BAYESIAN GROUP DECISIONS: ALGORITHMS AND COMPLEXITY

Ali Jadbabaie , Elchanan Mossel and M. Amin Rahimian

We address the computations that Bayesian agents undertake to realize their optimal actions, as they repeatedly observe each other's actions, following an initial private observation. We use iterated eliminations of infeasible signals (IEIS) to model the thinking process as well as the calculations of a Bayesian agent in a group decision scenario. We show that IEIS runs in exponential time; however, when the group structure is a partially ordered set, the Bayesian calculations simplify and polynomial-time computation of the Bayesian recommendations is possible.

We next shift attention to the case where agents reveal their beliefs (instead of actions) at every decision epoch. We analyze the computational complexity of the Bayesian belief formation in groups and show that it is $\mathcal{NP}$-hard. We also investigate the factors underlying this computational complexity and show how belief calculations simplify in special network structures or cases with strong inherent symmetries. We finally give insights about the statistical efficiency (optimality) of the beliefs and its relations to computational efficiency.

1. INTRODUCTION

Social learning or learning form actions of others is an important area of research in microeconomics. Many important real-world decision-making problems involve group interactions among individuals with purely informational externalities. Such situations arise for example in jury deliberations, expert committees, medical diagnoses, etc. We model the purely informational interactions of rational agents in a group, where they receive private information and act based on that information while also observing other people's beliefs or actions. As a Bayesian agent attempts to infer the true state of the world from her sequence of observations of actions of others as well as her own private signal, she recursively refines her belief on the signals that other players could have observed and actions that they could have taken given the assumption that other players are also rational. The existing literature addresses asymptotic and equilibrium properties of Bayesian group decisions and important questions such as convergence to consensus and learning (cf. Golub and Jackson (2010); Rosenberg et al. (2009) and their references).

The seminal work of Aumann (1976) studies the interactions of two rational agents with common prior beliefs and concludes that if the values of their posterior beliefs are common knowledge between the two agents, then the two values should be the same: rational agents cannot agree to disagree. The later work of Geanakoplos and Polemarchakis (1982) investigates how rational agents reach an agreement by communicating back and forth and refining their information partitions. Following Aumann (1976) and Geanakoplos and Polemarchakis (1982), a large body of literature studies the strategic interaction of agents in a social network, where they receive private information and act based on that information while also observing each other's actions (Acemoglu et al. (2011); Bala and Goyal (1998); Banerjee (1992); Bikhchandani et al. (1998); Gale and Kariv (2003); Mossel et al. (2014); Mueller-Frank (2013)). These observations are in turn informative about other agents' private signals; information that can be then used in making future decisions. In this line of work, it is important to understand the effectiveness of information sharing/exchange through observed actions and the effectiveness of decision-making using the available information; indeed, the quality of decision-making depends on the quality of information exchange and vice versa. In this paper, we model the purely informational interactions of rational agents in a group, where they make a private initial observation and act based upon that information while also observing other people's recommendations repeatedly; such lack of strategic externalities in group interactions arise since people are interested in each other's action, only to learn what others know which they do not know, for example, in jury deliberations, expert committees, medical diagnosis, etc.

Bayesian calculations in social settings are notoriously difficult. Successive applications of Bayes rule to the entire history of past observations leads to forebodingly complex inferences: due to unavailability of private observations as well as third party interactions that precede every decision. In general, when a rational agent observes her neighbors in a network, she should compensate for redundancies in information: the same neighbors' actions are repeatedly observed and neighboring actions may be affected by the past actions of the agent herself. Hence major challenges of Bayesian inference for social learning are due to the private signals and third party interactions that are hidden from the agent. Moreover, the existence of loops in the network causes dependencies and correlations in the information received from different neighbors, which further complicates the inference task. Failure to account for such structural dependencies subjects the agents to mistakes and inefficiencies such as redundancy neglect (by neglecting the fact that several of the neighboring agents may have been influenced by the same source of information, cf. Eyster and Rabin (2010)), and data incest (by neglecting the fact that neighboring actions may have been affected by the past actions of the agent herself, Krishnamurthy and Hoiles (2014)).

Our focus in this work is on the development of algorithms for Bayesian decision-making in groups and characterizing their complexity. We are interested in the computations that the Bayesian agent should undertake to achieve the goal of producing best recommendations at every decision epoch during a group discussion. We are further interested in determining how the complexity of these computations scale up with the increasing network size. In Section 2, we explain the Bayesian model of decision-making in groups, the so-called group decision process, and the kind of calculations that it entails. In a forward reasoning approach, to interpret observations of the actions of others, the agent considers the causes of those actions and is able to form a Bayesian posterior by weighing all contingencies that could have lead to those actions according to their probabilities. This requires the rational agent to simulate the inferences of her neighbors at all possible actions that they could have observed, and which she cannot observe directly but can only learn about partially (and indirectly) after knowing what her neighbors do. Although this forward reasoning about causes of the actions is natural to human nature (Aguilar and Buckareff (2010)), it is extremely difficult to adapt to the complexities of a partially observed setting where hidden causes lead to a multiplicity of contingencies.

In recent works of Harel et al. (2014); Kanoria and Tamuz (2013); Mossel and Tamuz (2010), recursive techniques have been applied to analyze Bayesian decision problems with partial success. In this paper, we will use the framework of iterated eliminations to model the thinking process of a Bayesian agent in a group decision-making scenario. As the Bayesian agent attempts to infer the true state of the world from her private signal and sequence of observations of actions of others, her decision problems at every epoch can be cast recursively, as a dynamic program. By the same token, the private signals of all agents constitute the state space of the problem and with every new observation, the agent refines her knowledge about the private signals that other agents have observed, by eliminating all cases that are inconsistent with her observations under the assumption that other agents are acting rationally. In Section 3, we formalize these calculations as an iterated elimination of infeasible signals (IEIS) algorithm. The IEIS approach curbs some of the complexities of the group decision process, but only to a limited extent. In a group decision scenario, the initial private signals of the agents constitute a search space that is exponential in the size of the network. The ultimate goal of the agents is to get informed about the private signals of each other and use that information to produce the best actions. A Bayesian agent is initially informed of only her own signal; however, as the history of interactions with other group members becomes enriched, her knowledge of the possible private signals that others may have observed also gets refined; thus enabling her to make better decisions. While the search over the feasible signal profiles in the IEIS algorithm runs in exponential time, these calculations may simplify in special highly connected structures: in Subsection 3.1, we give an efficient algorithm that enables a Bayesian agent to compute her posterior beliefs at every decision epoch, where the graph structure is a partially ordered set (POSET), cf. Definition 1 for the POSET property and the respective constraints that are imposed on the network topology. We thus provide a partial answer to one of the questions raised by Mossel and Tamuz (2013), who provide an efficient algorithm for computing the Bayesian binary actions in a complete graph: we show that efficient computation is possible for non-complete graphs (POSETs) with general finite action spaces.

The Bayesian iterations during a group decision process can be cast into the framework of a partially observed Markov decision process (POMDP). Thereby, the private signals of all agents constitute the state space of the problem and the decision maker only has access to a deterministic function of the state, the so-called partial observations. In a group decison process the actions or beliefs of the neighbors constitute the partial observations. The partially observed problem and its relations to the decentralized

and team decision problems have been the subject of major contributions by Tsitsiklis and Athans (1985) and Radner (1962); in particular, the partially observed problem is known to be PSPACE-hard in the worst case (Papadimitriou and Tsitsiklis, 1987, Theorem 6). However, unlike the general POMDP, the state (private signals) in a group decision process do not undergo Markovian jumps as they are fixed at the initiation of the process. Hence, determining the complexity of the group decision process requires a different analysis. To address this requirement, in Section 4 we shift focus to a case where agents repeatedly exchange their beliefs (as opposed to announcing their best recommendations); subsequently in Section 5, we are able to show that computing the Bayesian posterior beliefs in a group decision problem is $\mathcal{NP}$-hard with respect to the increasing network size.[1] This result complements and informs the existing literature on Bayesian learning over networks; in particular, those which offer efficient algorithms for special settings such as Gaussian signals in a continuous state space (Mossel and Tamuz (2010)), or with binary actions in a complete graph (Mossel and Tamuz (2013)).

In organization science, the economic theory of teams has a rich history devoted to choosing optimal information instruments subject to limited and dispersed resources in organizations, cf. Marschak and Radner (1972) and the references therein. Some of the main issues that arise in the study of decision-making organizations are information aggregation (Csaszar and Eggers (2013)) and architecture (Ioannides (1987)).[2] Visser (2000) compares the performance of hierarchical and polyarchical organization structures in a project selection task, where each agent possesses a private signal about the quality of the projects and acts rationally (maximizing the expected pay-off from subject to her information). Limiting attention to two decision-makers, the author shows how each agent's decision reflects the organizational structure while accounting for the rationality of the other actor. Algorithmic and complexity aspects of organizational decision-making are relatively unexplored. Vassilakis (1997) uses the formalism of constraint satisfaction problems to model the product development process in organizations. The author is thus able to identify some algorithmic and structural features that help reduce backtracking and rework costs of the design process in the organization. In this paper, we address issues pertaining to the complexity of rational choice in decision-making organizations. Addition of new results in this domain can further facilitate scalable cooperation among colleagues in large organizations (cf. Remark 1).

The conflict and interplay between rationality and computational tractability in economic models of human behavior has been a focus of attention by both the earlier and the contemporary scholars of the field: for example in the early works of Herbert Simon on bounded rationality, artificial intelligence and cognitive psychology (Simon (1990)), and in the contemporary research of Vela Velupillai on the computable foundations for economics (Velupillai (2000)). The present work can be regarded as an effort in this direction; a particularly relevant recent study along these lines is due to Aaronson (2005) on complexity of agreement, who investigates the question of convergence of beliefs to a consensus and the number of messages (bits) that needs to be exchanged before one can guarantee that everybody's beliefs are close to each other. There is another relevant body of literature that is dedicated to computation of the Nash equilibria in games and characterizing their complexity (cf. Daskalakis et al. (2009) and the references therein). Our results also enrich the evolving body of literature on various inference

---

[1] We provide two reductions to known $\mathcal{NP}$-complete problems. One reduction relies on the increasing number of different types of signals that are observed by different agents in the the network. The other reduction relies on the increasing size of the agent's neighborhood (with i.i.d signals).

[2] The organizational economics literature devotes considerable attention to incentive issues and agency problems that arise in organizational decision-making (Gibbons (1998)); however, issues relating to distributed information processing and communication are less explored, cf. Bolton and Dewatripont (1994).

problems over graphs and networks. The interplay between statistical and computational complexity in such problems, as well as their complexity and algorithmic landscapes, are interesting issues, on which we elaborate in Section 6, along with other concluding remarks and discussion of future directions.

2. RATIONAL DECISION-MAKING IN GROUPS

While some analytical properties of rational learning are deduced and studied in the literature Acemoglu et al. (2011); Gale and Kariv (2003); Harel et al. (2014); Mossel et al. (2014); Mueller-Frank (2013), analysis of such problems continues to attract attention. Some of the earliest results addressing the problem of social learning are due to Banerjee (1992), and Bikhchandani et al. (1998) who consider a complete graph structure where the agent's observations are public information and also ordered in time, such that each agent has access to the observations of all the past agents. These assumptions help analyze and explain the interplay between public and private information leading to fashion, fads, herds, etc. Later results by Gale and Kariv (2003) relax some of these assumptions by considering the agents that make simultaneous observations of only their neighbors rather than the whole network, but the computational complexities limit the analysis to networks with only two or three agents. In more recent results, Mueller-Frank (2013) provides a framework of rational learning that is analytically amenable and applies to general choice and information structures. Mossel et al. (2014) analyze the problem of estimating a binary state of the world from a single initial private signal that is independent and identically distributed among the agents conditioned on the true state. The authors show that by repeatedly observing each other's best estimates of the unknown, as the size of the network increases, Bayesian agents asymptotically learn the true state with high probability. Hence, the agents are able to combine their initial private observations and learn the truth. This setting is very close to our formulation of group decision processes; however, rather than the asymptotic analysis of the probability of mistakes with the increasing network size, we are interested in the computations that each agent should undertake to realize her rational choices during the group decision process. In particular, we investigate how the complexity of these computations scale up with the increasing network size.

We now proceed to present the elements of the rational model for decision-making in a group. We assume that the signal, state, and action spaces are finite sets. Because some of our algorithms rely critically on the ability of the Bayesian agent to enumerate all possible private signals that the other network agents may have observed. We may relax this assumption in special cases where calculations are possible without resorting to exhaustive enumerations. The ultimate goal of a Bayesian agent can be described as learning enough about the private signals of all other agents in the network to be able to compute the Bayesian posterior belief about the true state, given her local observations; this, however, can be extremely complex, if not impossible.

2.1. *The Bayesian Model*

We consider a group of n agents, labeled by $[n] = \{1, \ldots, n\}$, whose interactions are represented by a fixed directed graph $\mathcal{G}$. For each agent $i \in [n]$, $\mathcal{N}_i$ denotes a neighborhood $\mathcal{N}_i \subset [n]$, whose actions are observed by agent i. We use $\delta(j, i)$ to denote the length (number of edges) of the shortest path in $\mathcal{G}$ that connects j to i. In a group decision process, each agent $i \in [n]$ receives a private signal $\mathbf{s}_i$ at the beginning and then engages in repeated interactions with other group members in the ensuing decision epochs: choosing actions and observing neighbors' choices every time. We model the topic of the discussion/group decision process by a state θ belonging to a finite set Θ. For example, in the course

of a political debate, Θ can be the set of all political parties. The state θ would, therefore, take a binary value in a bipartisan system. The value/identity of θ is not known to the agents; they all start with a prior belief about the value of θ, which is a probability distribution over the set Θ with probability mass function $\nu(\cdot) : \Theta \to [0, 1]$. We assume that this prior is common to all agents. At each time t, we denote the Bayesian posterior belief of agent i given her history of observations by its probability mass function $\boldsymbol{\mu}_{i,t}(\cdot) : \Theta \to [0, 1]$. Initially, every agent receives a private signal about the unknown θ. Each signal $\mathbf{s}_i$ belongs to a finite set $\mathcal{S}_i$ and its distribution conditioned on θ is given by $\mathbb{P}_{i,\theta}(\cdot)$ which is referred to as the *signal structure* of agent i. We use $\mathbb{P}_\theta(\cdot)$ to denote the joint distribution of the private signals of all agents, signals being independent across the agents. This independence of private signals allows us to exploit a decomposition property of feasible signal profiles in case of POSETs in Subsection 3.1 to achieve polynomial-time Bayesian computations.

Associated with every agent i is an action space $\mathcal{A}_i$ that represents all the choices available to her at every point of time $t \in \mathbb{N}_0$, and a utility $u_i(\cdot, \cdot) : \mathcal{A}_i \times \Theta \to \mathbb{R}$ which in expectation represents her preferences regarding lotteries with independent draws from $\mathcal{A}_i$ and/or Θ.[1] We assume that the preferences of agents across time are myopic. At every time $t \in \mathbb{N}$, agents i takes action $\mathbf{a}_{i,t}$ to maximize her expected utility, $\mathbb{E}_{i,t}\{u_i(\mathbf{a}_{i,t}, \theta)\}$, where the expectation is with respect to $\boldsymbol{\mu}_{i,t}$. This myopia is rooted in the underlying group decision scenario that we are modeling: the agents' goal for interacting with other group members is to come up with a decision that is more informed than if they were to act solely based on their own private data; hence, by observing the recommendations of their neighboring agents $\mathbf{a}_{j,t}$ they hope to augment their information with what their neighbors, as well as other agents in the network, know that they do not. In particular, the agent does not have the freedom to learn from consequences of their recommendations, not before committing to a choice. Specifically in the group decision scenario, the agents do not learn from the realized values of the utilities of their previous recommendations (unless they commit to their choice and end the group discussion); rather the purpose of the group discussion is to augment their information by learning from recommendations of others as much as possible before committing to a choice. The network externalities that arise in above settings are purely informational. People are therefore interacting with each other, only to learn from one another, and to improve the quality of their decisions; for example, in jury deliberations, after jurors are each individually exposed to the court proceedings, the jury enters deliberations to decide on a verdict. In another case, several doctors may examine a patient and then engage in group discussions to determine the source of an illness; lack of strategic externalities is another important characteristic of the kind of human interactions that we investigate in this paper.[2]

Accordingly, at every time t, agent i observes the most recent actions of her neighbors, $\{\mathbf{a}_{j,t-1}$ for all $j \in \mathcal{N}_i\}$, and chooses an action $\mathbf{a}_{i,t} \in \mathcal{A}_i$, maximizing her expected utility given all her observations up to time t, $\{\mathbf{a}_{j,\tau}$ for all $j \in \mathcal{N}_i$, and $\tau \leq t - 1\}$. For example in the case of two communicating

---

[1]The signal, action, and utility structures, as well as the priors, are all common knowledge among the players; this assumption of common knowledge, in particular, implies that given the same access to each other's behavior or private information distinct agents would make identical inferences; in the sense that starting from the same belief about the unknown θ, their updated beliefs given the same observations would be the same.

