, as well as the possible imperfections of aggregation model $\mathscr{M}$. In other words, a strategic behaviour is often desirable.

Theorem 16. *If an agent a maximizes her expected payoff $\mathbb{E}(\tau_G(Y_a, X_{trust})) = \mathbb{E}(score_{a,t})$, then she also maximizes her expected impact $\mathbb{E}(\bar{\Delta}_{a,t})$.*

Proof. The myopic impact of agent a, $\bar{\Delta}_{a,t}$, is proportional to her score: $\bar{\Delta}_{a,t} = \frac{\rho_{a,t}}{\rho_{a,t}+1} \cdot G_{a,t} = \frac{\rho_{a,t}}{\rho_{a,t}+1} \cdot score_{a,t}$. Hence, an agent a that aims to maximize her expected score, is also incentivized to submit a report that maximizes her expected impact. □

5.6 Application to community sensing

We consider a community sensing scenario where the center aggregates crowdsensed information in an online manner, from both public and private sensors, to provide real time estimates of air pollution over a certain urban area. In this scenario, the center controls a few accurate sensors that provide spatially or temporally sparse measurements (e.g., very accurate particle sensors are slow; similarly NO_2 can be sensed chemically but it's again slow and expensive), so to properly monitor the localized features of air pollution, it complements its own measurements with those obtained by crowd-participants who own ubiquitous sensor devices.

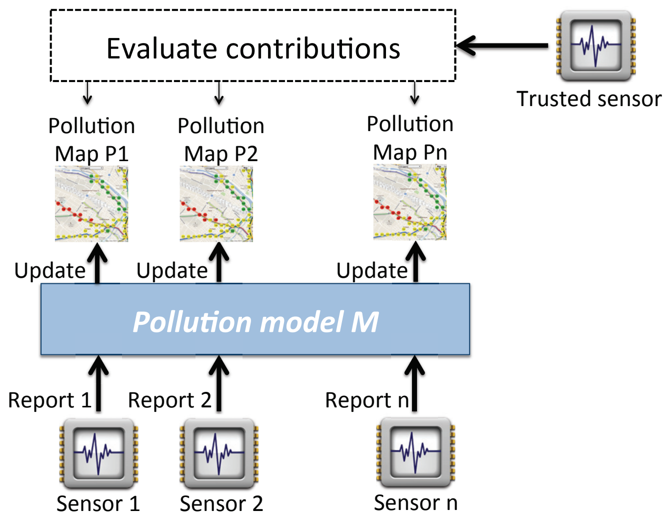


Figure 5.2 – Community sensing scenario with online information fusion

The model of the information fusion process that we consider in this section is depicted in Figure 5.2, and it follows the formal model of this chapter. At the beginning, the center has only prior information about air pollution over an urban area. After some time, the center receives a report from a crowd-sensor and merges it with the current pollution map using pollution (aggregation) model $\mathscr{M}$. Clearly, apart from its measurement, the sensor also reports its location, which we do not explicitly emphasize in the further text. Moreover, pollution model $\mathscr{M}$ is assumed to capture correlations among measurements taken at different locations well. The described process repeats until a trusted sensor reports its measurement, after which the center can evaluate the reports of the crowd-sensors. The crowd sensing process then continues in the same manner.

One of the main challenges in the described scenario is how to cope with untrustworthy information. For example, a factory owner who wants to hide her own pollution traces could install sensors that misreport values of pollution. Clearly, incentive schemes alone cannot provide quality control that would solve the problem of participants with ulterior motives. Reputation systems provide such a guarantee: bad reports lead to low reputation, which limits the influence of the later reports. Therefore, we investigate the application of the stochastic influence limiter in the considered setting and compare its performance to the performance of the Beta reputation system with the thresholding principle, which is a state of the art reputation system for sensing.

5.6.1 Simulation setup

Considering that in a real dataset one cannot identify upfront the strategies adopted by different sensors, we simulate different malicious strategies to experimentally validate our approach. Our pollution sensing scenario is based on the testbed from Chapter 4, section 4.5.2, but now sensors' measurements are not discretized. In total, the dataset contains approximately one month of hourly measurements - the larger sensing periods can be simulated by looping over the dataset several times, which we do 12 times to obtain the sensing time of $T = 12 \cdot 4 \cdot 7 \cdot 24$ hours. Our main reputation system is SIL with the initial reputation set to $\rho_0 = 0.1$.

Pollution Model

We use a probabilistic air pollution model that is based on Gaussian process regression, as described in [RW05]. For any point of interest (in our case 116 locations), the pre-trained Gaussian Process (GP) model produces a probability distribution function over the possible levels of pollution from the reports of sensors placed at different locations. This posterior distribution is a normal distribution $\mathcal{N}(\mu, \sigma)$, with parameters μ and σ derived from the GP model. We are interested in predicting the value of pollution level measures by a trusted sensor at its location, so we denote the corresponding prediction by $p(X_{trust})$.

Performance Measure

We measure the quality of the aggregates of model $\mathscr{M}$ by how well they predict the measurement X_{trust} of a trusted sensor. Since model $\mathscr{M}$ outputs a normal distribution $\mathscr{N}(\mu,\sigma)$ for a point of interest (x, y), we apply scoring rule (2.3) (see Chapter 2) on probability density function p of the form $p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ to obtain:

$$S_E(p, X_{trust}) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(X_{trust}-\mu)^2}{2\sigma^2}} - \frac{1}{4\sigma\sqrt{\pi}} \tag{5.2}$$

The score takes values in $[-\frac{1}{4\sigma\sqrt{\pi}}, \frac{1}{\sigma\sqrt{\pi}}(\frac{1}{\sqrt{2}} - \frac{1}{4})]$, and can be further scaled so that $score_{a,t} \in [-1, 1]$. In our case, no specific scaling was needed.

Sensors

We consider 40 mobile crowd-sensors and 1 trusted mobile sensor that are at each time period placed at one of 116 available locations. The 40 crowd-sensors are either honest (25% of them) or malicious sensors (75% of them). Malicious sensors report according to one of the following four strategies. In the *Vary* strategy, sensors build up their reputations by reporting honestly for the first 1000 iterations, and from then on, they report only a low level of pollution. In the *Deceive* strategy, sensors report honestly when their reputation is below 0.5; otherwise, they report a low level of pollution. *Vary and Deceive* is a mixed strategy where malicious sensors first build up their reputation by reporting honestly for 1000 iterations, and from then on, they use the *Deceive* strategy. *Cover* is a strategy that mimics a situation where malicious sensors try to boost up their reputation when it is not important for them to misreport, and then, on specific events, they report wrong values. In our case, malicious sensors boost up their reputation for 1000 iterations. Then they report honestly whenever the pollution is below 35 ppb of NO_2 or their reputation is lower than 0.5; otherwise, they report a low level of pollution. The low level of pollution in the above strategies is defined as 10 ppb of NO_2 plus a Gaussian noise with 0 mean and the standard deviation equal to 5.

Theoretical Bound

By Theorem 14, it follows that $0.75 \cdot 40 = 30$ malicious sensors can cause an immediate damage of at most $2 \cdot 30 \cdot 0.1 = 6$ score units (units used in (5.2)). To calculate the bound from Theorem 15, one needs to decide on parameter η and estimate parameters z, h and d. We set $\eta = \frac{1}{2}$, and approximate z, h and d, by investigating averages of $score_{a,t}$, $\log(1 + \frac{1}{2} \cdot score_{a,t})$ and $\max_a \log(1 + \frac{1}{2} \cdot score_{a,t}) - \min_a \log(1 + \frac{1}{2} \cdot score_{a,t})$ over time t. Assuming that the scores of honest sensors are similar in most of the sensing periods, these averages lead to the estimates:[7]

[7] If $k << T$ sensing periods have significantly different values from the average values, to achieve a higher precision, one can exclude these k periods when estimating the upper bound from Theorem 15 and simply add to the calculated bound $k \cdot \max_{\tau \in kPeriods} \mathbb{E}(score_{a,\tau})$.

$z \approx 0.002$, $h \approx 0.001$ and $d \approx 0.005$, from which we can estimate the upper bound from Theorem 15: 10.39. By multiplying the estimate by the number of honest sensors (i.e., 10), we conclude that the total information loss should be no more than 103.9 score units. Notice that the bounds from Theorem 14 and Theorem 15 have different meanings: the bound from Theorem 14 describes how much a malicious sensor could intentionally shift the result, while the bound from Theorem 15 describes an implicit damage whose nature is not controlled by a malicious sensor. Nevertheless, it follows from the bounds that the quality degradation should not be more than 109.9 score units in total. This can be averaged over time, so that at each time step t, we have an average degradation of at most $\frac{109.9}{t}$ score units. The average goes to 0 as time increases implying a no-regret property in terms of sensors' myopic impact.

Baseline: Beta Reputation System

In the Beta reputation system, we quantify the behaviour of a sensor using two parameters, α and β, which represent the parameters of the beta distribution $B(\alpha, \beta)$[8]. In the setting we analyze, the parameters can be updated as follows (e.g., see [JI02]). If the marginal information gain $G_{a,t} = score_{a,t}$ of updating the current pollution map with a sensor a's report is positive, parameter $\alpha_{a,t}$ is updated to $\alpha_{a,t+1} = \alpha_{a,t} + G_{a,t}$. Otherwise, parameter β is updated to $\beta_{a,t+1} = \beta_{a,t} + G_{a,t}$. The reputation of sensor a is at time t calculated as the mean of beta distribution $B(\alpha_{a,t}, \beta_{a,t})$, i.e., $\rho_{a,t} = \frac{\alpha_{a,t}}{\alpha_{a,t}+\beta_{a,t}}$. In other words, the reputation of sensor a characterizes the fraction of the positive impact that the sensor had on the system. The decision on whether to include the report of sensor a is based on its reputation and determined using the thresholding principle. We set the initial values of α and β parameters to 0.01 and 0.1, respectively, with threshold $\Theta = 0.5$.