[2]Rosenberg et al. (2009) study the emergence of consensus under such purely informational externalities. They show that even with forward-looking agents the incentives to experiment disappear, thus leading them to a consensus subject to common utility and action structures. Mossel et al. (2015) also look at forward-looking agents with binary state and action space and propose an egalitarian condition on the topology of the network to guarantee learning in infinite networks. An egalitarian graph is one in which all degrees are bounded and every agent who is being observed by some agent i, observes her back, (possibly indirectly) through a path of bounded length.
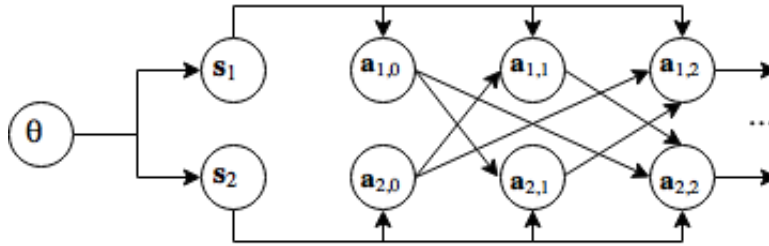
Figure 1: The Decision Flow Diagram for Two Bayesian Agents

agents the action of agent one at time two $\mathbf{a}_{1,2}$ is influenced by own private signal $\mathbf{s}_1$ as well as the neighboring action at times zero and one; part of the difficulty of the analysis is due to the fact that the action of agent two at time one is shaped not only by the private information of agent two but also by the action of agent one at time zero, cf. Fig. 1. A heuristic behavior may be justified as a mistake by interpreting actions of others as consequences of their private information, thus ignoring the history of observations when making inferences about the actions of others; in Fig. 1 this corresponds to ignoring all the arrows except those which are exiting the signal and state nodes: $\mathbf{s}_1$, $\mathbf{s}_2$, and θ. Rahimian and Jadbabaie (2016, 2017) have investigated the consequences of such heuristic behaviors in the contexts of Bayesian heuristics for group decision-making , and learning without recall when the agents observe a stream of private signals in addition to each other's actions.

In more general structures, there are also unobserved third party interactions that influence the decisions of agent two but are not available to agent one (and therefore should be inferred indirectly).

For each agent i, her history of observations $\mathbf{h}_{i,t}$ is an element of the set:

$$\mathcal{H}_{i,t} = \mathcal{S}_i \times \left(\prod_{j \in \mathcal{N}_i} \mathcal{A}_j \right)^{t-1}.$$

At every time t, the expected reward to agent i given her choice of action a_i and observed history $\mathbf{h}_{i,t}$ is given by the expected reward function $r_{i,t} : \mathcal{A}_i \times \mathcal{H}_{i,t} \to \mathbb{R}$, as follows:

$$r_{i,t}(a_i, \mathbf{h}_{i,t}) = \mathbb{E}_{i,t} \left\{ u_i(a_i, \theta) \mid \mathbf{h}_{i,t} \right\} = \sum_{\theta' \in \Theta} u_i(a_i, \theta') \boldsymbol{\mu}_{i,t}(\theta'),$$

for all $\mathbf{h}_{i,t} \in \mathcal{H}_{i,t}$, where $\boldsymbol{\mu}_{i,t}(\theta')$ is the Bayesian posterior of agent i about the truth θ given the observed history $\mathbf{h}_{i,t}$. The (myopic) optimal action of agent i at time t is then given by $\mathbf{a}_{i,t} = \arg\max_{a_i \in \mathcal{A}_i} r_{i,t}(a_i, \mathbf{h}_{i,t})$. Here we use the notation $\arg\max_{a \in \mathcal{A}}$ to include the following tie-breaking rule when the maximizer is not unique: we assume that all of the action spaces are ordered (arbitrarily) and whenever an agent is indifferent between a multitude of options she will choose the one that ranks lowest in her ordering. We further assume that all agents know about their tie breaking rules; in particular, the ordering of each action space $\mathcal{A}_i$ is known to all agents who observe agent i, directly or indirectly. This tie breaking rule induces a deterministic choice for all agents at all times as a function of their history. The restriction to deterministic tie-breaking rules is not without loss of generality. Because in the case of randomized tie-breaking rules, rational agents would have to make inferences about how past occurrences of ties have been resolved by other agents, whom they observe directly or indirectly. This is in addition to their inferences about private signals and other unknown random quantities whose values they are trying to

learn. Thus the agent's problem is to calculate her Bayesian posterior belief $\boldsymbol{\mu}_{i,t}$, given her history of past observations: $\mathbf{h}_{i,t} := \{\mathbf{s}_i, \mathbf{a}_{j,\tau}, j \in \mathcal{N}_i, \tau \in [t-1]\}$. Asymptotic properties of Bayesian group decisions, including convergence of the actions to a consensus and learning (convergence to an "optimal" aggregate action), can be studied using the Markov Bayesian equilibrium as a solution concept (cf. Appendix A); however, our main focus in this paper is on the computational and algorithmic aspects of the group decision process rather than its asymptotic properties.

Refinement of information partitions with the increasing observations is a key feature of rational learning problems and it is fundamental to major classical results that establish agreement (Geanakoplos and Polemarchakis (1982)) or learning (Blackwell and Dubins (1962); Lehrer and Smorodinsky (1996)) among rational agents. Several follow-up works of Geanakoplos and Polemarchakis (1982) have extended different aspects of information exchange among rational agents. In this line of work, it is of particular interest to derive conditions that ensure the refinement of information partitions would lead to the consensus on and/or the common knowledge of an aggregate decision.[1] In particular, Mueller-Frank (2013) points out that rational social learning requires all agents in every period to consider the set of possible information partitions of other agents and to further determine how each choice would impact the information partitions of others in the subsequent periods.

In the group decision setting, the list of feasible signals can be regarded as the information set representing the current understanding of the agent about her environment and the way additional observations are informative is by trimming the current information set and reducing the ambiguity in the set of initial signals that have caused the agent's history of past observations. In Section 3, we describe a recursive implementation for the refinement of the information sets (partitions) that relies on iterated elimination of infeasible signals (IEIS) for all the agents. The IEIS calculations scale exponentially with the network size; this is true with the exception of some very well-connected agents who have, indeed, direct access to all the observations of their neighbors and can thus analyze the decisions of each of their neighbors based on their respective observations. We expand on this special case (called POSETs) in Subsection 3.1 and explain how the Bayesian calculations simplify as a result.

3. ITERATED ELIMINATION OF INFEASIBLE SIGNALS (IEIS)

Building on the prior works of Kanoria and Tamuz (2013); Mueller-Frank (2013), we implement the refinement of information partitions for rational agents in a group decision process as an iterated elimination of infeasible signals. Accordingly, at every decision time, the signal profiles that are inconsistent with the most recent observations are removed, leading to a refined information set for next period. In this section, we analyze the Bayesian calculations that take place among the group members as they calculate their refined information partitions and the corresponding beliefs. To calculate their Bayesian posteriors, each of the agents keeps track of a list of possible combinations of private signals of all the other agents. At each iteration, they refine their list of feasible signal profiles in accordance with the most recent actions of their neighbors.

To proceed, let $\overline{s} = (s_1, \ldots, s_n) \in \mathcal{S}_1 \times \ldots \times \mathcal{S}_n$ be any profile of initial signals observed by each agent across the network, and denote the set of all private signal profiles that agent i regards as feasible

---

[1]Some notable examples include the works of Bacharach (1985); Cave (1983); Parikh and Krasucki (1990), which consider information exchange by repeatedly reporting the values of a general set function $f(\cdot)$ over the state space (rather than the conditional probabilities, which are the Bayesian beliefs). Bacharach (1985); Cave (1983) propose a condition of union consistency on $f(\cdot)$ and Parikh and Krasucki (1990) replace this union consistency condition with a convexity property for $f(\cdot)$, all ensuring that the value of $f(\cdot)$ become common knowledge among the agents after repeated exchanges.

at time t, i.e. her information set at time t, by $\boldsymbol{\mathcal{I}}_{i,t} \subset \mathcal{S}_1 \times \ldots \times \mathcal{S}_n$; this set is a random set, as it is determined by the random observations of agent i up to time t. Starting from $\boldsymbol{\mathcal{I}}_{i,0} = \{\mathbf{s}_i\} \times \prod_{j \neq i} \mathcal{S}_j$, at every decision epoch agent i removes those signal profiles in $\boldsymbol{\mathcal{I}}_{i,t-1}$ that are not consistent with her history of observations $\mathbf{h}_{i,t}$ and comes up with a trimmed set of signal profiles $\boldsymbol{\mathcal{I}}_{i,t} \subset \boldsymbol{\mathcal{I}}_{i,t-1}$ to form her Bayesian posterior belief and make her decision at time t. The set of feasible signals $\boldsymbol{\mathcal{I}}_{i,t}$ is mapped to a Bayesian posterior for agent i at time t as follows:

$$(3.1) \qquad \boldsymbol{\mu}_{i,t}(\theta) = \frac{\sum_{\overline{s} \in \boldsymbol{\mathcal{I}}_{i,t}} \mathbb{P}_\theta(\overline{s}) \nu(\theta)}{\sum_{\theta' \in \Theta} \sum_{\overline{s} \in \boldsymbol{\mathcal{I}}_{i,t}} \mathbb{P}_{\theta'}(\overline{s}) \nu(\theta')},$$

which in turn enables the agent to choose an optimal (myopic) action given her observations:[1]

$$(3.2) \qquad \mathbf{a}_{i,t} = \arg\max_{a_i \in \mathcal{A}_i} \sum_{\theta' \in \Theta} u_i(a_i, \theta') \boldsymbol{\mu}_{i,t}(\theta').$$

For example at time zero agent i learns her private signal $\mathbf{s}_i$, this enables her to initialize her list of feasible signals: $\boldsymbol{\mathcal{I}}_{i,0} = \{\mathbf{s}_i\} \times \prod_{k \in [n] \setminus \{i\}} \mathcal{S}_k$. Subsequently, her Bayesian posterior at time zero is given by:

$$\boldsymbol{\mu}_{i,0}(\theta) = \frac{\sum_{\overline{s} \in \boldsymbol{\mathcal{I}}_{i,0}} \mathbb{P}_\theta(\overline{s}) \nu(\theta)}{\sum_{\theta' \in \Theta} \sum_{\overline{s} \in \boldsymbol{\mathcal{I}}_{i,0}} \mathbb{P}_{\theta'}(\overline{s}) \nu(\theta')} = \frac{\mathbb{P}_{i,\theta}(\mathbf{s}_i) \nu(\theta)}{\sum_{\theta' \in \Theta} \mathbb{P}_{i,\theta'}(\mathbf{s}_i) \nu(\theta')}$$

and her optimal action (recommendation) at time one is as follows:

$$\mathbf{a}_{i,0} = \arg\max_{a_i \in \mathcal{A}_i} \frac{\sum_{\theta' \in \Theta} u_i(a_i, \theta') \mathbb{P}_{i,\theta'}(\mathbf{s}_i) \nu(\theta')}{\sum_{\theta'' \in \Theta} \mathbb{P}_{i,\theta''}(\mathbf{s}_i) \nu(\theta'')} = \arg\max_{a_i \in \mathcal{A}_i} \sum_{\theta' \in \Theta} u_i(a_i, \theta') \boldsymbol{\mu}_{i,1}(\theta').$$

In IEIS, the agent not only needs to keep track of the list of private signals that are consistent with her observations, denoted by $\boldsymbol{\mathcal{I}}_{i,t}$, but also she needs to consider what other agents regard as consistent with their own observations under the particular set of initial signals. The latter consideration enables the decision maker to calculate actions of other agents under any circumstances that arise at a fixed profile of initial signals, as she tries to evaluate the feasibility of that particular signal profile given her observations. In other words, the neighbors are acting rationally in accordance with what they regard as being a feasible set of initial signal profiles. Hence, with every new observation of the neighboring actions, agent i not only refines her knowledge of other people's private signals but also her knowledge of what signal profiles other agents would regard as feasible.

For any agent $j \neq i$ and at every signal profile $\overline{s}$, let $\boldsymbol{\mathcal{I}}_{j,t}^{(i)}(\overline{s})$ be the set of all signal profiles that agent i believes have not yet been rejected by agent j, given all her observations and conditioned on the initial private signals being $\overline{s}$. Consider the feasible action calculated by agent i for agent j under the assumption that the initial private signals are prescribed by $\overline{s} = (s_1, \ldots, s_n)$, i.e.

$$(3.3) \qquad a_{j,\tau}^{(i)}(\overline{s}) = \arg\max_{a_j \in \mathcal{A}_j} \sum_{\theta' \in \Theta} u_j(a_j, \theta') \frac{\sum_{\overline{s}' \in \boldsymbol{\mathcal{I}}_{j,\tau}^{(i)}(\overline{s})} \mathbb{P}_{\theta'}(\overline{s}') \nu(\theta')}{\sum_{\theta'' \in \Theta} \sum_{\overline{s}' \in \boldsymbol{\mathcal{I}}_{j,\tau}^{(i)}(\overline{s})} \mathbb{P}_{\theta''}(\overline{s}') \nu(\theta'')}, \forall \tau \in [t],$$

---

[1]In this sense, the Bayesian posterior is a sufficient statistic for the history of observations and unlike the observation history, it does not grow in dimension with time.

$\overline{s}$	$\overline{s} = (s_1, s_2, \ldots, s_n) \in \mathcal{S}_1 \times \ldots \times \mathcal{S}_n$ is a profile of initial private signals.
$\boldsymbol{\mathcal{I}}_{i,t}$	$\boldsymbol{\mathcal{I}}_{i,t} \subset \mathcal{S}_1 \times \ldots \times \mathcal{S}_n$ is the list of all signal profiles that are deemed feasible by agent i, given her observations up until time t.
$\mathcal{I}_{j,t}^{(i)}(\overline{s})$	$\mathcal{I}_{j,t}^{(i)}(\overline{s}) \subset \mathcal{S}_1 \times \ldots \times \mathcal{S}_n$ is the list of all signal profiles that agent i believes are deemed feasible by agent j, given what agent i believes agent j could have observed up until time t conditioned on the event that the initial signals of all agents are prescribed according to $\overline{s}$.
$a_{j,t}^{(i)}(\overline{s})$	$a_{j,t}^{(i)}(\overline{s}) \in \mathcal{A}_j$ is the action that agent i deems optimal for agent j, given what agent i believes agent j could have observed up until time t conditioned on the event that the initial signals of all agents are prescribed according to $\overline{s}$.

where $\mathcal{I}_{j,\tau}^{(i)}(\overline{s})$ is defined in Table I. Given $a_{j,t}^{(i)}(\overline{s})$ for all $\overline{s} \in \boldsymbol{\mathcal{I}}_{i,t-1}$ and every $j \in \mathcal{N}_i$, the agent can reject any $\overline{s}$ for which the observed neighboring action $\mathbf{a}_{j,t}$ for some $j \in \mathcal{N}_i$ does not agree with the simulated feasible action conditioned on $\overline{s}$: $\mathbf{a}_{j,t} \neq a_{j,t}^{(i)}(\overline{s})$. To proceed, we introduce the notation $\mathcal{N}_i^{\tau}$ as the τ-th order neighborhood of agent i comprising entirely of those agents who are connected to agent i through a walk of length τ: $\mathcal{N}_i^{\tau} = \{j \in [n] : j \in \mathcal{N}_{i_1}, i_1 \in \mathcal{N}_{i_2}, \ldots, i_{\tau-1} \in \mathcal{N}_{i_\tau}, i_\tau = i, \text{ for some } i_1, \ldots, i_{\tau-1} \in [n]\}$; in particular, $\mathcal{N}_i^1 = \mathcal{N}_i$ and we use the convention $\mathcal{N}_i^0 = \{i\}$. We further denote $\bar{\mathcal{N}}_i^t := \cup_{\tau=0}^t \mathcal{N}_i^{\tau}$ as the set of all agents who are within distance t of or closer to agent i; we sometimes refer to $\bar{\mathcal{N}}_i^t$ as her t-radius ego-net.

We now describe the calculations that agent i undertakes at every time t to update her list of feasible signal profiles from $\boldsymbol{\mathcal{I}}_{i,t}$ to $\boldsymbol{\mathcal{I}}_{i,t+1}$: agent i initializes her list of feasible signals $\boldsymbol{\mathcal{I}}_{i,0} = \{\mathbf{s}_i\} \times \prod_{j \neq i} \mathcal{S}_j$; at time t she would have access to $\boldsymbol{\mathcal{I}}_{i,t}$, the list of feasible signal profiles that are consistent with her observations, as well as all signal profiles that she thinks each of the other agents would regard as feasible conditioned on any profile of initial signals: $\mathcal{I}_{j,t-\tau}^{(i)}(\overline{s})$ for all $\overline{s} \in \mathcal{S}_1 \times \ldots \times \mathcal{S}_n$, all $j \in \mathcal{N}_i^{\tau}$, and all $\tau \in [t]$. Calculations of agent i at time t enables her to update her information at time t to incorporate the newly obtained data which constitute her observations of neighbors' most recent actions $\mathbf{a}_{j,t}$ for all $j \in \mathcal{N}_i$; whence she refines $\boldsymbol{\mathcal{I}}_{i,t}$ to $\boldsymbol{\mathcal{I}}_{i,t+1}$ and updates her belief and actions accordingly, cf. (3.1) and (3.2). This is achieved as follows (recall that we use $\delta(j,i)$ to denote the length of the shortest path connecting j to i):

(**I1: BAYES-GROUP**). The information available to agent i at time t:
- $\mathcal{I}_{j,t-\tau}^{(i)}(\overline{s})$ for all $\overline{s} \in \mathcal{S}_1 \times \ldots \times \mathcal{S}_n$, all $j \in \mathcal{N}_i^{\tau}$, and all $\tau \in [t]$.
- $\mathcal{I}_{i,t}$, i.e. all signal profiles that she regards as feasible given her observations.

(**A1: BAYES-GROUP**). Calculations of agent i at time t for deciding $\mathbf{a}_{i,t+1}$:

1. For all $\overline{s} := (s_1, \ldots, s_n) \in \mathcal{S}_1 \times \ldots \times \mathcal{S}_n$ and all $j \in \mathcal{N}_i^{t+1}$ do:

   - If $\delta(j,i) = t+1$, initialize $\mathcal{I}_{j,0}^{(i)}(\overline{s}) = \{s_j\} \times \prod_{k \neq j} \mathcal{S}_k$.

   - Else initialize $\mathcal{I}_{j,t+1-\delta(j,i)}^{(i)}(\overline{s}) = \mathcal{I}_{j,t-\delta(j,i)}^{(i)}(\overline{s})$ and for all $\overline{s}' \in \mathcal{I}_{j,t+1-\delta(j,i)}^{(i)}(\overline{s})$ do:

     - For all $k \in \mathcal{N}_j$ if $a_{k,t-\tau}^{(i)}(\overline{s}') \neq a_{k,t-\tau}^{(i)}(\overline{s})$, then $\mathcal{I}_{j,t+1-\tau}^{(i)}(\overline{s}) = \mathcal{I}_{j,t+1-\tau}^{(i)}(\overline{s}) \setminus \{\overline{s}'\}$, where $a_{k,t-\tau}^{(i)}(\overline{s}')$ and $a_{k,t-\tau}^{(i)}(\overline{s})$ are calculated using (3.3), based on $\mathcal{I}_{k,t-\tau}^{(i)}(\overline{s}')$ and $\mathcal{I}_{k,t-\tau}^{(i)}(\overline{s})$.

2. Initialize $\mathcal{I}_{i,t+1} = \mathcal{I}_{i,t}$ and for all $\overline{s} \in \mathcal{I}_{i,t+1}$ do:

   - For all $j \in \mathcal{N}_i$ if $\mathbf{a}_{j,t} \neq a_{j,t}^{(i)}(\overline{s})$, then $\mathcal{I}_{i,t+1} = \mathcal{I}_{i,t+1} \setminus \{\overline{s}\}$.

In Appendix B we describe the complexity of the computations that the agent should undertake using (A1) at any time t in order to calculate her posterior probability $\boldsymbol{\mu}_{i,t+1}$ and Bayesian decision $\mathbf{a}_{i,t+1}$ given all her observations up to time t. Subsequently, we prove that:

THEOREM 1 (Complexity of IEIS)   *There exists an IEIS algorithm with an $O(n^2 M^{2n-1} mA)$ running time, which given the private signal of agent i and the previous actions of her neighbors $\{\mathbf{a}_{j,\tau} : j \in \mathcal{N}_i, \tau < t\}$ in any network structure, calculates $\mathbf{a}_{i,t}$, the updated action of agent i at time t.*

REMARK 1 (Structure and Complexity in Decision Making Organizations)   Suppose the cardinality of the set of agents who influence the decisions of agent i (her cone of influence) remains bounded with the network size: $\bar{\mathcal{N}}_i^n \leq D$ for some fixed $D \in \mathbb{N}$. In such structures, where the growth is bounded, the Bayesian computations using (A1) become polynomial, upon replacing n with fixed D in (B.2). Such bounded structures can, for example, arise as a result of horizontal growth in organizations as shown in Fig. 2. The question of structure and its relation to performance receive considerable attention in organization studies. Through a series of seminal papers, Sah and Stiglitz (1985, 1986, 1988) popularized a model of project selection in organizations to study the effect of their structures, and in particular to compare the performance of hierarchies and polyarchies. Christensen and Knudsen (2010) consider the optimal decision making structures for reducing the probability of two error types in project evaluation tasks (rejecting profitable projects, type I error, or accepting unprofitable ones, type II error). They point out that either of the hierarchical or polyarchical organization structures are suitable for reducing one error type and they can be combined optimally to produce good overall performance. They further study the incremental improvement from the addition of new decision-makers and point out that polyarchical structures allow for the information to propagate throughout the organization, while in hierarchical organizations most information is filter out on the way to the top. Therefore, from a complexity standpoint, extending hierarchies to accommodate new members can lead to better tractability with the increasing organization size.
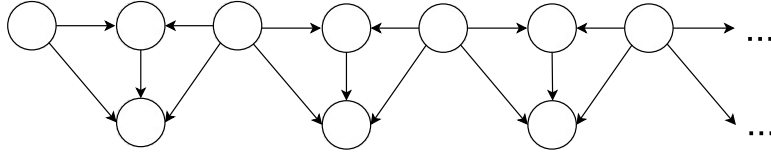
Figure 2: A structure with bounded growth: each agent is influenced by no more than three other agents even as the network (organization) size grows to infinity.

3.1. *IEIS over POSETs*

We now shift focus to the special case of POSET networks. A partially ordered set (POSET) consists of a set together with an order which is a reflexive, antisymmetric, and transitive binary relation (indicating that, for certain pairs of elements, one of them precedes the other in the ordering).

DEFINITION 1 (POSET Networks)   *We call a network structure a POSET if the directed neighborhood relationship between its nodes satisfies the reflexive and transitive properties (note that we relax the anti-symmetric property). In particular, the transitive property implies that anyone whose actions indirectly influences the observations of agent i is also directly observed by her, i.e. any neighbor of a neighbor of agent i is a neighbor of agent i as well.*

In such structures, any agnet whose actions indirectly influences the observations of agent i is also directly observed by her. Hence, any neighbor of a neighbor of agent i is a neighbor of agent i as well; the same is true for all neighbors of the neighbors of the neighbors of agent i, who would themselves be a neighbor of agent i, and so on and so forth. In particular, in a POSET network it is always true that $\mathcal{N}_i^t \subset \mathcal{N}_i^\tau \subset \mathcal{N}_i$ for all $t \leq \tau$; and in particular, $\bar{\mathcal{N}}_i^t = \mathcal{N}_i$ for all t: as time marches on, no new private signals will ever be discovered, only what is known about the private signals in the neighborhood $\mathcal{N}_i$ gets refined. In more tangible terms, the POSET requirement for an agent i would imply that in observing any of her neighbors $j \in \mathcal{N}_i$, she not only observes agent j but also observes anything that agent j observes (except for agent j's private signal).[1] The complete graph, where everybody sees the actions of everybody else is an example of a POSET structure. Such rich communication structures are in fact characteristic of round-table discussions, and they arise quite often in settings with purely informational externalities: for example in applications such as jury deliberations where jurors enter a group discussion after they each is independently exposed to the court proceedings. Other examples include group decision-makings among professionals such as medical doctors who have each made their own examination of a critically ill person and have come together to decide on the best course of treatment. A POSET structure may also arise as a result of a vested agent investigating the sources of her information; thus deliberately imposing a POSET structure on her neighborhood. Such scenarios arise when stakes are high enough as in gathering legal evidence, or documenting factual data for business decisions.

The special structure of POSET networks mitigates the issue of hidden observations, and as a result, Bayesian inferences in a POSET structure are significantly less complex. In particular, the fact that

---

[1] We can regard the directed neighborhood relationship as a binary relation on the set of vertices: i is in relation $\mathcal{R}_N$ with j iff $j \in \mathcal{N}_i$. Then the POSET property would ensure that $\mathcal{R}_N$ is a transitive relation on the set of vertices. Note that the neighborhood relationship as defined does not specify a partial order on the set of vertices because it does not satisfy the antisymmetry property. To provide for the anti-symmetry property, one needs to identify all pairs of vertices with a bidirectional link between them; thus by identifying agents who have a bidirectional link between them we obtain the neighborhood partial order $\preccurlyeq_N$ between the set of agents in a POSET group: $i \succcurlyeq_N j, \forall j \in \mathcal{N}_i$.

TABLE II

LIST OF THE VARIABLES FOR BAYESIAN CALCULATIONS IN POSET GROUPS (BAYES-POSET).

$\boldsymbol{\mathcal{S}}_{i,t}$	$\boldsymbol{\mathcal{S}}_{i,t} \subset \mathcal{S}_i$ is the list of all private signals that are deemed feasible for agent i at time t, by an agent who has observed her actions in a POSET network structure up until time t.
$\mathbf{a}_{i,t}(s_i)$	$\mathbf{a}_{i,t}(s_i) \in \mathcal{A}_i$ is the optimal choice of agent i at time t, given her observations in the POSET up until time t conditioned on the event that her initial private is s_i.
$\boldsymbol{\mathcal{I}}_{i,t}(s_i)$	$\boldsymbol{\mathcal{I}}_{i,t}(s_i) = \{s_i\} \times \prod_{j \in \mathcal{N}_i} \boldsymbol{\mathcal{S}}_{j,t}$ is the list of all signal profiles that are deemed feasible by agent i for the POSET of her neighbors, given her observations of their actions up until time t conditioned on own private signal being s_i.

the agent has access to all observations of her neighbors while observing their actions allows her to directly map an observed action to refined information about the private signals of the particular agent taking that action. We make this intuition precise in what follows by giving the exact description of the Bayesian calculations that an agent performs in a POSET structure.

Note from Table II that agent i needs only to keep track of $\boldsymbol{\mathcal{S}}_{j,t}$ for all $j \in \mathcal{N}_i$, i.e. the private signals that she deems feasible for each of her neighbors individually. This is due to the fact that in a POSET structure all agents whose actions may influence (directly or indirectly) the recommendations of a decision maker are already directly observed by her; any other agent's private signals would be immaterial to her decisions, as she would never make any observations that might have been influenced by those other agent's private signals. At the t-th decision epoch, the information that is at the disposal of agent i constitutes the list of private signals that agent i deems feasible for each of her neighbors $j \in \mathcal{N}_i$ given her observations up to time t. The goal of the agent at time t is to update her list of feasible signal profiles from $\boldsymbol{\mathcal{I}}_{i,t}$ to $\boldsymbol{\mathcal{I}}_{i,t+1}$ by incorporating her observations of her neighboring actions at time t: $\mathbf{a}_{j,t}, j \in \mathcal{N}_i$. The POSET structure allows the list of feasible signal profiles at time t to be decomposed according to the signals that are feasible for each of the neighbors individually, i.e. $\boldsymbol{\mathcal{I}}_{i,t} = \{\mathbf{s}_i\} \times \prod_{j \in \mathcal{N}_i} \boldsymbol{\mathcal{S}}_{j,t}$; the updating is thus achieved by incorporating the respective actions $\mathbf{a}_{j,t}$ for each $j \in \mathcal{N}_i$ individually and transforming the respective $\boldsymbol{\mathcal{S}}_{j,t}$ into $\boldsymbol{\mathcal{S}}_{j,t+1}$. Agent i could then refine her belief and come up with improved recommendations based on (3.1) and (3.2). After initializing $\boldsymbol{\mathcal{S}}_{\mathbf{j,0}} = \mathcal{S}_j$ for all $j \in \mathcal{N}_i$ and $\boldsymbol{\mathcal{I}}_{i,0} = \{\mathbf{s}_i\} \times \prod_{j \in \mathcal{N}_i} \boldsymbol{\mathcal{S}}_{j,0}$, at any time t the transformation from $\boldsymbol{\mathcal{I}}_{i,t}$ into $\boldsymbol{\mathcal{I}}_{i,t+1}$ in a POSET structure can be achieved as follows:

(I2: BAYES-POSET). The information available to agent i at time t:
- $\boldsymbol{\mathcal{S}}_{j,t} \subset \mathcal{S}_j$ for all $j \in \mathcal{N}_i$ is the list of private signals that agent i deems feasible for her neighbor $j \in \mathcal{N}_i$ given her observations up to time t.

(A2: BAYES-POSET). Calculations of agent i at time t for deciding $\mathbf{a}_{i,t+1}$ in a POSET:
1. For all $j \in \mathcal{N}_i$ do:
   - Initialize $\boldsymbol{\mathcal{S}}_{j,t+1} = \boldsymbol{\mathcal{S}}_{j,t}$, and for all $s_j \in \boldsymbol{\mathcal{S}}_{j,t+1}$ do:
     - Calculate $\mathbf{a}_{j,t}(s_j)$ given $\boldsymbol{\mathcal{I}}_{j,t}(s_j) = \{s_j\} \times \prod_{k \in \mathcal{N}_j} \boldsymbol{\mathcal{S}}_{k,t}$.
     - If $\mathbf{a}_{j,t} \neq \mathbf{a}_{j,t}(s_j)$, then set $\boldsymbol{\mathcal{S}}_{j,t+1} = \boldsymbol{\mathcal{S}}_{j,t+1} \setminus \{s_j\}$.
2. Update $\boldsymbol{\mathcal{I}}_{i,t+1} = \{\mathbf{s}_i\} \times \prod_{j \in \mathcal{N}_i} \boldsymbol{\mathcal{S}}_{j,t+1}$.

In Appendix C, we determine the computational complexity of (A2:BAYES-POSET) as follows:
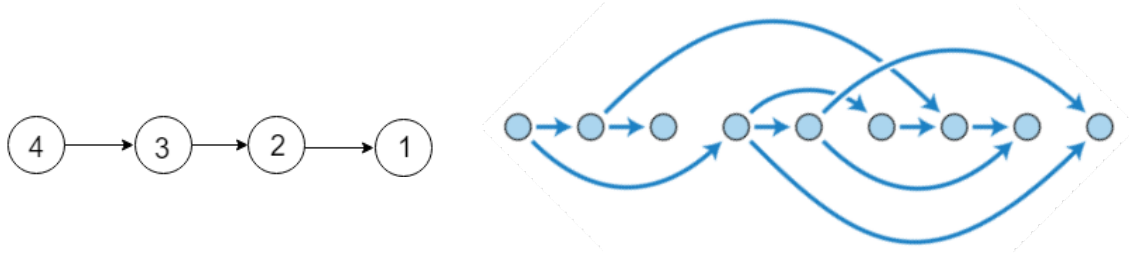
Figure 3: On the left, a directed path of length four. On the right, a directed graph is acyclic if and only if it has a topological ordering; a topological ordering of a DAG orders its vertices such that every edge goes from a lesser node (to the left) to a higher one (to the right).

THEOREM 2 (Efficient Bayesian group decisions in POSETs)   *There exists an algorithm with running time $O(Amn^2M^2)$ which given the private signal of agent i and the previous actions of her neighbors $\{\mathbf{a}_{j,\tau} : j \in \mathcal{N}_i, \tau < t\}$ in any POSET, calculates $\mathbf{a}_{i,t}$, the updated action of agent i at time t.*

The highly connected structure of POSETs leads to the rich history of observations from the neighboring actions that allows for efficient computation of Bayesian decisions in POSETs. On the other hand, one can also design efficient algorithms that are tailored to the special symmetries of the signal or network structure; for example, if all agents observe i.i.d. binary signals and take their best guess of the underlying binary state (cf. Appendix D). In Appendix D, we further observe that in a path of length n, the $(n-t)$-th agent gets fixed in her decisions after time t; and in particular, no agents will change their recommendations after $t \geq n-1$ (see the left graph in Fig. 3 for the case $n = 4$). The following proposition extends our above realization about the bounded convergence time of group decision process over paths to all directed acyclic graphs (DAGs), cf. e.g. Bang-Jensen and Gutin (2008). Such ordered structures include many cases of interest in real-world applications with a conceivable hierarchy among players: each agent observe her inferiors and is observed by her superiors or vice versa.[1] A simple examination of the dynamics in the case of two communicating agents (with a bidirectional link between them) reveals how the conclusion of this proposition can be violated in loopy structures.

PROPOSITION 1 (Bounded convergence time of group decision process over DAGs)   *Let $\mathcal{G}$ be a DAG on n nodes with a topological ordering $\prec$, and let the agents be labeled in accordance with this topological order as follows: $n \prec n-1 \prec \ldots \prec 1$. Then every agent $n-t$ gets fixed in her decisions after time t; and in particular, no agents will change their recommendations after $t \geq n-1$.*

---

[1]The key property of DAGs is their topological ordering (Bang-Jensen and Gutin, 2008, Proposition 2.1.3): a topological ordering of a directed graph is an ordering of its vertices into a sequence, such that for every edge the start vertex of the edge occurs earlier in the sequence than the ending vertex of the edge, and DAGs can be equivalently characterized as the graphs that have topological orderings. This topological ordering property allows for the influences of other agents to be addressed and analyzed in an orderly fashion, starting with the closest agents and expanding to farther and farther agents as time proceeds (see the right graph in Fig. 3). This topological ordering can be obtained by removing a vertex with no neighbors (which is guaranteed to exist in any DAG) and by repeating this procedure in the resultant DAG. Using a depth-first search (DFS) one can devise an algorithm that is linear-time in the number of nodes and edges and determines a topological ordering of a given DAG.

4. THE CASE OF REVEALED BELIEFS

Let us label $\theta_j \in \Theta := \{\theta_1, \ldots, \theta_m\}, j \in \{1, \ldots, m\}$ by $\bar{e}_j \in \mathbb{R}^m$ which is a column vector of all zeros except for its j-th element which is equal to one. Furthermore, we relax the requirement that the action spaces $\mathcal{A}_i, i \in [n]$ are finite sets; instead, for each agents $i \in [n]$ let $\mathcal{A}_i$ be the m-dimensional probability simplex: $\mathcal{A}_i = \{(x_1, \ldots, x_m)^T \in \mathbb{R}^m : \sum_1^m x_i = 1 \text{ and } x_i \geq 0, \forall i\}$. If the utility assigned to each action $\bar{a} := (a_1, \ldots, a_m)^T \in \mathcal{A}_i$ and at every state $\theta_j \in \Theta$, measures the quadratic squared distance between $\bar{a}$ and $\bar{e}_j$, then it is optimal for each agent i at any given time t to reveal her belief about the unknown state as $\sum_{\theta' \in \Theta} u_i(a_i, \theta') \boldsymbol{\mu}_{i,t}(\theta')$ in (3.2) is uniquely maximized over $a_i \in \mathcal{A}_i$ by $\mathbf{a}_{i,t} = (\boldsymbol{\mu}_{i,t}(\theta_1), \ldots, \boldsymbol{\mu}_{i,t}(\theta_m))^T$.

Thence under the prescribed quadratic utility and by taking actions over the probability simplex, agents announce their beliefs truthfully at every epoch; in practice, jurors may express their belief about the probability of guilt in a criminal case or more generally people may make statements that are expressive of their beliefs. Eyster and Rabin (2010) explain that rich-enough action spaces can reveal the underlying beliefs that lead to actions; subsequently, an individual's action is a fine reflection of her beliefs. Ali (2014) characterizes the distinction between coarse and rich action spaces using the concept of "*responsiveness*": the utility function is responsive, if a player with that utility chooses different actions at different beliefs (as is the case for the quadratic utility described above); Ali (2014) also discusses the role of responsiveness in determining the observational learning outcome.