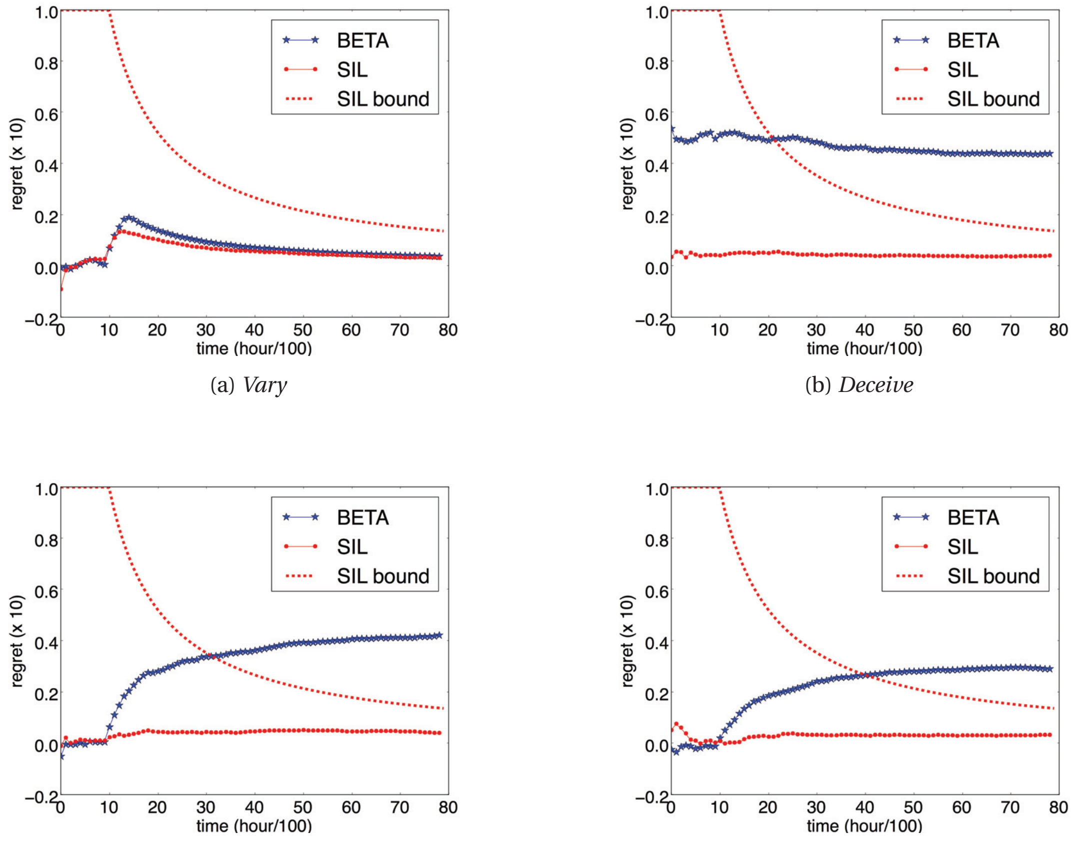
Evaluation Metric

We define a measure of an average regret that evaluates the quality of the aggregates produced by the center with respect to the aggregates obtained by fusing the reports of honest sensors. More precisely:

$$AvgRegret_t = \frac{Score_{honest,t} - Score_{center,t}}{t}$$

where $Score_{honest,t}$ is the total score (until the time period t) of the aggregates obtained from the reports of honest sensors, and $Score_{center,t}$ is the total score of the center (with a particular reputation system) until the time period t. Both scores are calculated using the quadratic scoring rule, as described in the previous subsections, applied on the pollution map published prior to the report of a trusted sensor. Therefore, the regret is measured in the same score units as the theoretical bound computed in subsection 'Theoretical bound'.

[8] Notice that α and β are parameters of the distribution, not the scaling parameters defined in Chapter 2



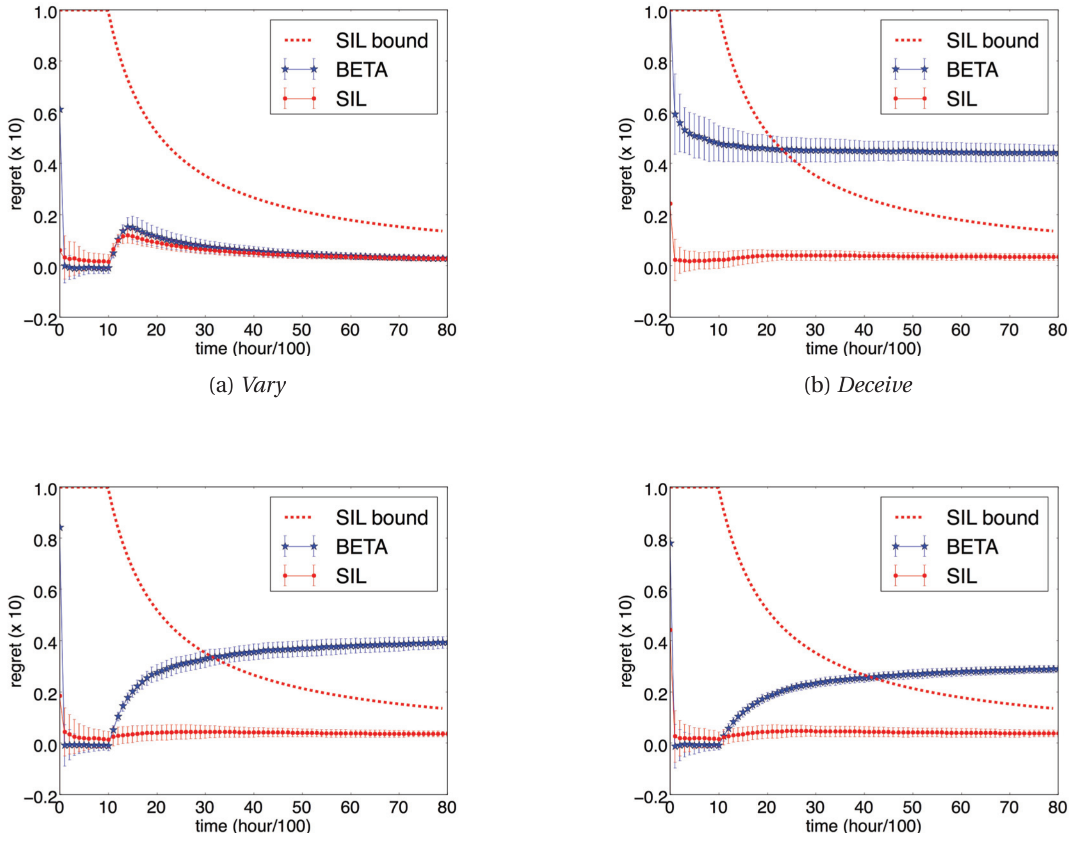
(a) *Vary* (b) *Deceive*

(c) *Vary and Deceive* (d) *Cover*

Figure 5.3 – Average regrets (times 10) for different strategies (single simulation run)

5.6.2 Simulation results

Figures 5.3 and 5.4 show the performance of the SIL algorithm and the Beta reputation system in terms of the average regret for four different misreporting strategies. Along with those results, we put the theoretical estimate of the upper bound on the regret of SIL algorithm ($\frac{109.9}{t}$), which is truncated to 0.1 for large values. The Beta reputation system is able to limit the negative influence of malicious sensors that use the *Vary* strategy. However, in the *Vary* strategy, malicious sensors misreport in a simple and consistent way. For the other three misreporting strategies, the Beta reputation system experiences an average regret that is clearly away from 0, and in two of the cases, the regret is increasing, which means that the total negative impact of malicious sensors is not bounded. The SIL algorithm is much better in dealing with malicious sensors: its average regret over a longer sensing period for all of the malicious strategies is close to 0, as expected by the theoretical results. Finally, the strategy independent upper bound on the SIL's regret is often below the regret of the Beta reputation system.



(a) *Vary*

(b) *Deceive*

(c) *Vary and Deceive*

(d) *Cover*

Figure 5.4 – Average regrets (times 10) for different strategies — the figure shows the mean and the 95% confidence interval of 50 simulation runs

5.7 Conclusion

In this chapter, we discussed a problem of having malicious agents in online information fusion. We designed a novel reputation system, called the stochastic influence limiter (SIL), that has a manageable complexity and puts an upper bound on the total negative impact that malicious agents can have on the fused result, regardless of their reporting strategy. This is in contrast to the standard reputation systems which do not provide any theoretical guarantees and for which the total negative impact of malicious agents can increase over time.

Due to its theoretical guarantees, we advocated the use of the SIL algorithm in the community sensing. We empirically confirmed that the theoretical results hold in a realistic air pollution sensing scenario, and showed that in an average-case simulation, SIL outperforms a state of the art reputation system for sensing, whose performance is often worse than the worst case performance of SIL.

The most interesting direction for future work would be to investigate under which conditions the SIL mechanism provides guarantees when the evaluation is based on a peer consistency approach. We expect that in this case, a majority of agents' population should provide accurate information. Furthermore, a payment mechanism needs to incentivize strategic agents to report honestly even when a fraction of agents is dishonest, which is not possible for an arbitrary population statistics. Namely, the presence of malicious agents changes the character of belief conditions necessary to achieve properness of peer consistency techniques. For example, the self-predicting condition, i.e., the belief condition under which the robust peer truth serum is proper, might no longer hold when a fraction of agents is dishonest.

6 Concluding remarks

Ensuring the accuracy of reported information is a major challenge for using crowds as part of intelligent systems. In this thesis, we focused on two aspects important to achieve a good quality control:

- incentivizing rational participants to acquire and report accurate information;
- filtering out low quality reports from participants that do not respond to incentives.

Incentive mechanism design

Instead of fixed rewards, participants should expect the highest rewards when they report accurate information. Such mechanisms can be appropriately scaled so that only participants who provide positive contributions profit from participating in the elicitation process. This is important for two reasons:

- to improve the accuracy of reported information, and thus complement filtering mechanisms such as gold tasks and reputation systems;
- to make participant self-selection help the mechanism by discouraging participants that do not contribute useful results.

We focused on an incentive mechanism design in which the center cannot directly verify the obtained information. Therefore, it has to compare reports in order to reward the participants — agents. Two elicitation settings were investigated, a single-task setting and a multi-task setting, which differ in the number of phenomena (tasks) that agents observe (solve).