Such agents engage in the discussion by repeatedly exchanging their beliefs about an issue of common interest, which is modeled by the state θ. For example in the course of a political debate, Θ can be the set of all political parties and it would take a binary value in a bipartisan system. The value/identity of θ is not known to the agents but they each receive a private signal about the unknown θ and starting from a full-support prior belief $\nu(\cdot)$, at any time $t \in \{0, 1, 2, \ldots\}$ each agent holds a belief $\boldsymbol{\mu}_{i,t}$, which is her Bayesian posterior on Θ given her knowledge of the signal structure and priors as well as her history of observations, which include her initial private signal as well as the beliefs that she has observed in her neighbors throughout past times $\tau < t$.

Consider the finite state space $\Theta = \{\theta_1, \ldots, \theta_m\}$ and for all $2 \leq k \leq m$, let:

$$\boldsymbol{\lambda}_i(\theta_k) := \log\left(\frac{\mathbb{P}_{i,\theta_k}(\mathbf{s}_i)}{\mathbb{P}_{i,\theta_1}(\mathbf{s}_i)}\right) , \ \boldsymbol{\phi}_{i,t}(\theta_k) := \log\left(\frac{\boldsymbol{\mu}_{i,t}(\theta_k)}{\boldsymbol{\mu}_{i,t}(\theta_1)}\right) , \ \gamma(\theta_k) := \log\left(\frac{\nu(\theta_k)}{\nu(\theta_1)}\right).$$

Here and throughout Sections 4 and 5, we assume the agents have started from uniform prior beliefs and the size of the state space is $m = 2$, thence we enjoy a slightly simpler notation: with uniform priors $\gamma_i(\theta_k) = \log\left(\nu_i(\theta_k)/\nu_i(\theta_1)\right) = 0$ for all i, k, whereas otherwise knowing the (common) priors the agents can always compensate for the effect of the priors as they observe each other's beliefs; with a binary state space $\Theta = \{\theta_1, \theta_2\}$, the agents need to only keep track of one set of belief and likelihood ratios corresponding to the pair (θ_1, θ_2), whereas in general the agents should form and calculate $m - 1$ ratio terms for each of the pairs (θ_1, θ_k), $k = 2, \ldots, m$ to have a fully specified belief. For a binary state space with no danger of confusion we can use the simplified notation $\boldsymbol{\lambda}_i = \boldsymbol{\lambda}_i(\theta_2) := \log\left(\mathbb{P}_i(\mathbf{s}_i|\theta_2)/\mathbb{P}_{i,\theta_1}(\mathbf{s}_i)\right)$, and $\boldsymbol{\phi}_{i,t} = \boldsymbol{\phi}_{i,t}(\theta_2) = \log\left(\boldsymbol{\mu}_{i,t}(\theta_2)/\boldsymbol{\mu}_{i,t}(\theta_1)\right)$.

PROBLEM 1 (GROUP-DECISION) *At any time t, given the graph structure $\mathcal{G}$, the private signal $\mathbf{s}_i$ and the history of observed neighboring beliefs $\boldsymbol{\mu}_{j,\tau}, j \in \mathcal{N}_i, \tau \in [t]$ determine the Bayesian posterior belief $\boldsymbol{\mu}_{i,t+1}$.*

In general GROUP-DECISION is a hard problem as we will describe in Section 5. Here, we introduce a special class of structures which play an important role in determining the type of calculations that agent i should undertake to determine her posterior belief (recall that the t-radius ego-net of agent i, $\bar{\mathcal{N}}_i^t$, is the set of all agents who are within distance t of or closer to agent i):

DEFINITION 2 (Transparency)   *The graph structure $\mathcal{G}$ is transparent to agent i at time t, if for all $j \in \mathcal{N}_i$ and every $\tau \leq t - 1$ we have that: $\phi_{j,\tau} = \sum_{k \in \bar{\mathcal{N}}_j^\tau} \lambda_k$, for any choice of signal structures and all possible initial signals.*

The initial belief exchanges reveal the likelihoods of the private signals in the neighboring agents. Hence, from her observations of the beliefs of her neighbors at time zero $\{\boldsymbol{\mu}_{j,0}, j \in \mathcal{N}_i\}$, agent i learns all that she ever needs to know regarding the private signals of her neighbors so far as they influence her beliefs about the unknown state θ. However, the future neighboring beliefs (at time one and beyond) are less "*transparent*" when it comes to reflecting the neighbors' knowledge of other private signals that are received throughout the network. In particular, the time one beliefs of the neighbors $\phi_{j,1}, j \in \mathcal{N}_i$ is given by $\phi_{j,1} = \sum_{k \in \bar{\mathcal{N}}_j^1} \lambda_k$; hence, from observing the time one belief of a neighbor, agent i would only get to know $\sum_{k \in \mathcal{N}_j} \lambda_k$, rather than the individual values of λ_k for each $k \in \mathcal{N}_j$, which her neighbor j had gotten to know before reporting the belief $\phi_{j,1} = \sum_{k \in \bar{\mathcal{N}}_j^1} \lambda_k$ to agent i. Indeed, this is a fundamental aspect of inference problems in observational learning (in learning from other actors): similar to responsiveness that Ali (2014) defines as a property of the utility functions to determine whether players' beliefs can be inferred from their actions, *transparency* in our belief exchange setup is defined as a property of the graph structure (see Remark 2 on why transparency is a structural property) which determines to what extent other players' private signals can be inferred from observing the neighboring beliefs. We also have the following simple consequence:

COROLLARY 1 (Transparency at time one)   *All graphs are transparent to all their agents at time one.*

REMARK 2 (Transparency, statistical efficiency, and impartial inference)   Such agents j whose beliefs satisfy the equation in Definition 2 at some time τ are said to hold a *transparent* or *efficient* belief; the latter signifies the fact that the such a belief coincides with the Bayesian posterior if agent j were given direct access to the private signals of every agent in $\bar{\mathcal{N}}_j^\tau$. This is indeed the best possible (or statistically efficient) belief that agent j can hope to form given the information available to her at time τ; it specializes the perfect aggregation property of Appendix D to the case of revealed beliefs. The same connection to the statistically efficient beliefs arise in the work of Eyster and Rabin (2014) who formulate the closely related concept of "impartial inference" in a model of sequential decisions by different players in successive rounds; accordingly, impartial inference ensures that the full informational content of all signals that influence a player's beliefs can be extracted and players can fully (rather than partially) infer their predecessors' signals. In other words, under impartial inference, players' immediate predecessors provide "sufficient statistics" for earlier movers that are indirectly observed (Eyster and Rabin, 2014, Section 3). Last but not least, it is worth noting that statistical efficiency or impartial inference are properties of the posterior beliefs, and as such the signal structures may be designed so that statistical efficiency or impartial inference hold true for a particular problem setting; on the other hand, transparency is a structural property of the network and would hold true for any choice of signal structures and all possible initial signals.

The following is a sufficient graphical condition for agent i to hold an efficient (transparent) belief at time t: there are no agents $k \in \bar{\mathcal{N}}_i^t$ that has multiple paths to agent i, unless it is among her neighbors (agent k is directly observed by agent i).

PROPOSITION 2 (Graphical condition for transparency) *Agent i will hold a transparent (efficient) belief at time t if there are no $k \in \bar{\mathcal{N}}_i^t \backslash \mathcal{N}_i$ such that for $j \neq j'$, both j and j' belonging to $\mathcal{N}_i$, we have $k \in \bar{\mathcal{N}}_j^{t-1}$ and $k \in \bar{\mathcal{N}}_{j'}^{t-1}$.*

PROOF: The proof follows by induction on t, i.e. by considering the agents whose information reach agent i for the first time at t. The claim is trivially true at time one, since agent i can always infer the likelihoods of the private signals of each of her neighbors by observing their beliefs at time one. Now consider the belief of agent i at time t, the induction hypothesis implies that $\phi_{i,t-1} = \sum_{k \in \bar{\mathcal{N}}_i^{t-1}} \lambda_k$, as well as $\phi_{j,t-1} = \sum_{k \in \bar{\mathcal{N}}_j^{t-1}} \lambda_k$ and $\phi_{j,t-2} = \sum_{k \in \bar{\mathcal{N}}_j^{t-2}} \lambda_k$ for all $j \in \mathcal{N}_i$. To form her belief at time t (or equivalently its log-ratio $\phi_{i,t}$), agent i should consider her most recent information $\{\phi_{j,t-1} = \sum_{k \in \bar{\mathcal{N}}_j^{t-1}} \lambda_k, j \in \mathcal{N}_i\}$ and use that to update her current belief $\phi_{i,t-1} = \sum_{k \in \bar{\mathcal{N}}_i^{t-2}} \lambda_k$. To prove the induction claim, it suffices to show that agent i has enough information to calculate the sum of log-likelihood ratios of all signals in her t-radius ego-net, $\bar{\mathcal{N}}_i^t$; i.e. to form $\phi_{i,t} = \sum_{k \in \bar{\mathcal{N}}_i^t} \lambda_k$. This is the best possible belief that she can hope to achieve at time t, and it is the same as her Bayesian posterior, had she direct access to the private signals of all agents in her t-radius ego-net. To this end, by using her knowledge of $\phi_{j,t-1}$ and $\phi_{j,t-2}$ she can form:

$$\hat{\phi}_{j,t-1} = \phi_{j,t-1} - \phi_{j,t-2} = \sum_{k \in \bar{\mathcal{N}}_j^{t-1} \backslash \bar{\mathcal{N}}_j^{t-2}} \lambda_k,$$

for all $j \in \mathcal{N}_i$. Since, $\phi_{i,t-1} = \sum_{k \in \bar{\mathcal{N}}_i^{t-1}} \lambda_k$ by the induction hypothesis, the efficient belief $\phi_{i,t} = \sum_{k \in \bar{\mathcal{N}}_i^t} \lambda_k$ can be calculated if and only if,

$$(4.1) \qquad \hat{\phi}_{i,t} = \phi_{i,t} - \phi_{i,t-1} = \sum_{k \in \bar{\mathcal{N}}_i^t \backslash \bar{\mathcal{N}}_i^{t-1}} \lambda_k,$$

can be computed. In the above formulation $\hat{\phi}_{i,t}$ is an innovation term, representing the information that agent i learns from her most recent observations at time t. We now show that under the assumption that any agent with multiple paths to an agent i is directly observed by her, the innovation term in (4.1) can be constructed from the knowledge of $\phi_{j,t-1} = \sum_{k \in \bar{\mathcal{N}}_j^{t-1}} \lambda_k$, and $\phi_{j,t-2} = \sum_{k \in \bar{\mathcal{N}}_j^{t-2}} \lambda_k$ for all $j \in \mathcal{N}_i$; indeed, we show that:

$$(4.2) \qquad \hat{\phi}_{i,t} = \sum_{j \in \mathcal{N}_i} \left(\hat{\phi}_{j,t-1} - \sum_{\substack{k \in \mathcal{N}_i: \\ \delta(k,j) = t-1}} \phi_{k,0} \right), \text{ for all } t > 1.$$

Consider any $k \in \bar{\mathcal{N}}_i^t \backslash \bar{\mathcal{N}}_i^{t-1}$, these are all agents which are at distance exactly t, $t > 1$, from agent i, and no closer to her. No such $k \in \bar{\mathcal{N}}_i^t \backslash \bar{\mathcal{N}}_i^{t-1}$ is a direct neighbor of agent i and the structural assumption therefore implies that there is a unique neighbor of agent i, call this unique neighbor $j_k \in \mathcal{N}_i$, satisfying $k \in \bar{\mathcal{N}}_{j_k}^{t-1} \backslash \bar{\mathcal{N}}_{j_k}^{t-2}$. On the other hand, consider any $j \in \mathcal{N}_i$ and some $k \in \bar{\mathcal{N}}_j^{t-1} \backslash \bar{\mathcal{N}}_j^{t-2}$, such an agent k is either a neighbor of i or else at distance exactly $t > 1$ from agent i and therefore $k \in \bar{\mathcal{N}}_i^t \backslash \bar{\mathcal{N}}_i^{t-1}$,

and element j would be the unique neighbor $j_k \in \mathcal{N}_i$ satisfying $k \in \bar{\mathcal{N}}_{j_k}^{t-1} \setminus \bar{\mathcal{N}}_{j_k}^{t-2}$. Subsequently, we can partition

$$\bar{\mathcal{N}}_i^t \setminus \bar{\mathcal{N}}_i^{t-1} = \uplus_{j \in \mathcal{N}_i} \bar{\mathcal{N}}_j^{t-1} \setminus \left(\bar{\mathcal{N}}_j^{t-2} \cup \mathcal{N}_i \right),$$

and therefore we can rewrite the left-hand side of (4.1) as follows:

$$\hat{\phi}_{i,t} = \sum_{\substack{k \in \\ \bar{\mathcal{N}}_i^t \setminus \bar{\mathcal{N}}_i^{t-1}}} \lambda_k = \sum_{\substack{k \in \uplus_{j \in \mathcal{N}_i} \\ \bar{\mathcal{N}}_j^{t-1} \setminus (\bar{\mathcal{N}}_j^{t-2} \cup \mathcal{N}_i)}} \lambda_k = \sum_{j \in \mathcal{N}_i} \sum_{\substack{k \in \bar{\mathcal{N}}_j^{t-1} \setminus \\ (\bar{\mathcal{N}}_j^{t-2} \cup \mathcal{N}_i)}} \lambda_k$$

$$= \sum_{j \in \mathcal{N}_i} \left(\sum_{\substack{k \in \bar{\mathcal{N}}_j^{t-1} \setminus \\ \bar{\mathcal{N}}_j^{t-2}}} \lambda_k - \sum_{\substack{k \in \mathcal{N}_i \cap \\ \bar{\mathcal{N}}_j^{t-1} \setminus \bar{\mathcal{N}}_j^{t-2}}} \lambda_k \right) = \sum_{j \in \mathcal{N}_i} \left(\hat{\phi}_{j,t-1} - \sum_{\substack{k \in \mathcal{N}_i: \\ \delta(k,j)=t-1}} \phi_{k,0} \right),$$

as claimed in (4.2), completing the proof.   Q.E.D.

Note that in the course of the proof of Proposition 2, for the structures that satisfy the sufficient condition for transparency, we obtain a simple algorithm for updating beliefs by setting the total innovation at every step equal to the sum of the most recent innovations observed at each of the neighbors, correcting for those neighbors who are being recounted:

1. Initialize: $\phi_{i,0} = \lambda_i$, $\hat{\phi}_{j,0} = \phi_{j,0} = \lambda_j$, $\phi_{i,1} = \sum_{j \in \bar{\mathcal{N}}_i^1} \phi_{j,0}$.

2. For $t > 1$ set: $\hat{\phi}_{j,t-1} = \phi_{j,t-1} - \phi_{j,t-2}$, $\hat{\phi}_{i,t} = \sum_{j \in \mathcal{N}_i} [\hat{\phi}_{j,t-1} - \sum_{\substack{k \in \mathcal{N}_i: \\ \delta(k,j)=t-1}} \phi_{k,0}]$, $\phi_{i,t} = \phi_{i,t-1} + \hat{\phi}_{i,t}$.

Rooted (directed) trees are a special class of transparent structures, which also satisfy the sufficient structural condition of Proposition 2; indeed, in case of a rooted tree for any agent k that is indirectly observed by agent i, there is a unique path connecting k to i. As such the correction terms for the sum of innovations observed in the neighbors is always zero, and we have $\hat{\phi}_{i,t} = \sum_{j \in \mathcal{N}_i} \hat{\phi}_{j,t-1}$), i.e. the innovation at every time step is equal to the total innovations observed in all the neighbors.

EXAMPLE 1 (Transparent structures)   Fig. 4 illustrates cases of transparent and nontransparent structures. We refer to them as first, second, third, and forth in their respective order from left to right. All structures except the first one are transparent. To see how the transparency is violated in the first structure, consider the beliefs of agent i: $\phi_{i,0} = \lambda_i$, $\phi_{i,1} = \lambda_i + \lambda_{j_1} + \lambda_{j_2}$; at time two, agent 1 observes $\phi_{j_1,1} = \lambda_{j_1} + \lambda_{k_1} + \lambda_{k_2}$ and $\phi_{j_2,1} = \lambda_{j_2} + \lambda_{k_2} + \lambda_{k_3}$. Knowing $\phi_{j_1,0} = \lambda_{j_1}$ and $\phi_{j_2,0} = \lambda_{j_2}$ she can infer the value of the two sub-sums $\lambda_{k_1} + \lambda_{k_2}$ and $\lambda_{k_1} + \lambda_{k_3}$, but there is no way for her to infer their total sum $\lambda_{j_1} + \lambda_{j_2} + \lambda_{k_1} + \lambda_{k_2} + \lambda_{k_3}$. Agent i cannot hold an efficient or transparent belief at time two. The issue is resolved in the second structure by adding a direct link so that agent k_2 is directly observed by agent i; the sufficient structural condition of Proposition 2 is thus satisfied and we have $\phi_{i,2} = \lambda_i + \lambda_{j_1} + \lambda_{j_2} + \lambda_{k_1} + \lambda_{k_2} + \lambda_{k_3}$. In structure three, we have $\phi_{i,2} = \lambda_i + \lambda_{j_1} + \lambda_{j_2} + \lambda_{k_1} + \lambda_{k_2} = \lambda_i + \phi_{j_1,1} + \phi_{j_2,0}$. Structure four is also transparent and we have $\phi_{i,2} = \lambda_i + \lambda_{j_1} + \lambda_{j_2} + \lambda_{k_1} + \lambda_{k_2} + \lambda_{k_3} + \lambda_{k_4} = \lambda_i + \phi_{j_1,1} + \phi_{j_2,1}$ and $\phi_{i,3} = \lambda_i + \lambda_{j_1} + \lambda_{j_2} + \lambda_{k_1} + \lambda_{k_2} + \lambda_{k_3} + \lambda_{k_4} + \lambda_l = \lambda_i + \phi_{j_1,1} + \phi_{j_2,1} + (\phi_{j_1,2} - \phi_{j_1,1})$, where in the last equality we have used the fact that $\lambda_l = (\phi_{j_1,2} - \phi_{j_1,1})$. In particular, note that structures three and four violate the sufficient structural condition laid out in Proposition 2, despite both being transparent.
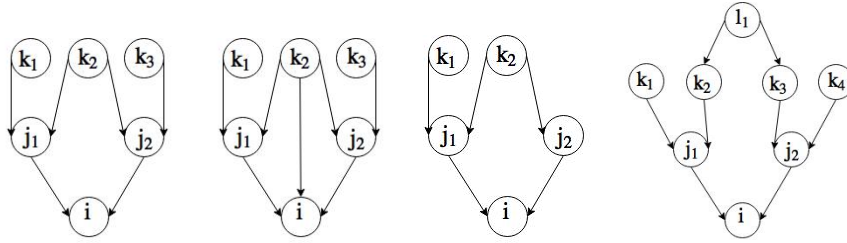
Figure 4: The last three structures are transparent but the first one is not.