For the single-task setting, in which agents observe a single phenomenon, we showed that an additional report is needed whenever the agents' beliefs are not highly constrained and the center does not know them. We designed several Bayesian truth serum (BTS) mechanisms that from each agent elicit targeted information and prediction report regarding what the other agents have reported. The most general mechanism, the divergence-based BTS, allows agents to have different belief systems as long as their beliefs are more similar when they observe the same value than when their observations are different. In the limit case when observations take real values, a version of the divergence-based BTS for continuous domains (the continuous BTS) is strictly proper if agents have a common belief system, and we show that this condition cannot be further relaxed under reasonable constraints on the agents' belief systems.

Unlike the single-task setting, the multi-task setting allows agents to have different beliefs, but it assumes that the observed phenomena are a priori similar. The multi-task mechanisms presented in the thesis achieve this result by appropriately sampling reports from different phenomena in order to calculate the statistics used for scoring an agent. Depending on the scoring technique, it is further possible to make truthful reporting the highest paying strategy profile.

Future directions

While the aforementioned mechanisms have proven theoretical properties in their own domains, and some of the mechanisms even offer experimental evidence of their superiority over their predecessors, there is still a lack of understanding on how to exactly implement these mechanisms in practice and how to make them robust enough for general use. An appropriate implementation choice is often tied to the application domain, especially considering the fact that some of the mentioned mechanisms are designed for specific scenarios, e.g., crowdsourcing.

Since data exchange systems are increasingly dealing with multi-dimensional information structures, one of the important future directions would be to apply peer consistency techniques in the elicitation of complex information. In order to examine consistency of peer reports, an elicitation process might have to rely on domain specific knowledge or, alternatively, be able to automatically discover relevant features of the elicited information. While some of the existing techniques provide relatively stable incentives even for large observation spaces (e.g., see [SAFP16]), there is still a need for proper incentives in a more general formal setting, for example, the one that includes a non-binary participation (effort) choice.

Furthermore, most of elicitation mechanisms either belong to the gold standard or the peer consistency mechanisms. At first glance, it does not seem sensible to use a hybrid approach as there is a gold standard, but an argument for doing so is more obvious when we go beyond a single shot elicitation scenario. Namely, in a scenario where participants repeatedly interact with the center, the center would require a large number of gold standard evaluations to frequently provide proper rewards to the participants. Therefore, by using a hybrid incentive mechanism, one can make agents' strategies converge towards truthful reporting in natural game playing dynamics (e.g., regret minimization) while minimizing the number of gold standard evaluations. This would also strengthen the incentives provided by multi-task peer consistency techniques, making them potentially resilient to a wider range of misreporting strategies, e.g., to the one discussed in [GWL16].

Finally, the existing mechanisms usually model participants as risk-neutral rational agents that maximize their rewards. While this model is fairly reasonable when dealing with, for example, intelligent software agents, it fails to capture different risk attitudes and bounded rationality of human participants. One of the relevant future steps would be to enrich the existing models by incorporating the aspects of behavioral game and economic theories. Such an approach has already been proposed for a specific scenario of designing optimal contests [EG15], but it is yet to be seen how *behavioral mechanism design* should be applied in the case of peer consistency mechanisms.

Information aggregation

To filter the low quality information coming from malicious participants, we designed a reputation system that has provable guarantees on the amount of negative impact that a malicious participant can have on the learned outcome. The novel reputation system, called the stochastic influence limiter (SIL), has two components that differentiate it from the traditional reputation system design: its reputation updating procedure and stochastic information fusion component. The former component has a form of exponential reputation boosting, while the latter one fuses reports probabilistically with the probability of fusion being dependent on the reputation of a participant. Consequently, SIL also discards some information coming from informed participants, but we show that the amount of discarded information is bounded from above. As an example of practical application, we considered sensing in which the center is in control of a few *trusted* sensors that periodically report their measurements, but supports these measurements with the community of sensors. We showed that the SIL algorithm outperforms a baseline algorithm often used in sensing.

Future directions

While our work addresses the issue of adversary participants, it is restricted to the settings where a mechanism can accurately evaluate the inputs provided by the users once the full aggregate is obtained. The next step is to remove the requirement of trusted information source and make an evaluation procedure based on peer consistency methods instead. In particular, the idea would be to construct an online information fusion process that is capable of limiting the negative influence of adversary participants but does not require trusted information to evaluate the influence of a participant on an aggregated result. Clearly, some constraints are needed either in terms of the percentage of adversary participants, as discussed in similar approaches [KOS13], or in terms of the strategy space of adversaries, as studied in the multi-task peer consistency mechanisms. Nevertheless, the empirical results in [PS06] show that it is possible to extract the correct information from crowds using a peer consistency score, even when the majority is wrong. This indicates that a combination of the stochastic influence limiter and multi-task peer consistency mechanisms has great potential in resolving this issue.

A Appendix

A.1 Predictemo game: contest with subjective information

We follow the design of the optimal contest for simple agents, introduced in [GR14], where agents strategize on participation due to the cost that they experience from participating in the contest. In the considered application, the cost of participating models privacy costs that the users experience by playing the Predictemo game, as they provide their identities along with their emotions. The privacy costs are expected to be the same for each participant, which, in the contest model of [GR14], means that the agents have a homogenous cost of participation.

Assuming that the participants do not strategize on the quality of their contributions, the contest that maximizes participation provides m equal rewards to the first m participants, where m depends on the participation cost and the budget V of the contest designer. We explore this approach in a repetitive scenario where the contest is run over a longer period of time. This enables us to eliminate the need of knowing the participation cost, and instead learn the optimal choice of m over time. Furthermore, by applying a BTS type of mechanism (co-BTS from Section 3.5.2 of Chapter 3), we assign a proper quality scores to reported information, which discourages the rational participants from falsely reporting their emotions. As already mentioned, the users of the Predictemo application also provide their coarse-grained locations. However, the reports contain only the information about the closest building to the location of a user, so we assume that this type of information is not misreported.

Optimizing the participation

We consider a contest design in which, at each time step t, participants report their subjective information to a contest designer. A participant can report several reports during a time period t, each report being scored with the co-BTS mechanism against another agent that reports at approximately the same time. Due to the fact that co-BTS rewards agents using a zero-sum reward structure, an agent effectively needs to outperform her peers in order to have a good relative score. At the end of the period, agents are ranked in decreasing order by their total score, and first m participants are rewarded with V/m, where V is the budget of a contest designer. The number m is a time dependent variable, chosen by the designer, and takes values from $\{1,...,m_{max}\}$. The maximum value m_{max} should be such that V/m_{max} represents a lower bound on a reward that covers the cost of participation. Notice that this bound does not need to be tight, so the center does not need to know the exact value of the cost.

We model the number of participants at a certain time step $t \in \{1,...,T\}$ as a random variable N_m that takes value in $\mathbb{N}$. In particular, we assume that N_m is a random sample of a Poisson distribution whose parameter λ_m depends on the number of the rewarded agents m. With this in mind, we would like to select m that leads to the greatest expected number of participants, i.e., that has the associated Poisson distribution with the greatest mean λ_m.

The problem of finding the optimal m in an online manner belongs to a general class of the multi-armed bandit problems (e.g. [ACBF02]), that investigate explore-exploit tradeoffs in online learning processes. We apply the KL-UCB algorithm [GC11] as it allows the objective quantity, in our case participation rate, to be distributed according to a Poisson distribution.

The algorithmic description of our approach is shown in Algorithm 4, and we call it KLUCoBits to indicate that it is based on the KL-UCB algorithm with the co-BTS evaluation procedure. The algorithm follows the steps explained in the above paragraphs, incorporated into the KL-UCB algorithm. The first m steps of the algorithm sample participation rate for different choices of m. Afterwards, the algorithm makes a more appropriate choice of m using procedure $KL\_UCB\_SelectM()$ that implements the arm choice function of KL-UCB, as explained in [GC11]. For each possible choice of m, the procedure takes into account the obtained participation rates for the considered m, but also the number of times the choice was made in order to achieve a good explore-exploit tradeoff.

A.1.1 Predictemo game

The Predictemo game represents an implementation of the KLUCoBits contest design in eliciting emotions across EPFL campus. Once logged in, a user chooses another player and challenges the player to play the Predictemo prediction task. As pointed out in Section 3.5.2 of Chapter 3, each player in the prediction task provides one of 20 possible emotions and the prediction about what the other player will report.

At the end of each period, m best users are rewarded with V/m points. m is selected upfront, before each period, and is known to players, as well as V. Users of the Predictemo application can access a simplified description of the KLUCoBits reward mechanism from the login page, or from their user profiles when they are logged in. The description outlines the basic concepts of the reward mechanism. It states that V points are periodically given to m best users in equal split, where users are ranked by the relative quality of the predictions they provide. The quality of a prediction is said to be measured by how accurate the prediction is plus how hard it was to predict the correct emotion, while the score that affects the ranking of a user is

Data: Time horizon $T > 0$, budget V, max number of rewards m_{max}
begin
  for $t = 1$ *to* $t = m_{max}$ **do**
    $m = t$;
    Publish the number of rewards m and their value V/m;
    Evaluate reports of agents using co-BTS;
    when period t **ends do**
      Get participation rate $N_{m,t}$;
      Rank agents according to their total co-BTS scores;
      Reward first m agents with V/m;
      $NumM[m] = 1$;
      $TotalN[m] = N_{m,t}$;
    endwhen
  end
  for $t = m_{max} + 1$ *to* $t = T$ **do**
    $m = KL\_UCB\_SelectM(NumM, TotalN, t)$;
    Publish the number of rewards m and their value V/m;
    Evaluate reports of agents using co-BTS;
    when period t **ends do**
      Get participation rate $N_{m,t}$;
      Rank agents according to their total co-BTS scores;
      Reward first m agents with V/m;
      $NumM[m] = NumM[m] + 1$;
      $TotalN[m] = TotalN[m] + N_{m,t}$;
    endwhen
  end
end

Algorithm 4: KLUCoBits: multi-step contest with subjective information explained as the difference between the qualities of the user's prediction and the prediction of her opponent in the considered prediction task.