When the transparency condition is violated, the neighboring agent's beliefs is a complex non-linear function of the signal likelihoods of the upstream (indirectly observed) neighbors. Therefore, making inferences about the unobserved private signals from such "*nontransparent*" beliefs is a very complex task: it ultimately leads to agent i reasoning about feasible signal profiles that are consistent with her observations similar to the IEIS algorithm (A1:BAYES-GROUP). We elaborate on the belief calculations for the nontransparent case in Subsection 4.2, where we provide the version of IEIS algorithm that is tailored to belief communications and it can be used in the most general cases with nontransparent structures. When the transparency conditions are satisfied, the beliefs of the neighboring agents reveal the sum of log-likelihoods for the private signals of other agents within a distance t of agent i. Nevertheless, even when the network is transparent to agent i, cases arise where efficient algorithms for calculating Bayesian posterior beliefs for agent i are unavailable and indeed impossible (if $\mathcal{P} \neq \mathcal{NP}$). In Subsection 4.1, we describe the calculations of the Bayesian posterior belief when the transparency condition is satisfied. In Section 5, we show that well-known $\mathcal{NP}$-complete problems are special cases of the GROUP-DECISION problem and as such the latter is $\mathcal{NP}$-hard; there we also describe special cases where a more positive answer is available and provide an efficient algorithm accordingly.

4.1. *Belief Calculations in Transparent Structures*

Here we describe calculations of a Bayesian agent in a transparent structure. If the network is transparent to agent i, she has access to the following information from the beliefs that she has observed in her neighbors at times $\tau \leq t$, before deciding her belief for time $t + 1$:

- Her own signal $\mathbf{s}_i$ and its log-likelihood $\boldsymbol{\lambda}_i$.
- Her observations of the neighboring beliefs: $\{\boldsymbol{\mu}_{j,\tau} : j \in \mathcal{N}_i, \tau \leq t\}$. Due to transparency, these beliefs reveal the following information about sums of log-likelihoods of private signals of subsets of other agents in the network: $\sum_{k \in \bar{\mathcal{N}}_j^\tau} \boldsymbol{\lambda}_k = \boldsymbol{\phi}_{i,\tau}$, for all $\tau \leq t$, and any $j \in \mathcal{N}_i$.

From the information available to her, agent i aims to learn as much as possible about the likelihoods of the private signals of others whom she does not observe; indeed, as she has already learned the likelihoods of the signals that her neighbors have observed from their reported beliefs at time one, at times $t > 1$ she is interested in learning about the agents that are further away from her up to the distance t. Her best hope for time $t + 1$ is to learn the sum of log-likelihoods of the signals of all agents that are within distance of at most $t+1$ from her in the graph and to set her posterior belief accordingly; this however is not always possible as demonstrated for agent i in the leftmost graph of Fig. 4. To decide her belief, agent i constructs the following system of linear equations in card $\left(\bar{\mathcal{N}}_{t+1}\right) + 1$ unknowns: $\{\boldsymbol{\lambda}_j : j \in \bar{\mathcal{N}}_{t+1}, \text{ and } \bar{\boldsymbol{\lambda}}_{i,t+1}\}$, where $\bar{\boldsymbol{\lambda}}_{i,t+1} = \sum_{j \in \bar{\mathcal{N}}_{t+1}} \boldsymbol{\lambda}_j$ is the best possible (statistically efficient) belief

for agent i at time $t+1$:

$$(4.3) \quad \begin{cases} \sum_{k \in \bar{\mathcal{N}}_j^\tau} \boldsymbol{\lambda}_k = \boldsymbol{\phi}_{j,\tau}, \text{ for all } \tau \leq t, \text{ and any } j \in \mathcal{N}_i, \\ \sum_{j \in \bar{\mathcal{N}}_i^{t+1}} \boldsymbol{\lambda}_j - \bar{\boldsymbol{\lambda}}_{i,t+1} = 0. \end{cases}$$

Agent i can apply the Gauss-Jordan method and convert the system of linear equations in $\mathrm{card}\left(\bar{\mathcal{N}}_i^{t+1}\right)$ $+\ 1$ variables to its reduced row echelon form. Next if in the reduced row echelon form $\bar{\boldsymbol{\lambda}}_{i,t}$ is a basic variable with fixed value (its corresponding column has a unique non-zero element that is a one, and that one belongs to a row with all zero elements except itself), then she sets her belief optimally such that $\boldsymbol{\phi}_{i,t+1} = \bar{\boldsymbol{\lambda}}_{i,t+1}$; this is the statistically efficient belief at time $t+1$. Recall that in the case of a binary state space, log-belief ratio $\boldsymbol{\phi}_{i,t+1}$ uniquely determines the belief $\boldsymbol{\mu}_{i,t+1}$.

Statistical versus Computational Efficiency

Having $\boldsymbol{\phi}_{i,t+1} = \bar{\boldsymbol{\lambda}}_{i,t+1}$ signifies the best achievable belief given the observations of the neighboring beliefs as it corresponds to the statistically efficient belief that the agent would have adopted, had she direct access to the private signals of every agent within distance $t+1$ from her; notwithstanding the efficient case $\boldsymbol{\phi}_{i,t+1} = \bar{\boldsymbol{\lambda}}_{i,t+1}$ does not necessarily imply that agent i learns the likelihoods of the signals of other agents in $\bar{\mathcal{N}}_i^{t+1}$; indeed, this was the case for agent i in the forth (transparent) structure of Example 1: agent i learns $\{\boldsymbol{\lambda}_i, \boldsymbol{\lambda}_{j_1}, \boldsymbol{\lambda}_{j_2}, \boldsymbol{\lambda}_{k_1} + \boldsymbol{\lambda}_{k_2}, \boldsymbol{\lambda}_{k_3} + \boldsymbol{\lambda}_{k_4}, \boldsymbol{\lambda}_l\}$ and in particular can determine the efficient beliefs $\bar{\boldsymbol{\lambda}}_{i,2} = \boldsymbol{\lambda}_i + \boldsymbol{\lambda}_{j_1} + \boldsymbol{\lambda}_{j_2} + \boldsymbol{\lambda}_{k_1} + \boldsymbol{\lambda}_{k_2} + \boldsymbol{\lambda}_{k_3} + \boldsymbol{\lambda}_{k_4}$ and $\bar{\boldsymbol{\lambda}}_{i,3} = \boldsymbol{\lambda}_i + \boldsymbol{\lambda}_{j_1} + \boldsymbol{\lambda}_{j_2} + \boldsymbol{\lambda}_{k_1} + \boldsymbol{\lambda}_{k_2} + \boldsymbol{\lambda}_{k_3} + \boldsymbol{\lambda}_{k_4} + \boldsymbol{\lambda}_l$, but she never learns the actual values of the likelihoods $\{\boldsymbol{\lambda}_{k_1}, \boldsymbol{\lambda}_{k_2}, \boldsymbol{\lambda}_{k_3}, \boldsymbol{\lambda}_{k_4}\}$, individually. In other words, it is possible for agent i to determine the sum of log-likelihoods of signals of agents in her higher-order neighborhoods even though she does not learn about each signal likelihood individually. The case where $\bar{\boldsymbol{\lambda}}_{i,t+1}$ can be determined uniquely so that $\boldsymbol{\phi}_{i,t+1} = \bar{\boldsymbol{\lambda}}_{i,t+1}$, is not only statistically efficient but also computationally efficient as complexity of determining the Bayesian posterior belief at time $t+1$ is the same as the complexity of performing Gauss-Jordan steps which is $O(n^3)$ for solving the $t \cdot \mathrm{card}(\mathcal{N}_i)$ equations in $\mathrm{card}(\bar{\mathcal{N}}_i^{t+1})$ unknowns. Note that here we make no attempt to optimize these computations beyond the fact that their growth is polynomial in n. This is an interesting alignment that emerges between statistical and computational efficiency in group decision process, and it is in contrast with the trade-off between statistical and computational performance that is reported in other graphical inference problems such as sparse principal component analysis, planted partition and stochastic block models, as well as sub-matrix localization, where there is an "*information-computation gap*" between what is achievable in polynomial-time and what is statistically optimal (achieves the information theoretic limit); cf. Chen and Xu (2016); Wang et al. (2016).

Next we consider the case where $\bar{\boldsymbol{\lambda}}_{i,t+1}$ is not a basic variable in the reduced row echelon form of system (4.3) or it is a basic variable but its value is not fixed by the system and depends on how the free variables are set. In such cases agent i does not have access to the statistically efficient belief $\bar{\boldsymbol{\lambda}}_{i,t+1}$. Instead she has to form her Bayesian posterior belief by inferring the set of all feasible signals for all agents in $\bar{\mathcal{N}}_i^{t+1}$ whose likelihoods are consistent with the system (4.3). To this end, she keeps track of the set of all signal profiles at any time t that are consistent with her information, system (4.3), at that time. Following the IEIS procedure of Section 3, let us denote the set of feasible signal profiles for agent i at time t by $\mathcal{I}_{i,t}$. The general strategy of agent i, would be to search over all elements of $\mathcal{I}_{i,t}$ and to eliminate (refute) any signal profile $\bar{s}$ that is inconsistent with (i.e. does not satisfy) the $\mathcal{N}_i$ new equations revealed to her from the transparent beliefs of her neighbors. For a signal profile $\bar{s} = (s_1, s_2, \ldots, s_n)$,

let $\lambda_i(\overline{s}) := \log(\mathbb{P}_{i,\theta_2}(s_i)/\mathbb{P}_{i,\theta_1}(s_i))$ denote the log-likelihood ratio of its i-th component private signal. Given the list of feasible signal profiles $\mathcal{I}_{i,t}$ for agent i at time t, we formalize the calculations of agent i, subject to observation of the transparent beliefs of her neighbors $\phi_{j,t}, j \in \mathcal{N}_i$, as follows:

(A3: BAYES-TRANSPARENT). Calculations of agent i at time t for deciding $\boldsymbol{\mu}_{i,t+1}$ in a structure that is transparent to her:

1. Initialize $\mathcal{I}_{i,t+1} = \mathcal{I}_{i,t}$.
2. For all $\overline{s} \in \mathcal{I}_{i,t+1}$ and any $j \in \mathcal{N}_i$ do:

   - If $\phi_{j,t} \neq \sum_{k \in \bar{\mathcal{N}}_j^t} \lambda_k(\overline{s})$, then set $\mathcal{I}_{i,t+1} = \mathcal{I}_{i,t+1} \setminus \{\overline{s}\}$.

3. Given $\mathcal{I}_{i,t+1}$, calculate the updated belief $\boldsymbol{\mu}_{i,t+1}$ according to (3.1).

Despite the relative simplification that is brought about by transparency, in general there is an exponential number of feasible signal profiles and verifying them for the new $\mathcal{N}_i$ equations would take exponential time. The belief calculations may be optimized by inferring the largest subset of individual likelihood ratios whose summation is fixed by system (4.3). The verification and refutation process can then be restricted to the remaining signals whose sum of log-likelihoods is not fixed by system (4.3). For example in leftmost structure of Fig. 4, agent i will not hold a transparent belief at time 2 but she can determine the sub-sum $\boldsymbol{\lambda}_i + \boldsymbol{\lambda}_{j_1} + \boldsymbol{\lambda}_{j_2}$ and her belief would involve a search only over the profile of the signals of the remaining agents $(s_{k_1}, s_{k_2}, s_{k_3})$. At time two, she finds all $(s_{k_1}, s_{k_2}, s_{k_3})$ that agree with the additionally inferred sub-sums $\phi_{j_1,1} - \phi_{j_1,0} = \boldsymbol{\lambda}_{k_1} + \boldsymbol{\lambda}_{k_2}$ and $\phi_{j_2,1} - \phi_{j_2,0} = \boldsymbol{\lambda}_{k_2} + \boldsymbol{\lambda}_{k_3}$; indeed we can express $\phi_{i,2}$ as follows:

$$\phi_{i,2} = \boldsymbol{\lambda}_i + \boldsymbol{\lambda}_{j_1} + \boldsymbol{\lambda}_{j_2} + \log \frac{\sum_{(s_{k_1}, s_{k_2}, s_{k_3}) \in \mathcal{I}_{i,2}} \mathbb{P}_{k_1,\theta_2}(s_{k_1}) \mathbb{P}_{k_2,\theta_2}(s_{k_2}) \mathbb{P}_{k_3,\theta_2}(s_{k_3})}{\sum_{(s_{k_1}, s_{k_2}, s_{k_3}) \in \mathcal{I}_{i,2}} \mathbb{P}_{k_1,\theta_1}(s_{k_1}) \mathbb{P}_{k_2,\theta_1}(s_{k_2}) \mathbb{P}_{k_3,\theta_1}(s_{k_3})},$$

where

$$\mathcal{I}_{i,2} = \{(s_{k_1}, s_{k_2}, s_{k_3}) : \log \frac{\mathbb{P}_{k_1,\theta_2}(s_{k_1})}{\mathbb{P}_{k_1,\theta_1}(s_{k_1})} + \log \frac{\mathbb{P}_{k_2,\theta_2}(s_{k_2})}{\mathbb{P}_{k_2,\theta_1}(s_{k_2})} = \boldsymbol{\lambda}_{k_1} + \boldsymbol{\lambda}_{k_2}, \text{ and}$$

$$\log \frac{\mathbb{P}_{k_1,\theta_2}(s_{k_1})}{\mathbb{P}_{k_1,\theta_1}(s_{k_1})} + \log \frac{\mathbb{P}_{k_3,\theta_2}(s_{k_3})}{\mathbb{P}_{k_3,\theta_1}(s_{k_3})} = \boldsymbol{\lambda}_{k_2} + \boldsymbol{\lambda}_{k_3}\}.$$

Here we make no attempt in optimizing the computations for the refutation process in transparent structures beyond pointing out that they can increase exponentially with the network size. In Section 5, we show that the GROUP-DECISION problem is $\mathcal{NP}$-hard; and as such there are no algorithms that will scale polynomial in network size for all network structures (unless $\mathcal{P} = \mathcal{NP}$).

When transparency is violated the neighboring beliefs are highly non-linear functions of the log-likelihoods and the forward reasoning approach of (A3: BAYES-TRANSPARENT) can no longer be applied; indeed, when transparency is violated then beliefs represent what signal profiles agents regard as feasible rather than what they know about the log-likelihoods of signals of others whom they have directly or indirectly observed. In particular, the agent cannot use the reported beliefs of the neighbors directly to make inferences about the original causes of those reports which are the private signals. Instead, to keep track of the feasible signal profiles that are consistent with her observations the agent employs a version of the IEIS algorithm of Section 3 that is tailored to the case of revealed beliefs. We describe these calculations of the Bayesian agents for nontransparent structures in Subsection 4.2.

4.2. *Belief Calculations in Nontransparent Structures*

In general nontransparent structures where one or more of the neighboring beliefs do not satisfy the transparency conditions in Definition 2, agent i would have to follow an IEIS strategy similar to (A1:BAYES-GROUP) to construct her Bayesian posterior belief given her observations of her neighbors' nontransparent beliefs. Accordingly, as in Table I, for every profile of initial signals $\overline{s} = (s_1, s_2, \ldots, s_n)$ she constructs a list of all signal profiles that she believes are deemed feasible by another agent j, given what she believes agent j may have observed up until time t conditioned on the initial signals being prescribed by $\overline{s}$. Subsequently, the information available to her at time t is the same as that in (I1: BAYES-GROUP); and she uses this information to update her list of feasible signal profiles from $\mathcal{I}_{i,t}$ to $\mathcal{I}_{i,t+1}$. Before presenting the exact calculations for determining the Bayesian posterior of agent i, note that rather than the conditionally feasible actions for each agent j, given by $a_{j,t}^{(i)}(\overline{s})$ in Table I, agent i in the case of revealed beliefs would instead keep track of $\bar{\mu}_{j,t}^{(i)}(\overline{s}) = (\mu_{j,t}^{(i)}(\overline{s}; \theta_1), \ldots, \mu_{j,t}^{(i)}(\overline{s}; \theta_m))$, i.e. the belief that she deems optimal for each agent j, given what she believes agent j could have observed up until time t conditioned on the event that the initial signals of all agents are prescribed according to $\overline{s}$. Note that following (3.3), we have:

$$\mu_{j,\tau}^{(i)}(\overline{s}; \theta_k) = \frac{\sum_{\overline{s}' \in \mathcal{I}_{j,\tau}^{(i)}(\overline{s})} \mathbb{P}_{\theta_k}(\overline{s}') \nu(\theta_k)}{\sum_{l=1}^{m} \sum_{\overline{s}' \in \mathcal{I}_{j,\tau}^{(i)}(\overline{s})} \mathbb{P}_{\theta_l}(\overline{s}') \nu(\theta_l)}.$$

Calculations of agent i at time t enables her to update her information at time t to incorporate her newly obtained data which constitute her observations of her neighbors' most recent beliefs $\bar{\mu}_{j,t}$ for all $j \in \mathcal{N}_i$; whence she refines $\mathcal{I}_{i,t}$ to $\mathcal{I}_{i,t+1}$ and updates her belief using (3.1). This is achieved as follows:

(A4: BAYES-NONTRANSPARENT). Calculations of agent i at time t for deciding her Bayesian posterior $\bar{\mu}_{i,t+1}$:

1. For all $\overline{s} := (s_1, \ldots, s_n) \in \mathcal{S}_1 \times \ldots \times \mathcal{S}_n$ and all $j \in \mathcal{N}_i^{t+1}$ do:
   - If $\delta(j,i) = t+1$, initialize $\mathcal{I}_{j,0}^{(i)}(\overline{s}) = \{s_j\} \times \prod_{k \neq j} \mathcal{S}_k$.
   - Else initialize $\mathcal{I}_{j,t+1-\delta(j,i)}^{(i)}(\overline{s}) = \mathcal{I}_{j,t-\delta(j,i)}^{(i)}(\overline{s})$ and for all $\overline{s}' \in \mathcal{I}_{j,t+1-\delta(j,i)}^{(i)}(\overline{s})$ do:
     - For all $k \in \mathcal{N}_j$ if $\bar{\mu}_{k,t-\tau}^{(i)}(\overline{s}') \neq \bar{\mu}_{k,t-\tau}^{(i)}(\overline{s})$, then $\mathcal{I}_{j,t+1-\tau}^{(i)}(\overline{s}) = \mathcal{I}_{j,t+1-\tau}^{(i)}(\overline{s}) \setminus \{\overline{s}'\}$.
2. Initialize $\mathcal{I}_{i,t+1} = \mathcal{I}_{i,t}$ and for all $\overline{s} \in \mathcal{I}_{i,t+1}$ do:
   - For all $j \in \mathcal{N}_i$ if $\bar{\mu}_{j,t} \neq \bar{\mu}_{j,t}^{(i)}(\overline{s})$, then $\mathcal{I}_{i,t+1} = \mathcal{I}_{i,t+1} \setminus \{\overline{s}\}$.
3. Given $\mathcal{I}_{i,t+1}$, calculate the updated belief $\bar{\mu}_{i,t+1}$ according to (3.1).