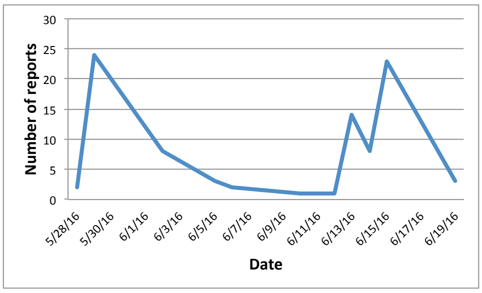


Figure A.1 – Report frequency

The preliminary version of the Predictemo application was deployed for a period of three weeks. A reporting period t was set to be three days, having in total $T = 7$ full periods, or equivalently, 21 days. We rewarded m best players of the Predictemo game with $V = 60$ points at each time step t. m was chosen according to KLUCoBits from the set $\{1, 2, 3, 4\}$. In this time horizon, a player could collect points and exchange 60 points for a gift card worth 20 CHF.

In total, there were 15 EPFL students using the Predictemo application in the non-anonymous (game) mode. Out of 89 submitted reports, 34.8% were anonymous reports, and the rest (65.2%) were reports from the users whose identity is known. As shown in Figure A.1, the participation rate fluctuated throughout the reporting time horizon T, but on average was higher at the end of the 3 week period.

The preliminary version of the application offers basic insights into students' emotional states, such as the location at which users are more likely to report their emotions or the types of the most frequent emotions, as shown in Figure A.2. The most interesting direction for future work would be to link the reported emotions to the EPFL courses, and see how the reported information compares to the traditional opinion polls that ask students to evaluate the quality of the courses.

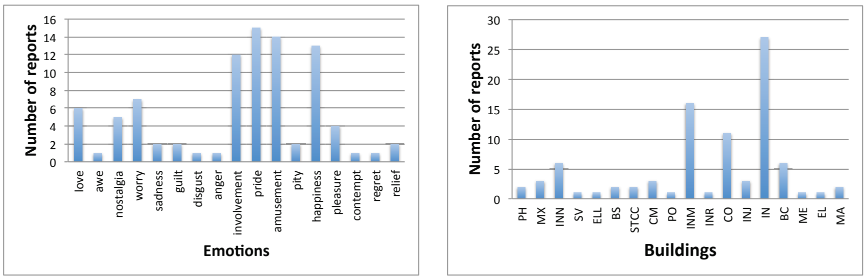
A.2 Geometric sequence

This section provides useful properties of geometric sequence $1 + x + x^2 + ... + x^{n-2}$, $x \in (0, 1)$. Its closed form is:

$$\frac{1 - x^{n-1}}{1 - x} = 1 + x + x^2 + ... + x^{n-2} \tag{A.1}$$

Another property that we will use is the derivative of the geometric sequence:

$$\frac{d}{dx}(1 + x + x^2 + ... + x^{n-2}) = \frac{d}{dx}\left(\frac{1 - x^{n-1}}{1 - x}\right) = \frac{(1 - x^{n-1}) - (n-1) \cdot x^{n-2} \cdot (1 - x)}{(1 - x)^2} \tag{A.2}$$



(a) Emotion distribution

(b) Location distribution

Figure A.2 – Distribution of reports

Next, we show that expression $(1+r\cdot x)\cdot(\sum_{i=0}^{n-2}(1-p-x)^i)$, where $1>p>0$, $r\in(0,\frac{1}{p})$ and $x\in(0,1-p)$, has a maximal non-negative derivative at $x=0$.

Lemma 9. *Consider function $f(x)=(1+r\cdot x)\cdot(\sum_{i=0}^{n-2}(1-p-x)^i)$, where $p\in(0,1)$, $x\in(0,1-p)$ and $r\in(0,\frac{1}{p})$. If there exists $x'\in(0,1-p)$ such that $\frac{df}{dx}(x')\geq 0$, then $\max_x\frac{df}{dx}(x)=\frac{df}{dx}(0)$.*

Proof. We can rewrite function $f(x)$ as:

$$f(x)=(1+r\cdot x)\cdot(\sum_{i=0}^{n-2}(1-p-x)^i)=(1+r-r\cdot p-r\cdot(1-p-x))\cdot(\sum_{i=0}^{n-2}(1-p-x)^i)$$
$$=(1+r-r\cdot p)\cdot\sum_{i=0}^{n-2}(1-p-x)^i-r\cdot\sum_{i=1}^{n-1}(1-p-x)^i$$
$$=r\cdot(1-(1-p-x)^{n-1})+(1-r\cdot p)\cdot\sum_{i=0}^{n-2}(1-p-x)^i$$

The derivative of $f(x)$ is equal to:

$$\frac{df}{dx}(x)=(n-1)\cdot r\cdot(1-p-x)^{n-2}-(1-r\cdot p)\cdot\sum_{i=1}^{n-2}i\cdot(1-p-x)^{i-1}$$
$$=(1-p-x)^{n-2}\cdot\left((n-1)\cdot r-(1-r\cdot p)\cdot\sum_{i=1}^{n-2}\frac{i}{(1-p-x)^{n-1-i}}\right)$$

Now, suppose there is $x'\in(0,1-p)$ such that $\frac{df}{dx}(x')\geq 0$. Since $x\in(0,1-p)$, the necessary condition for that is:

$$(n-1)\cdot r-(1-r\cdot p)\cdot\sum_{i=1}^{n-2}\frac{i}{(1-p-x')^{n-1-i}}\geq 0$$

Therefore, we obtain:

$$\begin{aligned}
\frac{df}{dx}(x') &= (1-p-x')^{n-2}\cdot\left((n-1)\cdot r-(1-r\cdot p)\cdot\sum_{i=1}^{n-2}\frac{i}{(1-p-x')^{n-1-i}}\right)\\
&\leq (1-p-x')^{n-2}\cdot\left((n-1)\cdot r-(1-r\cdot p)\sum_{i=1}^{n-2}\frac{i}{(1-p)^{n-1-i}}\right)\\
&\leq (1-p)^{n-2}\cdot\left((n-1)\cdot r-(1-r\cdot p)\cdot\sum_{i=1}^{n-2}\frac{i}{(1-p)^{n-1-i}}\right)=\frac{df}{dx}(0)
\end{aligned}$$

□

A.3 Dasgupta&Ghosh mechanism

We have seen that RPTS reduces to a simple score when statistic $\mathbf{x}_a$ is calculated based on only one phenomenon in addition to the phenomenon being observed by agent a. The form of score (4.14) is similar to the Dasgupta&Ghosh mechanism introduced in [DG13]. In fact, they are equivalent.

To see this, we first need to describe the basic structure of the Dasgupta&Ghosh mechanism. Let us assume for simplicity that an agent a and her peer p_i have only one common phenomena Φ_i that they observe (see [DG13] for how to transform the mechanism when this does not hold), and that they both observe m additional phenomena. Notice that now we have a larger batch of phenomena and each agent observes multiple of them. By carefully rearranging terms in the Dasgupta&Ghosh mechanism, we obtain that the mechanism is equivalent to:

$$\underbrace{\mathbb{1}_{Y_{a,i}=Y_{p_i,i}}}_{\textit{agreement A}}-\underbrace{\sum_{z\in\mathscr{X}}\sum_{\Phi_j\neq\Phi_i}\frac{\mathbb{1}_{Y_{a,j}=z}}{m}\sum_{\Phi_k\neq\Phi_i}\frac{\mathbb{1}_{Y_{p_i,k}=z}}{m}}_{\textit{statistic B}}=\underbrace{\mathbb{1}_{Y_{a,i}=Y_{p_i,i}}}_{\textit{agreement A}}-\underbrace{\sum_{\Phi_j\neq\Phi_i}\sum_{\Phi_k\neq\Phi_i}\frac{\mathbb{1}_{Y_{a,j}=Y_{p_i,k}}}{m^2}}_{\textit{statistic B}}$$

where $Y_{a,j}$ is agent a's report for phenomenon Φ_j, and summation Σ_{Φ_j} is over the phenomena observed by agent a (equivalent notation is used for agent p_i). Therefore, the total score is equal to:

$$\sum_{\Phi_i}\left[\mathbb{1}_{Y_{a,i}=Y_{p_i,i}}-\sum_{\Phi_j\neq\Phi_i}\sum_{\Phi_k\neq\Phi_i}\frac{\mathbb{1}_{Y_{a,j}=Y_{p_i,k}}}{m^2}\right]=\sum_{\Phi_i}\left[\mathbb{1}_{Y_{a,i}=Y_{p_i,i}}-\frac{1}{m}\sum_{l=1}^{m}\sum_{\Phi_l\in\{\Phi^l_{-a,\Phi_i}\}}\frac{\mathbb{1}_{Y_{a,i}=Y_{p_l,l}}}{m}\right]$$

where $\{\Phi^l_{-a,\Phi_i}\}$ is a group of m phenomena not observed by agent a and they are obtained from the previous step by rearranging the terms in the equation. This can be thought of as if the Dasgupta&Ghosh mechanism scores a report $Y_{a,i}$ with:

$$\underbrace{\mathbb{1}_{Y_{a,i}=Y_{p_i,i}}}_{agreement\ A} - \underbrace{\frac{1}{m}\sum_{l=1}^{m}\sum_{\Phi_l\in\{\Phi^l_{-a,\Phi_i}\}}\frac{\mathbb{1}_{Y_{a,i}=Y_{p_l,l}}}{m}}_{statistic\ B}$$

Now, since $Y_{a,i}$ and $Y_{p_l,l}$ are statistically independent (because agent a has not observed Φ_l), part B of the score is in the expectation equivalent to $\mathbf{x}'_a(Y_{a,i})$ (see Section 4.4.2). Therefore, the mechanism defined by (4.14) and the Dasgupta&Ghosh mechanism are equivalent, which means that the Dasgupta&Ghosh mechanism is a special case of RPTS obtained in the limit case when $\mathbf{x}_a$ is calculated from only two phenomena. Moreover, the equivalence implies that the Dasgupta&Ghosh mechanism requires non-correlated (categorical) observation values for the honest reporting strategy profile to result in a maximum payoff.

A.4 Detailed simulation results for RPTS and log-PTS

This section provides the detailed simulation results of the RPTS and log-PTS mechanisms in the community sensing simulation setup of Section 4.5.2 (Chapter 4), for the decreasing population of sensors.

Table A.1 – Average payoffs — RPTS (100 sensors, 13 peers)

Strategy	mean	min	max	median	1st quartile	3rd quartile
honest	0.138	-0.107	0.584	0.095	0.037	0.175
collude	0.051	-0.025	0.252	0.031	0.007	0.06
colludeLow	0	0	0	0	0	0
colludeExtraHigh	0	0	0	0	0	0
random	-0.507	-0.658	-0.185	-0.517	-0.579	-0.459
randomAll	-0.002	-0.051	0.05	0.0	-0.016	0.01

Table A.2 – Average payoffs — RPTS (80 sensors, 11 peers)

Strategy	mean	min	max	median	1st quartile	3rd quartile
honest	0.126	-0.126	0.569	0.087	0.035	0.168
collude	0.044	-0.023	0.216	0.029	0.011	0.057
colludeLow	0	0	0	0	0	0
colludeExtraHigh	0	0	0	0	0	0
random	-0.513	-0.669	-0.276	-0.529	-0.577	-0.469
randomAll	0.0	-0.065	0.066	0.0	-0.018	0.017

Table A.3 – Average payoffs — RPTS (60 sensors, 9 peers)

Strategy	mean	min	max	median	1st quartile	3rd quartile
honest	0.126	-0.098	0.589	0.089	0.032	0.17
collude	0.043	-0.042	0.206	0.029	0.008	0.062
colludeLow	0	0	0	0	0	0
colludeExtraHigh	0	0	0	0	0	0
random	-0.52	-0.664	-0.228	-0.523	-0.576	-0.469
randomAll	0.0	-0.08	0.069	-0.004	-0.018	0.018

Table A.4 – Average payoffs — RPTS (40 sensors, 7 peers)

Strategy	mean	min	max	median	1st quartile	3rd quartile
honest	0.098	-0.112	0.522	0.068	0.013	0.161
collude	0.035	-0.031	0.239	0.02	0.0	0.052
colludeLow	0	0	0	0	0	0
colludeExtraHigh	0	0	0	0	0	0
random	-0.535	-0.697	-0.308	-0.536	-0.586	-0.493
randomAll	0.0	-0.018	0.117	-0.003	-0.027	0.033

Table A.5 – Average payoffs — log-PTS (100 sensors, 13 peers)

Strategy	mean	min	max	median	1st quartile	3rd quartile
honest	0.025	-1.281	0.302	0.042	-0.02	0.104
collude	0.01	-0.352	0.109	0.015	-0.01	0.034
colludeLow	0	0	0	0	0	0
colludeExtraHigh	0	0	0	0	0	0
random	-0.847	-1.718	-0.452	-0.808	-0.997	-0.658
randomAll	-0.329	-0.5	-0.178	-0.332	-0.38	-0.279

Table A.6 – Average payoffs — log-PTS (80 sensors, 11 peers)

Strategy	mean	min	max	median	1st quartile	3rd quartile
honest	0.014	-1.178	0.3	0.039	-0.03	0.09
collude	0.007	-0.283	0.12	0.012	-0.01	0.027
colludeLow	0	0	0	0	0	0
colludeExtraHigh	0	0	0	0	0	0
random	-0.874	-1.425	-0.366	-0.858	-1.038	-0.694
randomAll	-0.479	-0.701	-0.232	-0.473	-0.55	-0.409

Table A.7 – Average payoffs — log-PTS (60 sensors, 9 peers)

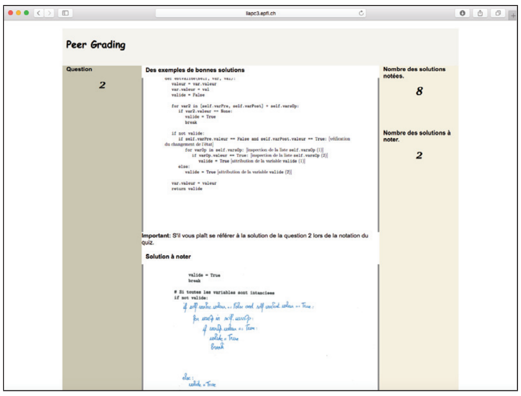
Strategy	mean	min	max	median	1st quartile	3rd quartile
honest	-0.021	-1.389	0.277	0.011	-0.06	0.078
collude	-0.004	-0.254	0.09	0.003	-0.015	0.023
colludeLow	0	0	0	0	0	0
colludeExtraHigh	0	0	0	0	0	0
random	-0.887	-1.524	-0.465	-0.836	-1.026	-0.715
randomAll	-0.759	-1.111	-0.406	-0.745	-0.846	-0.659

Table A.8 – Average payoffs — log-PTS (40 sensors, 7 peers)

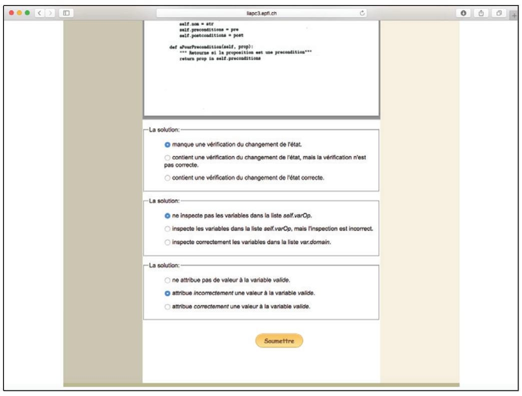
Strategy	mean	min	max	median	1st quartile	3rd quartile
honest	-0.069	-1.862	0.269	-0.006	-0.1	0.046
collude	-0.016	-0.345	0.123	-0.0003	-0.025	0.017
colludeLow	0	0	0	0	0	0
colludeExtraHigh	0	0	0	0	0	0
random	-0.994	-1.607	-0.563	-0.992	-1.139	-0.823
randomAll	-1.224	-1.767	-0.773	-1.195	-1.364	-1.077

A.5 Web interfaces for the peer grading task in Section 4.5.1

This section provides figures of the web interfaces for the second question of the peer grading assignment and for the tutorials about the constant reward mechanism and the RPTS mechanism.

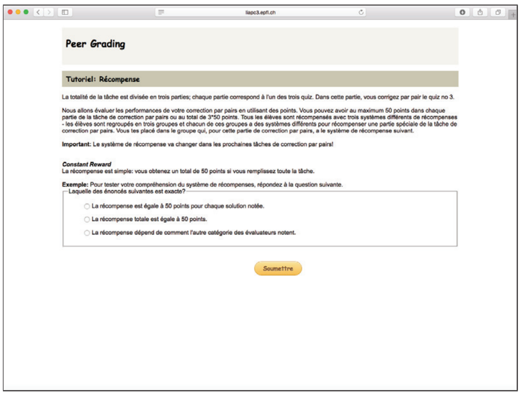


(a) The correct solution to the quiz question and a student's solution

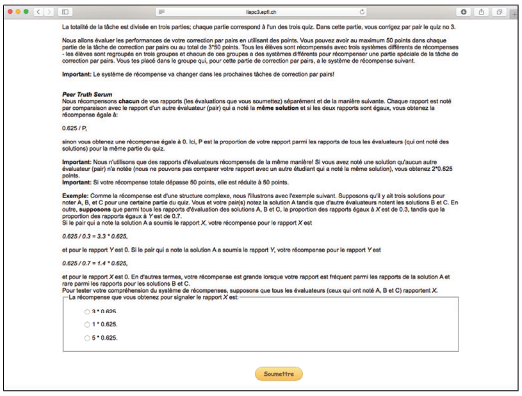


(b) Input form for corrections of the student's solution

Figure A.3 – Peer grading task - the second question



(a) Explanation of the constant reward mechanism



(b) Explanation of the RPTS mechanism

Figure A.4 – Tutorial about the reward mechanisms

Bibliography

[ACBF02] Peter Auer, Nicolò Cesa-Bianchi, and Paul Fischer. Finite-time analysis of the multiarmed bandit problem. *Machine Learning*, 47(2-3):235–256, 2002.

[Arr70] Kenneth Arrow. *Social choice and individual values.* Social choice and individual values, 1970.

[ASC+10] Karl Aberer, Saket Sathe, Dipanjan Chakraborty, Alcherio Martinoli, Guillermo Barrenetxea, Boi Faltings, and Lothar Thiele. Opensense: Open community driven sensing of environment. In *ACM SIGSPATIAL International Workshop on GeoStreaming*, 2010.

[ASP13] ASPA. l'association pour la surveillance et l'etude de la pollution atmosphérique en alsace. www.atmo-alsace.net, 2013.

[BB03] Sonja Buchegger and Jean-Yves Le Boudec. A robust reputation system for mobile ad-hoc networks. Technical report, Proceedings of P2PEcon, 2003.

[BEH+06] Jeff Burke, Deborah Estrin, Mark Hansen, Andrew Parker, Nithya Ramanathan, Sasank Reddy, and Mani B. Srivastava. Participatory sensing. In *Workshop on World-Sensor-Web: Mobile Device Centric Sensor Networks and Applications*, 2006.

[Bis06] Christopher M. Bishop. *Pattern Recognition and Machine Learning.* Springer, 2006.

[BLB02] Sonja Buchegger and Jean-Yves Le Boudec. Performance analysis of the CONFIDANT protocol. In *Proceedings of the 3rd ACM International Symposium on Mobile Ad Hoc Networking and Computing*, 2002.

[Bri50] Glenn W. Brier. Verification of forecasts expressed in terms of probability. *Monthly weather review*, 78(13), 1950.

[CBL06] Nicolo Cesa-Bianchi and Gabor Lugosi. *Prediction, Learning and Games.* Cambridge University Press, 2006.

[CBMS07] Nicolò Cesa-Bianchi, Yishay Mansour, and Gilles Stoltz. Improved second-order bounds for prediction with expert advice. *Mach. Learn.*, 66(2-3):321–352, 2007.

[Che09] Haiguang Chen. Task-based trust management for wireless sensor networks. *International Journal of Security and Its Applications*, 3(2):21–26, 2009.

[CM85] Jacques Cremer and Richard P. McLean. Optimal selling strategies under uncertainty for a discriminating monopolist when demands are interdependent. *Econometrica*, 53:345–361, 1985.