Following Appendix B, we know the computational complexity of (A4: BAYES - NONTRANSPARENT) increases exponentially in n and can be bounded as $O\left(n^2\ M^{2n-1}\ m\right)$. There is a key difference between the refutation process in step (2) of (A3: BAYES-TRANSPARENT), when the network is transparent to agent i, and the pruning that takes place in step (2) of the (A4 : BAYES- NONTRANSPARENT) for general (non-transparent) networks. In the latter case, the agent needs to consider the beliefs of other far way agents at any possible signal profile and to simulate the subsequent beliefs of her neighbors conditioned on the particular signal profile; cf. step (1) of (A4: BAYES-NONTRANSPARENT). Each

signal profile will be rejected and removed from the feasible set if the simulated belief of a neighbor conditioned on that signal profile does not agree with the actual (observed) beliefs at that time. On the other hand, in a transparent structure, the agent does not need to simulate the beliefs of other agents conditioned on a signal profile to investigate its feasibility; compare step (1) of (A3: BAYES-TRANSPARENT) with step (1) of (A4: BAYES-NONTRANSPARENT). She can directly verify whether the individual signals likelihoods satisfy the most recent set of constraints that are revealed to the agent at time t from the transparent beliefs of her neighbors; and if any one of the new equations is violated, then that signal profile will be rejected and removed from the feasible set. This constitutes an interesting bridge between statistical and computational efficiency in group decision processes.

5. HARDNESS OF GROUP-DECISION

In this section we prove:

THEOREM 3 (Hardness of GROUP-DECISION) *The GROUP-DECISION (Problem 1) is $\mathcal{NP}$-hard.*

We provide two reductions for the proof of Theorem 3, one reduction is to the SUBSET-SUM problem and the other reduction is to the EXACT-COVER problem. In both reductions, we use binary signal spaces for all the agents; however, the first reduction requires agents to receive different signals with non-identical probabilities whose variety increases with the network size so that we can accommodate the increasing set size. Our second reduction, on the other hand, works with i.i.d. binary signals but still relies on a complex network structure (with unbounded degrees, i.e. node degrees increase with the increasing network size) to realize arbitrary instances of EXACT-COVER.

The particular structures in which the two problems are realized are depicted in Fig. 5. The graph on the left is used for the SUBSET-SUM reduction and the graph on the right is used for the EXACT-COVER problem. The SUBSET-SUM problem asks if given a set of n positive integers $\ell_1, \ldots, \ell_n$ and another positive integer q, there is a non-empty subset of $\{\ell_1, \ldots, \ell_n\}$ that sum to $\mathcal{L}$. We encode the n parameters $\{\ell_1, \ldots, \ell_n\}$ of the SUBSET-SUM problem using the log-likelihood ratios of binary signals that the n agents in the top layer of the left graph in Fig. 5 receive. The encoding is such that $\ell_h = \overline{\ell}_h - \underline{\ell}_h$ for all $h \in [n]$, where $\overline{\ell}_h$ and $\underline{\ell}_h$ are the log-likelihood ratios of the one and zero signals for each of the n agents $w_h, h \in [n]$. Throughout this section and when working with binary signals, we use the over and under bars to indicate the log-likelihood ratios of the one and zero signals, respectively. Similarly, we denote the log-likelihood ratios of the signals of the two agents k_1 and k_2 by $\overline{\ell}^\star$ and $\underline{\ell}^\star$ and set them such that $-\mathcal{L} = \overline{\ell}^\star - \underline{\ell}^\star$. The crux of our reduction is in designing the aggregate beliefs of agents j_1 and j_2 at time one in such a way that agent i needs to decide whether the observed aggregates are caused by all of the indirectly observed agents $w_1, \ldots, w_n$ and k_1, k_2 having reported zero signals to j_1 and j_2; or else it is possible that the contributions from some of the one signals among $w_1, \ldots, w_n$ is canceled out in the aggregate by the one signals in k_1 and k_2. In the latter case, those agents, w_h, who have received one signals, $\mathbf{s}_{w_h} = 1$, constitute a feasibility certificate for the SUBSET-SUM problem, as their respective values of ℓ_h sum to $\mathcal{L}$. In Appendix E, we show that the decision problem of agent i in the designed scenario (after her observations of the beliefs of j_1 and j_2) simplifies to the feasibility of the SUBSET-SUM problem with parameters $\ell_1, \ldots, \ell_n$ and $\mathcal{L}$.

We use the right-hand side structure of Fig. 5 for the EXACT-COVER reduction. Unlike the SUBSET-SUM reduction, in the EXACT-COVER reduction we do not rely on unboundedly many types of signals. Given a set of n elements $\{j_1, \ldots, j_n\}$ and a family of m subsets $\{\mathcal{W}_1, \ldots, \mathcal{W}_m\}$, $\mathcal{W}_h \subset \{j_1, \ldots, j_n\}$ for
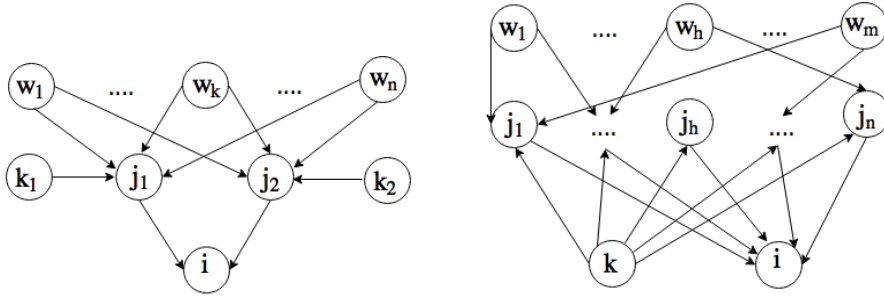
Figure 5: The graph structure on the left is used for the SUBSET-SUM reduction. The graph structure on the right is used for the EXACT-COVER reduction. In the particular instance of EXACT-COVER that is depicted on the right, we have that $j_1 \in \mathcal{W}_1$ and $j_1 \in \mathcal{W}_m$, as the links in the top two layers indicate the inclusion relations among subsets $\mathcal{W}_1, \ldots, \mathcal{W}_m$ and elements $j_1, \ldots, j_n$.

all $h \in [m]$, the EXACT-COVER problem asks if it is possible to construct a non-intersecting cover (partition) of $\{j_1, \ldots, j_n\}$ using a (disjoint) subfamily of $\{\mathcal{W}_1, \ldots, \mathcal{W}_m\}$. Given any instances of the EXACT-COVER problem, we can encode the inclusion relations between the n elements $\{j_1, \ldots, j_n\}$ and m subsets $\{\mathcal{W}_1, \ldots, \mathcal{W}_m\}$ using the bipartite graph in the first two layers of the right structure in Fig. 5. Here each node represents the respective entity (element j_r or subset $\mathcal{W}_h$, $r \in [n]$ and $h \in [m]$) of the same name: an edge from a node w_h to a node j_r for some $h \in [m]$ and $r \in [n]$ indicates that element j_r is included in the subset $\mathcal{W}_h$. Our strategy is again to take any instance of the EXACT-COVER problem and design the signal structures such that agent i's belief in the corresponding instance of GROUP-DECISION problem (with the network structure given in the right hand side of Fig. 5) would indicate her knowledge of the feasibility of the (arbitrarily chosen) instance of the EXACT-COVER problem (that is encoded by the first two layers of the right hand side graph in Fig. 5). We use $\bar{\ell}$ and $\underline{\ell}$ for the log-likelihood ratios of the one and zero signals of the $w_1, \ldots, w_m$ nodes and set these parameters such that $\bar{\ell} - \underline{\ell} = 1$. Similarly, we denote the log-likelihood ratios of the one and zero signals in the node k by $\bar{\ell}^\star$ and $\underline{\ell}^\star$, and set them such that $\bar{\ell}^\star - \underline{\ell}^\star = -1$. In Appendix F, we design a set of observations for agent i such that her belief at time 2 would require her to know whether her observations of the beliefs of $j_1, \ldots, j_n$ are caused by all agents $w_1, \ldots, w_m$ as well as agent k having received zero signals, or else whether it is possible that some of the agents among $w_1, \ldots, w_m$ have received one signals and their aggregate effects on the beliefs of $j_1, \ldots, j_n$ are canceled out by the one signal that agent k has received. The latter happens only when the corresponding instance of the EXACT-COVER problem (coded by the right hand graph of Fig. 5) is feasible. In such cases, those sets among $\mathcal{W}_1, \ldots, \mathcal{W}_m$ whose respective agents have receive one signals, $\{\mathcal{W}_h : h \in [m], \mathbf{s}_{w_h} = 1\}$, constitute a disjoint subfamily that covers $\{j_1, \ldots, j_n\}$.

The detailed reductions are presented in Appendices E and F. It is worth highlighting that our $\mathcal{NP}$-hardness reductions show that the GROUP-DECISION problem is hard to solve in the worst case. In other words, there exist network structures and particular profiles of private signals that lead to specific observations of the neighboring beliefs, such that making an inference about the observed beliefs and forming a Bayesian posterior belief conditioned on those observations is not possible in computation times that increase polynomially with the network size (unless $\mathcal{P} = \mathcal{NP}$). Alternatively, one may be interested in the complexity of computations in specific network structures with increasing size, such as trees, cycles, or complete graphs for which we know that beliefs can be computed efficiently by

virtue of their transparency. Moreover, one may also be interested in the complexity of computations in an average sense (for "typical" network structures and "typical" private signals). Deriving complexity notions in these alternative settings is much more difficult; indeed, development of such alternative notions of complexity is an active area of research in theory of computation, cf. e.g. Bogdanov et al. (2006) for average-case complexity with respect to random inputs, and cf. e.g. Motwani and Raghavan (2010) for the relevant complexity notions that apply to randomized algorithms.

REMARK 3 (Beyond $\mathcal{NP}$-hardness)   Both reductions are set up such that the feasibility of the corresponding $\mathcal{NP}$-complete problem (SUBSET-SUM or EXACT-COVER) is reflected in the time-two beliefs of agent i. However, the beliefs in both cases contain more information than the simple yes or no answer to the feasibility questions. Effectually, the information content of beliefs amounts to a weighted sum over all the feasibility certificates (each certificate is represented by a particular signal profile and is weighted by the likelihood of that particular signal profile). One possibility is to prove hardness in a class of functional problems such as $\#\mathcal{P}$. The class $\#\mathcal{P}$ is comprised of the counting problems associated with the problems in $\mathcal{NP}$. The latter is a class of decision problems for which the positive instances have an efficiently verifiable proof. While $\mathcal{NP}$ captures the difficulty of finding any certificates, the class $\#\mathcal{P}$ captures the difficulty of counting the number of all valid certificates (if any). As such the problems in $\#\mathcal{P}$ are naturally harder than those in $\mathcal{NP}$, cf. e.g. (Arora and Barak, 2009, Chapter 17).

The EXACT-COVER reduction relies critically on the fact that the number of directly observed neighbors (corresponding to the number of equations in the EXACT-COVER reduction) are allowed to increase. If the number of neighbors is fixed but different agents receive different signals with varying distributions, then our SUBSET-SUM reduction in Appendix E again verifies that the hardness property holds. Our next example shows that either of the two structural features (increasing size of the neighborhood or infinitely many types of private signals among the indirectly observed agents) are needed to obtain a hard problem; indeed, an efficient calculation of beliefs may be possible when the neighborhood sizes are kept fixed and agents receive i.i.d. private signals.

For example in the left structure of Fig. 5 we can efficiently compute the beliefs if $\{w_1, \ldots, w_n\}$ are receiving i.i.d. binary signals. To see how the belief of agent i at time two can be computed efficiently in the number of indirectly observed neighbors (n), suppose that the signal structures for agent i, her neighboring agents $\{j_1, j_2\}$, and the indirectly observed agents $\{k_1, k_2, w_1, \ldots, w_n\}$ are the same as $\{j_h, h \in [n]\}$ and $\{w_h, h \in [m]\}$ in EXACT-COVER reduction of Appendix F: $\{i, j_1, j_2\}$ receiving non-informative signals and $\{k_1, k_2, w_1, \ldots, w_n\}$ receiving i.i.d. binary signals, whose likelihoods satisfy $\boldsymbol{\lambda}_r = \mathbf{s}_r(\bar{\ell} - \underline{\ell}) + \underline{\ell}$ for all $r \in \{k_1, k_2, w_1, \ldots, w_n\}$ as in (F.1) of Appendix F. Subsequently, $\boldsymbol{\phi}_{i,0} = \boldsymbol{\phi}_{j_1,0} = \boldsymbol{\phi}_{j_2,0} = \boldsymbol{\phi}_{i,1} = 0$, due to their initial noninformative signals. At time two, agent i has to incorporate the time one beliefs of her neighbors, which are themselves caused by the time zero beliefs of $k_1, k_2, w_1, \ldots, w_n$: Given $\boldsymbol{\phi}_{j_r,1} = \boldsymbol{\lambda}_{k_r} + \sum_{h=1}^{n} \boldsymbol{\lambda}_{w_h}$, for $r = 1, 2$, agent i aims to determine her belief at time two (or equivalently $\boldsymbol{\phi}_{i,2}$). Using (F.1), we can write

$$\boldsymbol{\psi}_{j_r} = \mathbf{s}_{k_r} + \sum_{h=1}^{n} \mathbf{s}_{w_h} = \frac{1}{\bar{\ell} - \underline{\ell}} \left(\boldsymbol{\phi}_{j_r,1} - \operatorname{card}\left(\mathcal{N}_{j_r}\right) \underline{\ell} \right), r \in \{1, 2\},$$

where $\boldsymbol{\psi}_{j_r}$ are necessarily non-negative integers belonging to $[n + 1]_0 = \{0\} \cup [n + 1]$, due to their generation process; i.e. the fact that they count the number of one signals that are received in the neighborhood $\mathcal{N}_{j_r}$ of each of the neighbors $j_r, r = 1, 2$. To proceed, let $\eta \in [n]_0$ be the number of agents

among $\{w_1, \ldots, w_n\}$ who have received one signals. Depending on the relative values of $\boldsymbol{\psi}_{j_1}$ and $\boldsymbol{\psi}_{j_2}$ three cases may arise, and the subsequent beliefs $\boldsymbol{\phi}_{i,2}$ in each case are listed below:

1. If $\boldsymbol{\psi}_{j_1} \neq \boldsymbol{\psi}_{j_2}$, then exactly one of the two signals $\mathbf{s}_{k_1}$ and $\mathbf{s}_{k_2}$ is a one and the other one is zero, the latter corresponding to the lower of the two counts $\boldsymbol{\psi}_{j_1}$ and $\boldsymbol{\psi}_{j_2}$. We further have that $\eta = \min\{\boldsymbol{\psi}_{j_1}, \boldsymbol{\psi}_{j_2}\}$ and $\boldsymbol{\phi}_{i,2} = (\eta + 1)\overline{\ell} + (n - \eta + 1)\underline{\ell}$.

2. If $\boldsymbol{\psi}_{j_1} = \boldsymbol{\psi}_{j_2} = 0$, then every body in the second order neighborhood of i has received a zero and we have $\boldsymbol{\phi}_{i,2} = (n + 2)\underline{\ell}$.

3. If $\boldsymbol{\psi}_{j_1} = \boldsymbol{\psi}_{j_2} \geq 1$, then either $\mathbf{s}_{k_1} = \mathbf{s}_{k_2} = 0$ and $\eta = \eta_0 = \boldsymbol{\psi}_{j_1}$ or $\mathbf{s}_{k_1} = \mathbf{s}_{k_2} = 1$ and $\eta = \eta_1 = \boldsymbol{\psi}_{j_1} - 1$. In this case, the belief of agent i at time two is given by: $\boldsymbol{\phi}_{i,2} = \log(f_{\theta_2}/f_{\theta_1})$, where f_{θ_r}, $r = 1, 2$ is defined as: $f_{\theta_r} = \binom{n}{\eta_1}\mathbb{P}_{\theta_r}(\mathbf{s} = 1)^{\eta_1+2}\mathbb{P}_{\theta_r}(\mathbf{s} = 0)^{n-\eta_1} + \binom{m}{\eta_0}\mathbb{P}_{\theta_r}(\mathbf{s} = 1)^{\eta_0}\mathbb{P}_{\theta_r}(\mathbf{s} = 0)^{n-\eta_0+2}$.