[CP07] Yiling Chen and David M. Pennock. A utility framework for bounded-loss market makers. In *Proceedings of the Twenty-Third Conference on Uncertainty in Artificial Intelligence*, 2007.

[CRKH11] Delphine Christin, Andreas Reinhardt, Salil S. Kanhere, and Matthias Hollick. A survey on privacy in mobile participatory sensing applications. *The Journal of Systems and Software*, 84(11):1928–1946, 2011.

[CV10] Yiling Chen and Jennifer Wortman Vaughan. A new understanding of prediction markets via no-regret learning. In *Proceedings of the 11th ACM Conference on Electronic Commerce (EC '10)*, 2010.

[CWC+02] Roy N. Colvile, Nicky K. Woodfield, David J. Carruthers, Bernard E. A. Fisher, Anna Rickard, S. Neville, and A. Hughes. Uncertainty in dispersion modelling and urban air quality mapping. *Environmental Science and Policy*, 5(3):207–220, 2002.

[dASP16] Luca de Alfaro, Michael Shavlovsky, and Vassilis Polychronopoulos. Incentives for truthful peer grading. *CoRR*, abs/1604.03178, 2016.

[DBFH09] Akshay Dua, Nirupama Bulusu, Wu-Chang Feng, and Wen Hu. Towards trustworthy participatory sensing. In *Proceedings of the 4th USENIX conference on Hot topics in security*, 2009.

[Del05] Chrysanthos Dellarocas. Reputation mechanism design in online trading environments with pure moral hazard. *Info. Sys. Research*, 16(2), 2005.

[DG13] Anirban Dasgupta and Arpita Ghosh. Crowdsourced judgement elicitation with endogenous proficiency. In *Proceedings of the 22nd ACM International World Wide Web Conference (WWW'13)*, 2013.

[EG15] David Easley and Arpita Ghosh. Behavioral mechanism design: Optimal crowdsourcing contracts and prospect theory. In *Proceedings of the Sixteenth ACM Conference on Economics and Computation (EC'15)*, 2015.

[FCK15] Rafael Frongillo, Yiling Chen, and Ian Kash. Elicitation for aggregation. In *Proceedings of the 29th AAAI Conference on Artificial Intelligence (AAAI'15)*, 2015.

[FLJ14] Boi Faltings, Jason J. Li, and Radu Jurca. Incentive mechanisms for community sensing. *IEEE Transactions on Computers*, 63:115–128, 2014.

[FPTJ14] Boi Faltings, Pearl Pu, Bao Duy Tran, and Radu Jurca. Incentives to counter bias in human computation. In *Proceedings of the 2nd AAAI Conference on Human Computation and Crowdsourcing (HCOMP'14)*, 2014.

[FW16] Rafael Frongillo and Jens Witkowski. A geometric method to construct minimal peer prediction mechanisms. In *Proceedings of the 30th AAAI Conference on Artificial Intelligence (AAAI'16)*, 2016.

[GC11] Aurélien Garivier and Olivier Cappé. The KL-UCB algorithm for bounded stochastic bandits and beyond. In *The 24th Annual Conference on Learning Theory (COLT'11)*, pages 359–376, 2011.

[GF14] Florent Garcin and Boi Faltings. Swissnoise: Online polls with game-theoretic incentives. In *Proceedings of the 26th Conference on Innovative Applications of AI*, 2014.

[Gho13] Arpita Ghosh. Game theory and incentives in human computation systems. In *Handbook of Human Computation*. Springer, 2013.

[GJL+11] Peter Gilbert, Jaeyeon Jung, Kyungmin Lee, Henry Qin, Daniel Sharkey, Anmol Sheth, and Landon P. Cox. Youprove: Authenticity and fidelity in mobile sensing. In *Proceedings of the 9th ACM Conference on Embedded Networked Sensor Systems*, 2011.

[GMCA14] Xi Alice Gao, Andrew Mao, Yiling Chen, and Ryan P. Adams. Trick or treat: Putting peer prediction to the test. In *Proceedings of the 15th ACM Conference on Economics and Computation (EC'14)*, 2014.

[Goo52] Irving J. Good. Rational decisions. *Journal of the Royal Statistical Society*, 14:107–114, 1952.

[GR07] Tilmann Gneiting and Adrian E. Raftery. Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association*, 102:359–378, 2007.

[GR14] Arpita Ghosh and Kleinberg Robert. Optimal contest design for simple agents. In *Proceedings of the Fifteenth ACM Conference on Economics and Computation (EC '14)*, 2014.

[GRP09] Sharad Goel, Daniel M. Reeves, and David M. Pennock. Collective revelation: A mechanism for self-verified, weighted, and truthful predictions. In *Proceedings of the 10th ACM conference on Electronic commerce (EC'09)*, 2009.

[GS04] Saurabh Ganeriwal and Mani B. Srivastava. Reputation-based framework for high integrity sensor networks. In *Proceedings of the 2nd ACM workshop on Security of ad hoc and sensor networks*, 2004.

[GWL16] Alice Gao, James R. Wright, and Kevin Leyton-Brown. Incentivizing evaluation via limited access to ground truth: Peer-prediction makes things worse. *CoRR*, abs/1606.07042, 2016.

[Han03] Robin D. Hanson. Combinatorial information market design. *Information Systems Frontiers*, 5(1):107–119, 2003.

[Han07] Robin D. Hanson. Logarithmic market scoring rules for modular combinatorial information aggregation. *Journal of Prediction Markets*, 1(1):1–15, 2007.

[Har11] Christopher Harris. You're hired!: An examination of crowdsourcing incentive models in human resource tasks. In *Proceedings of the Workshop on Crowdsourcing for Search and Data Mining (CSDM) at the Fourth ACM International Conference on Web Search and Data Mining (WSDM)*, 2011.

[HF13a] Shih-Wen Huang and Wai-Tat Fu. Don't hide in the crowd!: Increasing social transparency between peer workers improves crowdsourcing outcomes. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2013.

[HF13b] Shih-Wen Huang and Wai-Tat Fu. Enhancing reliability using peer consistency evaluation in human computation. In *Proceedings of the 2013 conference on Computer supported cooperative work*, 2013.

[HKH14] Kuan Lun Huang, Salil S. Kanhere, and Wen Hu. On the need for a reputation system in mobile phone based sensing. *Ad Hoc Networks*, 12:130–149, 2014.

[HPZ06] Nan Hu, Paul A. Pavlou, and Jennifer Zhang. Can online reviews reveal a product's true quality?: Empirical findings and analytical modeling of online word-of-mouth communication. In *Proceedings of the 7th ACM Conference on Electronic Commerce (EC '06)*, 2006.

[HSW+14] David Hasenfratz, Olga Saukh, Christoph Walser, Christoph Hueglin, Martin Fierz, and Lothar Thiele. Pushing the spatio-temporal resolution limit of urban air pollution maps. In *Proceedings of the 12th International Conference on Pervasive Computing and Communications (PerCom 2014)*, 2014.

[Iha93] Shunsuke Ihara. *Information Theory for Continuous Systems*. World Scientific Publisher Co. Pte. Ltd., 1993.

[JF06a] Radu Jurca and Boi Faltings. Minimum payments that reward honest reputation feedback. In *Proceedings of the 7th ACM Conference on Electronic Commerce (EC'06)*, 2006.

[JF06b] Radu Jurca and Boi Faltings. Robust incentive-compatible feedback payments. In *Agent-Mediated Electronic Commerce*, pages 204–218, 2006.

[JF08] Radu Jurca and Boi Faltings. Incentives for expressing opinions in online polls. In *Proceedings of the 9th ACM conference on Electronic commerce (EC' 08)*, 2008.

[JF09] Radu Jurca and Boi Faltings. Mechanisms for making crowds truthful. *Journal of Artificial Intelligence Research (JAIR)*, 34:209–253, 2009.

[JF11] Radu Jurca and Boi Faltings. Incentives for answering hypothetical questions. In *Workshop on Social Computing and User Generated Content (EC' 11)*, 2011.

[JI02] Audun Josang and Roslan Ismail. The beta reputation system. In *Proceedings of the Electronic Commerce Conference*, 2002.

[JLF14] Arnaud Jutzeler, Jason Jingshi Li, and Boi Faltings. A region-based model for estimating urban air pollution. In *Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence*, AAAI'14, 2014.

[JLP12] Leslie K. John, George Loewenstein, and Drazen Prelec. Measuring the prevalence of questionable research practices with incentives for truth telling. *Psychological Science*, 23:524–532, 2012.

[JSV14] Srikanth Jagabathula, Lakshminarayanan Subramanian, and Ashwin Venkataraman. Reputation-based worker filtering in crowdsourcing. In *Proceeedings of the 27th Annual Conference on Neural Information ProcessingSystems (NIPS'14)*, 2014.

[KH12] Ece Kamar and Eric Horvitz. Incentives for truthful reporting in crowdsourcing. In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems (AAMAS'12)*, 2012.

[KHKZ08] Andreas Krause, Eric Horvitz, Aman Kansal, and Feng Zhao. Toward community sensing. In *Proceedings of the 7th International Conference on Information Processing in Sensor Networks (IPSN'08)*, 2008.

[KOS11] David R. Karger, Sewoong Oh, and Devavrat Shah. Iterative learning for reliable crowdsourcing systems. In *Proceeedings of the 24th Annual Conference on Neural Information ProcessingSystems (NIPS'11)*, 2011.

[KOS13] David R. Karger, Sewoong Oh, and Devavrat Shah. Efficient crowdsourcing for multi-class labeling. In *Proceedings of the ACM SIGMETRICS/International Conference on Measurement and Modeling of Computer Systems (SIGMETRICS '13)*, 2013.

[KS10] Sindhu Kutty and Rahul Sami. A prediction market approach to learning with sequential advice. In *NIPS Workshop on Computational Social Science and the Wisdom of Crowds*, 2010.

[KS11] Sindhu Kutty and Rahul Sami. Bounded regret sequential learning using prediction markets. In *NIPS Workshop on Relations between Machine Learning Problems*, 2011.

[KS16a] Yuqing Kong and Grant Schoenebeck. Equilibrium selection in information elicitation without verification via information monotonicity. *CoRR*, 2016.

[KS16b] Yuqing Kong and Grant Schoenebeck. A framework for designing information elicitation mechanisms that reward truth-telling. *CoRR*, 2016.

[KSG08] Andreas Krause, Ajit Singh, and Carlos Guestrin. Near-optimal sensor placements in gaussian processes: Theory, efficient algorithms and empirical studies. *The Journal of Machine Learning Research*, 9:235–284, 2008.

[KSL16] Yuqing Kong, Grant Schoenebeck, and Katrina Ligett. Putting peer prediction under the micro(economic)scope and making truth-telling focal. *CoRR*, 2016.

[KSM$^+$15] Vijay Kamble, Nihar Shah, David Marn, Abhay Parekh, and Kannan Ramachandran. Truth serums for massively crowdsourced evaluation tasks. In *Proceedings of the 5th Workshop on Social Computing and User Generated Content*, 2015.

[KSMP14] Renato Kempter, Valentina Sintsova, Claudiu Cristian Musat, and Pearl Pu. Emotionwatch: Visualizing fine-grained emotions in event-related tweets. In *Proceedings of the Eighth International Conference on Weblogs and Social Media (ICWSM'14)*, 2014.

[KWL$^+$13] Chinmay Kulkarni, Koh Pang Wei, Huy Le, Daniel Chia, Kathryn Papadopoulos, Justin Cheng, Daphne Koller, and Scott R. Klemmer. Peer and self assessment in massive online classes. *ACM Trans. Comput.-Hum. Interact.*, 20(6), 2013.

[LPI12] Qiang Liu, Jian Peng, and Alex Ihler. Variational inference for crowdsourcing. In *Proceeedings of the 25th Annual Conference on Neural Information Processing Systems (NIPS'12)*, 2012.

[LPS08] Nicolas S. Lambert, David M. Pennock, and Yoav Shoham. Eliciting properties of probability distributions. In *Proceedings of the 9th ACM Conference on Electronic Commerce (EC '08)*, 2008.

[LR85] T. L. Lai and Herbert Robbins. Asymptotically efficient adaptive allocation rules. *Advances in Applied Mathematics*, 6(1):4–22, 1985.

[LRSH11] Jan Lorenz, Heiko Rauhut, Frank Schweitzer, and Dirk Helbing. How social influence can undermine the wisdom of crowd effect. *Proceedings of the National Academy of Sciences*, 108:9020–9025, 2011.

[LS08] Nicolas Lambert and Yoav Shoham. Truthful surveys. In *Proceedings of the 3rd International Workshop on Internet and Network Economics (WINE 2008)*, 2008.

[LS09] Nicolas Lambert and Yoav Shoham. Eliciting truthful answers to multiple-choice questions. In *Proceedings of the tenth ACM conference on Electronic Commerce*, 2009.

[LW94] Nick Littlestone and Manfred K. Warmuth. The weighted majority algorithm. *Inf. Comput.*, 108(2):212–261, 1994.

[MM02] Pietro Michiardi and Refik Molva. CORE: A collaborative reputation mechanism to enforce node cooperation in mobile ad hoc networks. In *Proceedings of the Sixth Joint Working Conference on Communications and Multimedia Security*, 2002.

[MRZ05] Nolan Miller, Paul Resnick, and Richard Zeckhauser. Eliciting informative feedback: The peer-prediction method. *Management Science*, 51:1359–1373, 2005.

[NRTV07] Noam Nisan, Tim Roughgarden, Eva Tardos, and Vijay V. Vazirani. *Algorithmic Game Theory*. Cambridge University Press, 2007.

[OSL+11] David Oleson, Alexander Sorokin, Greg Laughlin, Vaughn Hester, John Le, and Lukas Biewald. Programmatic gold: Targeted and scalable quality assurance in crowdsourcing. In *Proceedings of the AAAI Workshop on Human Computation*, 2011.

[PHC+13] Chris Piech, Jonathan Huang, Zhenghao Chen, Chuong B. Do, Andrew Y. Ng, and Daphne Koller. Tuned models of peer assessment in MOOCs. *CoRR*, abs/1307.2579, 2013.

[PLGN01] David M. Pennock, Steve Lawrence, Lee Giles, and Finn Nielsen. The real power of artificial markets. *Science*, 291(5506):987–988, 2001.

[Pre04] Drazen Prelec. A Bayesian truth serum for subjective data. *Science*, 34(5695):462–466, 2004.

[PS06] Drazen Prelec and Sebastian Seung. An algorithm that finds truth even if most people are wrong. Working paper, 2006.

[RF13] Goran Radanovic and Boi Faltings. A robust Bayesian truth serum for non-binary signals. In *Proceedings of the 27th AAAI Conference on Artificial Intelligence (AAAI'13)*, 2013.

[RF14] Goran Radanovic and Boi Faltings. Incentives for truthful information elicitation of continuous signals. In *Proceedings of the 28th AAAI Conference on Artificial Intelligence (AAAI'14)*, 2014.

[RF15a] Goran Radanovic and Boi Faltings. Incentive schemes for participatory sensing. In *Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems (AAMAS'15)*, 2015.

[RF15b] Goran Radanovic and Boi Faltings. Incentives for subjective evaluations with private beliefs. In *Proceedings of the 29th AAAI Conference on Artificial Intelligence (AAAI'15)*, 2015.

[RF16a] Goran Radanovic and Boi Faltings. Incentives for effort in crowdsourcing using the peer truth serum. *ACM Transactions on Intelligent Systems and Technology*, 7(4):48:1–48:28, 2016.

[RF16b] Goran Radanovic and Boi Faltings. Learning to scale payments in crowdsourcing with properboost. In *Proceedings of the 4th AAAI Conference on Human Computation and Crowdsourcing (HCOMP'16) (to appear)*, 2016.

[RF16c] Goran Radanovic and Boi Faltings. Limiting the influence of low quality information in community sensing. In *Proceedings of the 2016 International Conference on Autonomous Agents and Multiagent Systems (AAMAS'16)*, 2016.

[Ril14] Blake Riley. Minimum truth serums with optional predictions. In *Proceedings of the 4th Workshop on Social Computing and User Generated Content*, 2014.

[RRCN09] Steven Reece, Stephen J. Roberts, Christopher Claxton, and David Nicholson. Multi-sensor fault recovery in the presence of known and unknown fault types. In *12th International Conference on Information Fusion*, 2009.

[RS07] Paul Resnick and Rahul Sami. The influence limiter: Provably manipulation-resistant recommender system. In *Proceedings of the 2007 ACM Conference on Recommender Systems*, 2007.

[RW05] Carl Edward Rasmussen and Christopher K. I. Williams. *Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning)*. The MIT Press, 2005.

[RYZ+10] Vikas C. Raykar, Shipeng Yu, Linda H. Zhao, Gerardo Hermosillo Valadez, Charles Florin, Luca Bogoni, and Linda Moy. Learning from crowds. *J. Mach. Learn. Res.*, 11:1297–1322, 2010.

[SAFP16] Victor Shnayder, Arpit Agarwal, Rafael Frongillo, and David C. Parkes. Informed truthfulness in multi-task peer prediction. In *Proceedings of the 2016 ACM Conference on Economics and Computation (EC '16)*, 2016.

[Sav71] Leonard J. Savage. Elicitation of personal probabilities and expectations. *Journal of the American Statistical Association*, 66(336):783–801, 1971.

[Sch05] Klaus Scherer. What are emotions? And how can they be measured? *Social science information*, 44:695–729, 2005.

[SCH11] Aaron Shaw, Daniel L. Chen, and John Horton. Designing incentives for inexpert human raters. In *Proceedings of the ACM 2011 Conference on Computer Supported Cooperative Work (CSCW'11)*, 2011.

[SK13] Adish Singla and Andreas Krause. Incentives for privacy tradeoff in community sensing. In *Proceedings of the 1st AAAI Conference on Human Computation and Crowdsourcing (HCOMP'13)*, 2013.

[SLB08] Yoav Shoham and Kevin Leyton-Brown. *Multiagent Systems: Algorithmic, Game-Theoretic, and Logical Foundations*. Cambridge University Press, 2008.

[SP16a] Victor Shnayder and David C. Parkes. Measuring performance of peer prediction mechanisms using replicator dynamics. In *Proceedings of the 25th International Joint Conference on Artificial Intelligence (IJCAI'16)*, 2016.

[SP16b] Victor Shnayder and David C. Parkes. Practical peer prediction for peer assessment. In *Proceedings of the 4th AAAI Conference on Human Computation and Crowdsourcing (HCOMP'16) (to appear)*, 2016.

[Sur05] James Surowiecki. *The Wisdom of Crowds*. Anchor, 2005.

[SW10] Stefan Saroiu and Alec Wolman. I am a sensor, and i approve this message. In *Proceedings of the Eleventh Workshop on Mobile Computing Systems and Applications*, 2010.

[SZ15] Nihar B. Shah and Dengyong Zhou. Double or nothing: Multiplicative incentive mechanisms for crowdsourcing. In *Advances in Neural Information Processing Systems (NIPS'15)*, 2015.

[SZP15] Nihar Shah, Dengyong Zhou, and Yuval Peres. Approval voting and incentives in crowdsourcing. In *Proceedings of the 32nd International Conference on Machine Learning (ICML'15)*, 2015.

[UDG15] Johan Ugander, Ryan Drapeau, and Carlos Guestrin. The wisdom of multiple guesses. In *Proceedings of the Sixteenth ACM Conference on Economics and Computation (EC '15)*, 2015.

[vAD04] Luis von Ahn and Laura Dabbish. Labeling images with a computer game. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, 2004.

[vAD08] Luis von Ahn and Laura Dabbish. Designing games with a purpose. *Commun. ACM*, 51(8):58–67, 2008.

[VRJ13] Matteo Venanzi, Alex Rogers, and Nicholas R. Jennings. Trust-based fusion of untrustworthy information in crowdsourcing applications. In *Proceedings of the 2013 International Conference on Autonomous Agents and Multi-agent Systems (AAMAS'13)*, 2013.

[VVV13] Aditya Vempaty, Lav R. Varshney, and Pramod K. Varshney. Reliable crowdsourcing for multi-class labeling using coding theory. *CoRR*, 2013.