In Appendix G, we devise a similar algorithm to calculate the time two belief of agent i in the left-hand-side structure of Fig. 5 (by counting the number of agents $h \in [m]$ for which $\mathbf{s}_{w_h} = 1$), with time-complexity $O(n2^n m^{2^n} + m^{1+2^n}(2^{2n})(3m+2))$: increasing polynomially in m for a fixed neighborhood size (n).

6. DISCUSSION, CONCLUDING REMARKS, AND FUTURE DIRECTIONS

In this paper, we analyzed recommendations of rational agents in a group decision process, as they each observe an exogenous initial private signal and are exposed to the recommendations of (a subset) of other group members in the ensuing decision epochs. Such agents in a group decision process have purely informational externalities, but they still need to interpret (and learn from) the actions of others subject to the fact that they are acting rationally. Other members' actions reveal additional information about the state of the world, which can be then used to make better future recommendations. Indeed, the actions of neighbors are informative only to the extent that they reveal information about their private signals; and as time marches on, more information is revealed about the private signals of neighbors, and neighbors of neighbors, etc. Hence, after a long enough time, all players would be (partially) informed about the private signals of all other players if the graph is connected. We analyzed the complexity of decision-making in this information structure. Iterated elimination of infeasible signals (IEIS) curbs some of the complexities of inference in group decision-making, although its running time is exponential. These computations simplify and become efficient in a POSET structure where the agent has direct access to all observations of her neighbors (except their private signals). The computations also simplify in special symmetric settings, for example with i.i.d. binary signals over a directed path, or a rooted (directed) tree (cf. Appendix D). An open problem is to investigate other configurations and structures for which the computation of Bayesian actions is achievable in polynomial-time. It is also of interest to know the quality of information aggregation; i.e. under what conditions on the signal structure and network topology, Bayesian actions coincide with the best action given the aggregate information of all agents.

In the special case that agents reveal their beliefs to each other, we introduce and analyze a structural property of the graph, referred to as transparency, which plays a critical role in characterizing the complexity of the computations when forming Bayesian posterior beliefs. Bayesian beliefs in transparent structures are both easy to compute and statically efficient; in the sense that they coincide with the Bayesian posterior of the agent, had she direct access to the private signals of all agents whom she has observed, either directly, or indirectly through their influences on her neighbors, neighbors of neighbors, etc.

We proved the $\mathcal{NP}$-hardness of the Bayesian belief exchange problem by providing reductions that show well-known $\mathcal{NP}$-complete problems such as SUBSET-SUM and EXACT-COVER are special cases

of the group decision problem. The former relies on the increasing variety of signal types and the latter relies on the increasing neighborhood size. Here we note that the nature of the two reductions in Appendices E and F leave space for strengthening the complexity class of the GROUP-DECISION problem beyond $\mathcal{NP}$, cf. Remark 3 in Section 5. Another possibility is to prove that the beliefs are even hard to approximate, by exploiting the gap that exists between the log-ratio of beliefs, depending on whether the underlying instance of the decision problem is feasible or not.

Transparency of the network structure to agent i allows her to trace the reported beliefs of her neighbors directly to their root causes which are the private signals of other agents. When transparency is violated, the neighboring beliefs are complicated highly non-linear functions of the signal likelihoods and the forward reasoning approach can no longer be applied to search for possible signals that lead to observed beliefs; indeed, when transparency is violated then the observed beliefs represent what signal profiles are regarded as feasible by the neighbors. This is quite different from the transparent case where the beliefs of neighbors directly reflect their knowledge about the likelihoods of signals that occur in the higher-order neighborhoods. In other words, in a nontransparent structure, agent i cannot use the reported beliefs of her neighbors to make direct inferences about the original causes of those reports which are the private signals in the higher-order neighborhoods. Instead, to keep track of the feasible signal profiles that are consistent with her observations agent i should consider what beliefs other agents would hold under each of the possible signal profiles and to prune the infeasible ones following an IEIS procedure. A similar observation can be made in the case of POSETs and actions: as compared with general graphs, POSETs remove the need to simulate the network at a given signal profile to reject or approve it. Instead, we can directly verify if each individual private signal agrees with the observed action of its respective agent and if it does not, then it is rejected and removed from the list of feasible private signals.

Although determining the posterior beliefs during a group decisions process is, in general, $\mathcal{NP}$-hard, for transparent structures the posterior belief at each step can be computed efficiently using the reported beliefs of the neighbors. Furthermore, the optimality of belief exchange over transparent structures is a unique structural feature of the inference set up in group decision processes. It provides an interesting and distinct addition to known optimality conditions for inference problems over graphs. In particular, the transparent structures over which efficient and optimal Bayesian belief exchange is achievable include many loopy structures in addition to trees.[1] It would be particularly interesting if one can provide a tight graphical characterization for transparency or provide other useful sufficient conditions that ensure transparency and complement our Proposition 2. More importantly, one would look for other characterizations of the complexity landscape and find other notions of simplicity that are different from transparency.

---

[1]It is well known that if a Bayesian network has a tree (singly connected or polytree) structure, then efficient inference can be achieved using belief propagation (message passing or sum-product algorithms), cf. Pearl (2014). However, in general loopy structures, belief propagation only gives a (potentially useful) approximation of the desired posteriors, cf. Murphy et al. (1999). Notwithstanding, our Bayesian belief exchange set up also greatly simplifies in the case of tree structures, admitting a trivial sum of innovations algorithm. Chandrasekaran et al. (2012) study the complexity landscape of inference problems over graphical models in terms of their treewidth. For bounded treewidth structures the junction-tree method (performing belief propagation on the tree decomposition of the graph) works efficiently (Cowell et al. (2006)) but there is no class of graphical models with unbounded treewidth in which inference can be performed in time polynomial in treewidth.

APPENDIX A: CONVERGENCE TO EQUILIBRIUM AND CONSENSUS IN SYMMETRIC GROUPS

Following Molavi et al. (2015) and Rosenberg et al. (2009), the asymptotic outcome of Bayesian group decision can be characterized as a Markov perfect Bayesian equilibrium in a repeated game of incomplete information that is played by successive generations of short-lived players. Short-lived agents inherit the beliefs of the player playing in the previous stage in their role while also observing the last stage actions of the players in their social neighborhood. Rational myopia arise by nature of short-lived agents, and the equilibrium concept can be used to study the rational myopic decisions, subject to the assumption that other players are also playing their myopic best responses given their own information. Markov perfect Bayesian equilibrium is the appropriate solution concept for the study of Bayesian group decision process because a Markovian strategy for agent i can depend on the information available to her, $\mathbf{h}_{i,t}$, in her role as agent i at time t, only to extent that $\mathbf{h}_{i,t}$ is informative about θ, the pay-off relevant state of the world. Molavi et al. (2015) provide the following recursive construction of Markovian strategy profiles: consider the probability triplet $(\Omega, \mathscr{B}, \mathbb{P})$, where $\Omega = \Theta \times \prod_{i \in [n]} \mathcal{S}_i$, $\mathscr{B}$ is the Borel sigma algebra, and $\mathbb{P}$ assigns probabilities to the events in $\mathscr{B}$ consistently with the common prior ν and the product of the signal likelihoods; for each i, let $\boldsymbol{\sigma}_{i,0} : \Omega \to \mathcal{A}_i$ be a measurable map defined on $(\Omega, \mathscr{B}, \mathbb{P})$ that specifies the time zero action of agent i as a function of her private signal, and let $\mathscr{H}_{i,0}$ denote the information available to agent i at time zero which is the smallest sub-sigma algebra of $\mathscr{B}$ that makes $\mathbf{s}_i$ measurable. Then for any time t, we can define a Markovian strategy $\boldsymbol{\sigma}_{i,t}$,recursively, as a random variable which is measurable with respect to $\mathscr{H}_{i,t}^{\overline{\sigma}_{t-1}}$, where $\overline{\sigma}_{t-1} = (\overline{\sigma}_{1,t-1}, \ldots, \overline{\sigma}_{n,t-1})$, $\overline{\sigma}_{i,t-1} = (\sigma_{i,0}, \ldots, \sigma_{i,t-1})$ for all i, and $\mathscr{H}_{i,t}^{\overline{\sigma}_{t-1}}$ is the smallest sub-sigma algebra of $\mathscr{B}$ that makes $\mathbf{s}_i$ and $\overline{\sigma}_{j,t-1}, j \in \mathcal{N}_i$ measurable. The contributions of Molavi et al. (2015) and Rosenberg et al. (2009) consist of proving convergence to an equilibrium profile $\overline{\sigma}_\infty$ and showing consensus properties for the equilibrium profile, the former (convergence result) relies on the compactness of the action space, while the latter (asymptotic consensus result) replies on an imitation principle argument that works for common (symmetric among the agents) utility and action structures.[1] Both results rely on some analytical properties of the utility function as well, such as supermodularity[2]T in Molavi et al. (2015) or continuity (where the action spaces are metric compact spaces) and boundedness between L_2 integrable functions in Rosenberg et al. (2009). Other works have looked at different asymptotics; in particular, information aggregation and learning as the number of agents grows Mossel et al. (2014, 2015). In this work we are interested in the computations that are required of a Bayesian agent in order for her to achieve her optimal recommendations at every (finite) step during a group decision process, rather than the asymptotic and equilibrium properties of such recommendations.

---

[1]In symmetric groups all agents have the same action space $\mathcal{A}_i = \mathcal{A}_j$ for all i, j and identical utility functions $u_i(a, \theta) = u_j(a, \theta)$ for all $a \in \mathcal{A}$ and any $\theta \in \Theta$. Symmetric settings arise very naturally in group-decision scenarios where people have similar preferences about the group-decision outcome and seek the same truth or a common goal. In such scenarios, the question of consensus or unanimity is of particular importance, as it gives a sharp prediction about the group decision outcome and emergence of agreement among individual decision makers.

[2]In general, the supermodularity of the utilities signifies strategic complementarity between the actions of the players, as is the case for Molavi et al. (2015); however, in the absence of strategic externalities (as is the case for group decisions) super-modularity implies a case of diminishing returns: $u_i(\cdot, \cdot)$ is strictly supermodular iff $u_i(\min\{a, a'\}, \theta) + u + i(\max\{a, a'\}, \theta) > u_i(a, \theta) + u_i(a', \theta)$, for all $a \neq a' \in \mathcal{A}_i$ and each $\theta \in \Theta$.

APPENDIX B: COMPLEXITY OF BAYESIAN DECISIONS USING (A1: BAYES-GROUP)

Suppose that agent i has reached her t-th decision epoch in a general network structure. Given her information (I1) at time t, for all $\bar{s} = (s_1, \ldots, s_n) \in \mathcal{S}_1 \times \ldots \times \mathcal{S}_n$ and any $j \in \bar{\mathcal{N}}_i^{t+1}$ she has to update $\mathcal{I}_{j,t-\delta(j,i)}^{(i)}(\bar{s})$ into $\mathcal{I}_{j,t+1-\delta(j,i)}^{(i)}(\bar{s}) \subset \mathcal{I}_{j,t-\delta(j,i)}^{(i)}(\bar{s})$. If $\delta(j,i) = t+1$ then agent j is being considered for the first time at the t-th decision epoch and $\mathcal{I}_{j,0}^{(i)}(\bar{s}) = \{s_j\} \times \prod_{k \neq j} \mathcal{S}_k$ is initialized without any calculations. However if $\delta(j,i) \leq t$, then $\mathcal{I}_{j,t-\delta(j,i)}^{(i)}(\bar{s})$ can be updated into $\mathcal{I}_{j,t+1-\delta(j,i)}^{(i)}(\bar{s}) \subset \mathcal{I}_{j,t-\delta(j,i)}(\bar{s})$ only by verifying the condition $a_{k,t-\delta(j,i)}^{(i)}(\bar{s}') = a_{k,t-\delta(j,i)}^{(i)}(\bar{s})$ for every $\bar{s}' \in \mathcal{I}_{j,t-\delta(j,i)}^{(i)}(\bar{s})$ and $k \in \mathcal{N}_j$: any $\bar{s}' \in \mathcal{I}_{j,t-\delta(j,i)}^{(i)}(\bar{s})$ that violates this condition for some $k \in \mathcal{N}_j$ is eliminated and $\mathcal{I}_{j,t+1-\delta(j,i)}^{(i)}(\bar{s})$ is thus obtained by pruning $\mathcal{I}_{j,t-\delta(j,i)}^{(i)}(\bar{s})$.

Verification of $a_{k,t-\delta(j,i)}^{(i)}(\bar{s}') = a_{k,t-\delta(j,i)}^{(i)}(\bar{s})$ involves calculations of $a_{k,t-\delta(j,i)}^{(i)}(\bar{s}')$ and $a_{k,t-\delta(j,i)}^{(i)}(\bar{s})$ according to (3.3). The latter requires the addition of $\text{card}(\mathcal{I}_{k,t-\delta(j,i)}^{(i)}(\bar{s}))$ product terms $u_k(a_k, \theta') \, \mathbb{P}_{\theta'}(\bar{s}') \, \nu(\theta') = u_k(a_k, \theta') \, \mathbb{P}_{1,\theta'}(s_1') \ldots \mathbb{P}_{n,\theta'}(s_n') \, \nu(\theta')$ for each $\bar{s}' \in \mathcal{I}_{k,t-\delta(j,i)}^{(i)}(\bar{s})$, $\theta' \in \Theta$, and $a_k \in \mathcal{A}_k$ to evaluate the left hand-side of (3.3). Hence, we can estimate the total number of additions and multiplications required for calculation of each (conditionally) feasible action $a_{k,t-\delta(j,i)}^{(i)}(\bar{s})$ as $A \cdot (n+2) \cdot m \cdot \text{card}(\mathcal{I}_{k,t-\delta(j,i)}^{(i)}(\bar{s}))$, where $m := \text{card}(\Theta)$ and $A = \max_k \in [n]\text{card}(\mathcal{A}_k)$. Hence the total number of additions and multiplications undertaken by agent i at time t for determining actions $a_{k,t-\delta(j,i)}^{(i)}(\bar{s})$ can be estimated as follows:

$$(B.1) \qquad A \cdot (n+2) \cdot \text{card}(\Theta) \cdot \sum_{j:\delta(j,i) \leq t,} \sum_{k \in \mathcal{N}_j} \text{card}(\mathcal{I}_{k,t-\delta(j,i)}(\bar{s})) \leq A \cdot (n+2) \cdot n \cdot M^{n-1} \cdot m,$$

where we upper-bound the cardinality of the union of the higher-order neighborhoods of agent i by the total number of agents: $\text{card}(\bar{\mathcal{N}}_i^{t+1}) \leq n$ and use the inclusion relationship $\mathcal{I}_{k,t-\delta(j,i)}^{(i)}(\bar{s}) \subset \mathcal{I}_{k,0}^{(i)}(\bar{s}) = \{s_k\} \times \prod_{j \neq k} S_j$ to upper-bound $\text{card}(\mathcal{I}_{k,t-\delta(j,i)}^{(i)}(\bar{s}))$ by M^{n-1} where M is the largest cardinality of finite signal spaces, $S_j, j \in [n]$. As the above calculations are performed at every signal profile $\bar{s} \in \mathcal{S}_1 \times \ldots \mathcal{S}_n$ the total number of calculations (additions and multiplications) required for the Bayesian decision at time t can be bounded as follows:

$$(B.2) \qquad A \cdot M^n \leq A \cdot C_1 \leq (n+2) \cdot n \cdot M^{2n-1} \cdot m,$$

where we apply (B.1) for the right-hand side. In particular, the calculations grow exponential in the number of agents n. Once agent i calculates the action sets $a_{k,t-\delta(j,i)}^{(i)}(\bar{s})$ for all $k \in \mathcal{N}_j$ with $\delta(j,i) \leq t$ she can then update the feasible signal profiles $\mathcal{I}_{j,t-\delta(j,i)}^{(i)}(\bar{s})$, following step 1 of (A1), to obtain $\mathcal{I}_{j,t+1-\delta(j,i)}^{(i)}(\bar{s})$ for all $j : \delta(j,i) \leq t+1$ and any $\bar{s} \in \mathcal{S}_1 \times \ldots \times \mathcal{S}_n$. This in turn enables her to calculate the conditional actions of her neighbors $a_{j,t}^{(i)}(\bar{s})$ at every signal profile and to eliminate any $\bar{s}$ for which the conditionally feasible action set $a_{j,t}^{(i)}(\bar{s})$ does not agree with the observed action $\mathbf{a}_{j,t}$ for some $j \in \mathcal{N}_i$. She can thus update her list of feasible signal profiles from $\mathcal{I}_{i,t}$ to $\mathcal{I}_{i,t+1}$ and adopt the corresponding Bayesian belief $\boldsymbol{\mu}_{i,t+1}$ and action $\mathbf{a}_{i,t+1}$. The latter involves an additional $(n+2)mA\text{card}(\mathcal{I}_{i,t+1})$ additions and multiplication which are nonetheless dominated by the number calculations required in (B.2) for the simulation of other agents' actions at every signal profile. $Q.E.D.$

APPENDIX C: COMPUTATIONAL COMPLEXITY OF (A2:BAYES-POSET)