[WC13] Bo Waggoner and Yiling Chen. Information elicitation sans verification. In *Information Elicitation Sans Verification. In Proceedings of the 3rd Workshop on Social Computing and User Generated Content*, 2013.

[WC14] Bo Waggoner and Yiling Chen. Output agreement mechanisms and common knowledge. In *Proceedings of the 2nd AAAI Conference on Human Computation and Crowdsourcing (HCOMP'14)*, 2014.

[WDK+15] William Wu, Constantinos Daskalakis, Nicolaas Kaashoek, Christos Tzamos, and Matthew Weinberg. Game theory based peer grading mechanisms for MOOCs. In *Proceedings of the Second (2015) ACM Conference on Learning @ Scale (L@S '15)*, 2015.

[Wit14] Jens Witkowski. *Robust Peer Prediction Mechanisms*. PhD thesis, Albert-Ludwigs-Universitat Freiburg: Institut fur Informatik, 2014.

[WLSH09] Wenkai Wang, Husheng Li, Yan Sun, and Zhu Han. Catchit: Detect malicious nodes in collaborative spectrum sensing. In *Proceedings of the 28th IEEE Conference on Global Telecommunications*, 2009.

[WP12a] Jens Witkowski and David C. Parkes. Peer prediction without a common prior. In *Proceedings of the 13th ACM Conference on Electronic Commerce (EC' 12)*, pages 964–981, 2012.

[WP12b] Jens Witkowski and David C. Parkes. A robust Bayesian truth serum for small populations. In *Proceedings of the 26th AAAI Conference on Artificial Intelligence (AAAI'12)*, 2012.

[WP13a] Ray Weaver and Drazen Prelec. Creating truth-telling incentives with the Bayesian truth serum. *Journal of Marketing Research*, 50:289–302, 2013.

[WP13b] Jens Witkowski and David C. Parkes. Learning the prior in minimal peer prediction. In *Proceedings of the 3rd Workshop on Social Computing and User Generated Content (SC 2013)*, 2013.

[WTL15] James R. Wright, Chris Thornton, and Kevin Leyton-Brown. Mechanical TA: partially automated high-stakes peer grading. In *Proceedings of the 46th ACM Technical Symposium on Computer Science Education*, 2015.

[XS15] Yuanzhang Xiao and Mihaela Van Der Schaar. Socially-optimal design of service exchange platforms with imperfect monitoring. *ACM Trans. Econ. Comput.*, 3(4):25:1–25:25, 2015.

[YS10] Kuldeep Yadav and Avinash Srinivasan. iTrust: An integrated trust framework for wireless sensor networks. In *Proceedings of the 2010 ACM Symposium on Applied Computing*, 2010.

[ZC14] Peter Zhang and Yiling Chen. Elicitability and knowledge-free elicitation with peer prediction. In *Proceedings of the 2014 international conference on Autonomous agents and multi-agent systems (AAMAS '14)*, 2014.

[ZVdS12] Yu Zhang and Mihaela Van der Schaar. Reputation-based incentive protocols in crowdsourcing applications. In *IEEE INFOCOM*, 2012.

Goran Radanović

(as of September 2016)

EPFL - Artificial Intelligence Laboratory
Batiment IN, Station 14
1015 Lausanne, Switzerland
Email: goran.radanovic@epfl.ch
Phone: +41 21 69 36682

EDUCATION

- September 2011 – present: PhD degree in computer science, *École polytechnique fédérale de Lausanne* (EPFL)
    - Thesis supervisor: Prof. Boi Faltings.

- October 2008 – July 2010: Master's degree in computer science, *Faculty of Electrical Engineering and Computing at University of Zagreb*
    - GPA 5.0/5.0
    - graduated with the highest honor (summa cum laude).
    - Thesis supervisor: Prof. Marin Golub

- October 2005 – July 2008: Bachelor degree in computer science, *Faculty of Electrical Engineering and Computing at University of Zagreb*,
    - GPA 4.966/5.0.
    - Thesis supervisor: Prof. Marin Golub

WORK EXPERIENCE

- January 2013 – March 2016: system administrator for the Artificial Intelligence Laboratory, EPFL

- April 2011 – July 2011: temporary employee in Kron d.o.o. company
    - software development for a mobile payment architecture

- October 2010 - April 2011: independent contractor for Kron d.o.o. company
    - software development for a mobile payment architecture

- November 2010 – February 2011: volunteer assistant on an image analysis project in astronomy

- Supervisors:
  - Prof. Sven Loncaric, *Faculty of Electrical Engineering and Computing at University of Zagreb*
  - Prof. Dejan Vinkovic, *Physics Department at University of Split*

- September 2007 – February 2008: undergraduate teaching assistant in Mathematics 3-C, *Faculty of Electrical Engineering and Computing at University of Zagreb*

ACADEMIC EXPERIENCE

- Teaching assistant for the courses:
  - *Quantum Information Processing* (AY 2012/2013)
  - *Linear Algebra* (AY 2012/2013)
  - *Artificial Intelligence* (AY 2013/2014, 2014/2015)
  - *Intelligent Agents* (AY 2013/2014, 2014/2015, 2015/2016)

- Lecturing for the course:
  - *Intelligent Agents* (AY 2014/2015, 2015/2016)

- Supervising semester projects:
  - *Evaluating incentives in crowdsourcing using games* (AY 2013/2014 and 2014/2015)
  - *Bootstrapping recommender system with crowdsourcing* (AY 2013/2014 and 2014/2015)
  - *Predicting the success of altruistic requests* (AY 2014/2015).
  - *EmoMap: Emotion sensing of the EPFL campus* (AY 2015/2016)
  - *Do I hear 100?: Bidding platform for semester projects* (AY 2015/2016)
  - *Solving the codecup challenge with deep reinforcement learning* (AY 2015/2016)
  - *Auction platform for advertisement in users' pictures* (AY 2015/2016)

- Supervising master thesis:
  - *Improving accuracy of the peer truth serum mechanism* (AY 2015/2016)

- Reviewing papers in:
  - AAAI conference on Artificial Intelligence (2015)
  - Conference on Economics and Computation (EC) (2015)
  - Transactions on Economics and Computation (TEAC) (2013, 2014, 2015, 2016)

THESIS

- PhD Thesis (in progress): *Elicitation and aggregation of crowd information*
  Thesis supervisor: Prof. Boi Faltings
  Thesis topic: Mechanism design for crowdsourcing scenarios, in particular, incentive schemes that encourage accurate reports and reputation systems that limit the effectiveness of malicious reports.

- Master's Thesis: *Image steganography and steganalysis*
  Thesis supervisor: Prof. Marin Golub
  Thesis topic: Investigation of steganography procedures based on substitution and domain transformation and statistical steganalysis procedures.
  Defended on: 6 July 2010

- Bachelor thesis: *ZCS and XCS classifier systems*
  Thesis supervisor: Prof. Marin Golub
  Thesis topic: Investigation of Michigan-style learning classifier systems, in particular, ZCS and XCS algorithms, and their application to maze problems.
  Defended on: 11 July 2010

ACHIEVEMENTS

- May 2010, Second place in physics competition at international competition of electrical engineering students "Elektrijada".

- July 2007, Award "Josip Lončar" for outstanding performance in the second year of undergraduate study.

- July 2006, Award "Josip Lončar" for outstanding performance in the first year of undergraduate study

PROFESSIONAL EXPERTISE

- Fields of interest:
  - Artificial intelligence
  - Machine learning
  - Game theory and mechanism design

- Current research:
  - Algorithmic game theory: incentive schemes and reputation systems for crowdsourcing settings

- Languages:
    - Croatian: native speaker
    - English: full professional
    - French: basic

PUBLICATIONS

<u>*Conference papers*</u>:

- Radanovic G., and Faltings B. 2016. Learning to scale payments in crowdsourcing with PropeRBoost. To appear: In Proceedings of the Fourth AAAI Conference on Human Computation and Crowdsourcing (HCOMP'16).
- Radanovic G., and Faltings B. 2016. *Limiting the influence of low quality information in community sensing*. In Proceedings of the 15th International Conference on Autonomous Agents and Multiagents Systems (AAMAS'16).
- Radanovic G., and Faltings B. 2015. *Incentive schemes for participatory sensing*. In Proceedings of the 14th International Conference on Autonomous Agents and Multiagents Systems (AAMAS'15).
- Radanovic G., and Faltings B. 2015. *Incentives for subjective evaluations with private beliefs*. In Proceedings of the 29th AAAI Conference on Artificial Intelligence (AAAI'15).
- Radanovic G., and Faltings B. 2014. *Incentives for truthful information elicitation of continuous signals*. In Proceedings of the 28th AAAI Conference on Artificial Intelligence (AAAI'14).
- Radanovic G., and Faltings B. 2013. *A robust Bayesian truth serum for non-binary signals*. In Proceedings of the 27th AAAI Conference on Artificial Intelligence (AAAI'13).
- Radanovic G., and Pranjic M. 2010. *Neural network based speech synthesis for Croatian language*. MIPRO, 2010 Proceedings of the 33rd International Convention.

<u>*Journal paper*</u>:

- Radanovic G., Faltings B., and Jurca R. 2016. *Incentives for effort in crowdsourcing using the peer truth serum*. ACM Transactions on Intelligent Systems and Technology, 7(4):48:1-48:28.

<u>*Workshop paper*</u>:

- Radanovic G., and Faltings B. 2015. *Incentivizing truthful responses with the logarithmic peer truth serum*. Workshop on Mobile and Situated Crowdsourcing (WMSC'15).

Poster:

- Faltings B., Jurca R., and Radanovic G. 2014. *Eliciting truthful information with the peer truth serum*. Poster presented at: 15th ACM conference on economics and computation (EC'14).

Panel:

- Participated in the panel discussion on *Prediction markets: lessons from practice* that was a part of the 5th Workshop on Social Computing and User-Generated Content (SCUGC'15), 2015.

Tutorials:

- Presenting a tutorial on *Eliciting High-Quality Information* at the 25th International Joint Conference on Artificial Intelligence (IJCAI'16), 2016, and at the 22nd European Conference on Artificial Intelligence (ECAI'16), 2016.