According to (I2), in a POSET structure at time t agent i has access to the list of feasible private signals for each of her neighbors: $\boldsymbol{\mathcal{S}}_{j,t}$, $j \in \mathcal{N}_i$ given their observations up until that point in time. The feasible signal set for each agent $j \in \mathcal{N}_i$ is calculated based on the actions taken by others and observed by agent j until time $t - 1$ together with possible private signals that can explain her history of choices: $\mathbf{a}_{j,0}$, $\mathbf{a}_{j,1}$, and so on up until her most recent choice which is $\mathbf{a}_{j,t}$. At time t, agent i will have access to all the observations of every agent in her neighborhood and can vet their most recent choices $\mathbf{a}_{j,t}$ against their observations to eliminate the incompatible private signals from the feasible set $\boldsymbol{\mathcal{S}}_{j,t}$ and obtain an updated list of feasible signals $\boldsymbol{\mathcal{S}}_{j,t+1}$ for each of her neighbors $j \in \mathcal{N}_i$. This pruning is achieved by calculating $\mathbf{a}_{j,t}(s_j)$ given $\boldsymbol{\mathcal{I}}_{j,t}(s_j) = \{s_j\} \times \prod_{k \in \mathcal{N}_j} \boldsymbol{\mathcal{S}}_{j,t}$ for each $s_j \in \boldsymbol{\mathcal{S}}_{j,t}$ and removing any incompatible s_j that violates the condition $\mathbf{a}_{j,t} = \mathbf{a}_{j,t}(s_j)$; thus obtaining the pruned set $\boldsymbol{\mathcal{S}}_{j,t+1}$. The calculation of $\mathbf{a}_{j,t}(s_j)$ given $\boldsymbol{\mathcal{I}}_{j,t}(s_j) = \{s_j\} \times \prod_{k \in \mathcal{N}_j} \boldsymbol{\mathcal{S}}_{j,t}$ is performed according to (3.3) but the decomposition of the feasible signal profiles based on the relation $\boldsymbol{\mathcal{I}}_{j,t}(s_j) = \{s_j\} \times \prod_{k \in \mathcal{N}_j} \boldsymbol{\mathcal{S}}_{j,t}$ together with the independence of private signals across different agents help reduce the number of additions and multiplications involved as follows:

$$
\begin{aligned}
\mathbf{a}_{j,t}(s_j) &= \arg \max_{a_j \in \mathcal{A}_j} \sum_{\theta' \in \Theta} u_j(a_j, \theta') \frac{\sum_{\overline{s}' \in \boldsymbol{\mathcal{I}}_{i,t}(s_j)} \mathbb{P}_{\theta'}(\overline{s}') \nu(\theta')}{\sum_{\theta'' \in \Theta} \sum_{\overline{s}' \in \boldsymbol{\mathcal{I}}_{i,t}(s_j)} \mathbb{P}_{\theta''}(\overline{s}') \nu(\theta'')} \\
&= \arg \max_{a_j \in \mathcal{A}_j} \sum_{\theta' \in \Theta} u_j(a_j, \theta') \frac{\mathbb{P}_{\theta'}(s_j) \prod_{k \in \mathcal{N}_j} \sum_{s_k \in \boldsymbol{\mathcal{S}}_{k,t}} \mathbb{P}_{\theta'}(s_k) \nu(\theta')}{\sum_{\theta'' \in \Theta} \mathbb{P}_{\theta''}(s_j) \prod_{k \in \mathcal{N}_j} \sum_{s_k \in \boldsymbol{\mathcal{S}}_{k,t}} \mathbb{P}_{\theta''}(s_k) \nu(\theta'')}.
\end{aligned}
$$

Hence, the calculation of the conditionally feasible action $\mathbf{a}_{j,t}(s_j)$ for each $s_j \in \boldsymbol{\mathcal{S}}_{j,t}$ can be achieved through $\mathrm{card}(\Theta)$ $A \sum_{k \in \mathcal{N}_j} \mathrm{card}(\boldsymbol{\mathcal{S}}_{k,t})$ additions and $\mathrm{card}(\Theta) (\mathrm{card}(\mathcal{N}_j) + 2) A$ multiplications; subsequently, the total number of additions and multiplications required for agent i to update the feasible private signals of each of her neighbor can be estimated as follows:

$$
\text{(C.1)} \quad A \sum_{j \in \mathcal{N}_i} \mathrm{card}(\Theta) \mathrm{card}(\boldsymbol{\mathcal{S}}_{j,t}) \left[\sum_{k \in \mathcal{N}_j} \mathrm{card}(\boldsymbol{\mathcal{S}}_{k,t}) + \mathrm{card}(\mathcal{N}_j) + 2 \right] \leq A n^2 M^2 m + A n^2 M m + 2 n M m A,
$$

where M, n, m and A are as in (B.2). After updating her lists for the feasible signal profiles of all her agents the agent can refine her list of feasible signal profiles $\boldsymbol{\mathcal{I}}_{i,t+1} = \{\mathbf{s}_i\} \times \prod_{j \in \mathcal{N}_i} \boldsymbol{\mathcal{S}}_{j,t+1}$ and determine her belief $\boldsymbol{\mu}_{i,t+1}$ and refined choice $\mathbf{a}_{i,t+1}$. The latter is achieved through an extra $\mathrm{card}(\Theta) A \sum_{j \in \mathcal{N}_i} \mathrm{card}(\boldsymbol{\mathcal{S}}_{j,t+1})$ additions and $\mathrm{card}(\Theta) A (\mathrm{card}(\mathcal{N}_i) + 2)$ multiplications, which are dominated by the required calculations in (C.1). Most notably, the computations required of the agent for determining her Bayesian choices in a POSET increase polynomially in the number of agents n, whereas in a general network structure using (A1) these computations increase exponentially fast in the number of agents n. $\hfill$ Q.E.D.

APPENDIX D: EFFICIENT ALGORITHMS FOR SYMMETRIC BINARY ENVIRONMENTS

Suppose that the agents are in a binary environment with two states: $\theta_1 = 0$ and $\theta_2 = 1$ and uniform priors $\nu(\theta_1) = \nu(\theta_2) = 1/2$. They receive i.i.d. binary initial signals $\mathbf{s}_i \in \{0, 1\}$, $i \in [n]$, such that for some $p > 1/2$ fixed we have $\mathbb{P}_{i,\theta}(\mathbf{s}_i) = p$ if $\mathbf{s}_i = \theta$ and $\mathbb{P}_{i,\theta}(\mathbf{s}_i) = 1 - p$, otherwise. Since, $p > 1/2$, at

time zero all agents will act based on their signals by simply choosing $\mathbf{a}_{i,0} = \mathbf{s}_i$. At time two, agent i gets to learn the private signals of her neighbors from their time zero actions and therefore takes the action that indicates the majority over the signals observed in by her and everybody in her immediate neighborhood. Since the signals are i.i.d. the agent could be indifferent between her actions, thus in the sequel, we assume that the agent sticks with own signal whenever her information makes her indifferent between her actions. This assumption may seem natural and harmless but in fact, it leads to drastically different behaviors in the case of a directed path. Consider, the directed path of length four in Fig. 3, on the left. At time one, if any agent observes an action that contradicts own, then she will be indifferent between the two actions. If we assume that agents resolve their indifference by reverting to their private signals then no agent will ever shift her initial action in a directed path and we have $\mathbf{a}_{i,t} = \mathbf{s}_i$ for all t and i. However, a starkly different outcome will emerge if we instead assume that the agents will always shift their actions whenever they are indifferent. In particular, for a directed path we get that at any time the agents will take the optimal action given the initial signals of everybody in their t-radius ego-net (i.e. *prefect information aggregation*): $\mathbf{a}_{i,t} \in \arg\max_{x \in \{0,1\}} \sum_{j \in \bar{\mathcal{N}}_i^t} \mathbb{1}\{\mathbf{s}_j = x\}$, where we use the indicator function notation: $\mathbb{1}\{\mathscr{P}\}$ is one if $\mathscr{P}$ is true, and it is zero otherwise. At time t, $\bar{\mathcal{N}}_i^t$ is the set of all agents who directly or indirectly influence the decisions of agent i, and perfect aggregation ensures that the action of agent i at time t coincides with her optimal action if she had given direct access to all the signals of all agents who have influenced her decision directly or indirectly; hence, the name "*perfect aggregation*". We can verify that perfect aggregation holds true in any directed path by induction; in particular, consider agent one in the left graph of Fig. 3: she gets to know about the private signal of agent two at time one, after observing $\mathbf{a}_{2,0} = \mathbf{s}_2$, next at time two, she observes $\mathbf{a}_{2,1} = \mathbf{a}_{3,0} = \mathbf{s}_3$ (note that if agents switch actions when indifferent then, at time one in a directed path all agents will replicate their neighbor's time zero actions); hence, agent one learns the private signal of agent three at time two leading her to take the majority over all three signals $\{\mathbf{s}_1, \mathbf{s}_2 \mathbf{s}_3\}$. By the same token at time three agent one observes $\mathbf{a}_{2,2}$, which is the majority over $\{\mathbf{s}_2, \mathbf{s}_3, \mathbf{s}_4\}$ and having learned $\mathbf{s}_2$ and $\mathbf{s}_3$ from her observation at previous time steps, she can now infer the value of $\mathbf{s}_4$ and thus take the majority over all the private signals $\{\mathbf{s}_1, \mathbf{s}_2, \mathbf{s}_3, \mathbf{s}_4\}$, achieving the global optimum at time three. More generally, in any directed path, at time t agent i learns the values of the private signals of the agent at distance t from her. She achieves this by combining two pieces of information: (i) her knowledge at time $t-1$, which constitutes the values of the signals of all agent at distance $t-1$ from her, (ii) her observation at time t, which constitutes her neighbor's action at time $t-1$ and is equal to the majority over all signals within distance t of agent i, excluding herself. Put more succinctly, suppose $i > t$ and let the agents be labeled in accordance with the topological ordering of the directed path, then knowing the values of the all $t-1$ preceding signals and also the majority over all t preceding signals, agent i can learn the value of the t-th signal at time t.

Therefore, switching actions or staying with own past actions when indifferent, makes the difference between no aggregation at all ($\mathbf{a}_{i,t} = \mathbf{s}_i$ for all t , i) and perfect aggregation in the case of a directed path; indeed, by switching their actions at time one, after observing their neighbor's time zero action (or equivalently private signal) the agent can pass along her information about her neighbor's signal to the person who is observing her. In the case of a directed path, this indirect signaling is enough to ensure perfect aggregation. The exact same argument can be applied to the case of a rooted ordered tree, cf. Fig. 6 on the left; in such a structure the set of all agents who influence the actions of some agent i always constitute a directed path that starts with the root node and ends at the particular agent i. As such when computing her Bayesian actions in a rooted ordered tree agent i need only consider the

unique path that connects her to the root node; thus reducing her computations to those of an agent in a directed path.

In a general structure (beyond rooted trees and symmetric binary environments), perfect aggregation can be defined as:

$$\mathbf{a}_{i,t} \in \arg \max_{a_i \in \mathcal{A}_i} \frac{\sum_{\theta' \in \Theta} u_i(a_i, \theta') \prod_{j \in \bar{\mathcal{N}}_i^t} \mathbb{P}_{j,\theta'}(\mathbf{s}_j) \nu(\theta')}{\sum_{\theta'' \in \Theta} \prod_{j \in \bar{\mathcal{N}}_i^t} \mathbb{P}_{j,\theta''}(\mathbf{s}_j) \nu(\theta'')}, \text{ for all } i,t.$$

While many asymptotic characterizations are available for the efficiency of the group decision equilibrium outcomes (cf. Appendix A), deriving tight conditions that ensures perfect aggregation in general structures is a significant open problem. Our focus in this paper is on the computations of the Bayesian agent; hence we address the efficiency of information aggregation only to the extent that it relates to the computations of the Bayesian agent. In particular, when investigating the complexity of Bayesian belief exchange in Section 4, we introduce and study a graph property, called "transparency", that ensures perfect aggregation for beliefs.

We end our discussion of the symmetric binary environment by considering an oriented tree of depth three and focusing on the actions of the root node, call her agent i (cf. Fig. 6, on the right). At time zero all agents report their initial signals as their actions $\mathbf{a}_{i,0} = \mathbf{s}_i$; having learned her neighbors' private signals, at time one each agent takes a majority over all the signals in her immediate neighborhood (including own signal). Indeed, this is true for any graph structure in a symmetric binary environment that $\mathbf{a}_{i,0} = \mathbf{s}_i$ and $\mathbf{a}_{i,1} \in \arg \max_{x \in \{0,1\}} \sum_{j \in \bar{\mathcal{N}}_i^1} \mathbb{1}\{\mathbf{s}_j = x\}$ for all i. At time two, agent i is informed about the time-one actions of her neighbors which gives her the majority values over each of their respective local neighborhoods $\bar{\mathcal{N}}_j^1, j \in \mathcal{N}_i$. In a (singly connected) tree structure these neighborhoods are non-intersecting; hence, agent i can form a refined belief ratio at time two, by summing over all (mutually exclusive) signal profiles that lead to each of the observed majority values in each local neighborhood $\bar{\mathcal{N}}_j^1, j \in \mathcal{N}_i$ and then form their product, using the fact that signals are generated independently across the non-intersecting neighborhoods:

$$\frac{\boldsymbol{\mu}_{i,2}(0)}{\boldsymbol{\mu}_{i,2}(1)} = \frac{p^{1-\mathbf{s}_i}(1-p)^{\mathbf{s}_i} \prod_{j \in \mathcal{N}_i} p^{1-\mathbf{a}_{j,0}}(1-p)^{\mathbf{a}_{j,0}} f_p^{\mathbf{a}_{j,1}}(\lfloor d_j/2 \rfloor + 1, d_j)^{|\mathbf{a}_{j,1}-\mathbf{a}_{j,0}|} f_p^{\mathbf{a}_{j,1}}(\lceil d_j/2 \rceil, d_j)^{1-|\mathbf{a}_{j,1}-\mathbf{a}_{j,0}|}}{(1-p)^{1-\mathbf{s}_i}p^{\mathbf{s}_i} \prod_{j \in \mathcal{N}_i} (1-p)^{1-\mathbf{a}_{j,0}} p^{\mathbf{a}_{j,0}} f_{1-p}^{\mathbf{a}_{j,1}}(\lfloor d_j/2 \rfloor + 1, d_j)^{|\mathbf{a}_{j,1}-\mathbf{a}_{j,0}|} f_{1-p}^{\mathbf{a}_{j,1}}(\lceil d_j/2 \rceil, d_j)^{1-|\mathbf{a}_{j,1}-\mathbf{a}_{j,0}|}},$$

where we use $d_j := \text{card}(\mathcal{N}_j)$ and for non-negative integers x, y and $0 < p < 1$ we define:

$$(\text{D.1}) \quad f_p^a(x,y) = \sum_{\eta_j = x}^{y} \binom{y}{\eta_j} p^{\eta_j(1-a)+(d_j-\eta_j)a}(1-p)^{\eta_j a+(d_j-\eta_j)(1-a)},$$

where $\lfloor d_j/2 \rfloor$ and $\lceil d_j/2 \rceil$ are respectively, the greatest integer less than or equal to $d_j/2$, and the least integer greater than or equal to $d_j/2$. Note that the summations in (D.1) are over the set of signal profiles that agent i deems feasible for each of the disjoint neighborhoods $\mathcal{N}_j, j \in \mathcal{N}_i$. Computation of these summations and their use in the belief ratio $\boldsymbol{\mu}_{i,2}(0)/\boldsymbol{\mu}_{i,2}(1)$ are simplified by fixing the majority population η_j in each neighborhood $\mathcal{N}_j$: $\lfloor d_j/2 \rfloor + 1 \leq \eta_j \leq d_j$ if $\mathbf{a}_{j,1} \neq \mathbf{a}_{j,0}$ and $\lceil d_j/2 \rceil \leq \eta_j \leq d_j$ if $\mathbf{a}_{j,1} = \mathbf{a}_{j,0}$; then using the binomial coefficients to count the number of choices to form the fixed majority population η_j out of the total neighborhood size $d_j = \text{card}(\mathcal{N}_j)$. Given $\boldsymbol{\mu}_{i,2}(0)/\boldsymbol{\mu}_{i,2}(1)$, agent i can take actions as follows: $\mathbf{a}_{i,2} = 1$ if $\boldsymbol{\mu}_{i,2}(0)/\boldsymbol{\mu}_{i,2}(1) < 1$, $\mathbf{a}_{i,2} = 0$ if $\boldsymbol{\mu}_{i,2}(0)/\boldsymbol{\mu}_{i,2}(1) > 1$, and $\mathbf{a}_{i,2} = 1 - \mathbf{a}_{i,1}$ if $\boldsymbol{\mu}_{i,2}(0)/\boldsymbol{\mu}_{i,2}(1) = 1$.

We end this appendix by expanding on the above realization that in a path of length n, every agent $n - t$ gets fixed in decisions after time t; and in particular, no agents will change their recommendations
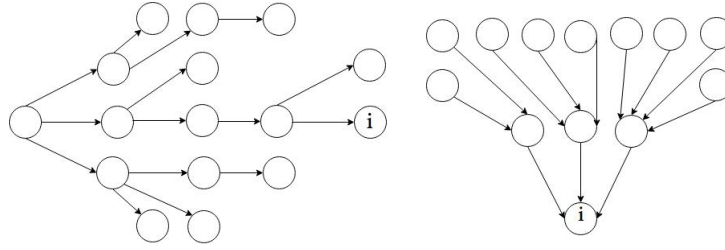
Figure 6: On the left, the Bayesian computations of agent i in a rooted ordered tree reduces to those in the unique path connecting her to the root (the leftmost node); On the right, an oriented (singly connected) tree with depth three.

after $t \geq n-1$. There is an easy inductive proof upon noting that indeed agent n, who is a leaf node with no access to the recommendations of anybody else, will never change initial action. Moreover, if agent $n-t+1$ fixes her decision at time $t-1$, then agent $n-t$ would necessarily fix her decision at time t as she receives no new information following her observation of $\mathbf{a}_{n-t+1,t-1}$. Proposition 1 extends this finite time convergence property of paths to more general structures where a "strict" partial order can be imposed on the set of agents, and in such a way that this order respects the neighborhood relationships among the agents. The strictness property restricts our method to structures without loops or bidirectional links, which are widely known as directed acyclic graphs (DAGs), cf. e.g. Bang-Jensen and Gutin (2008).

APPENDIX E: A SUBSET-SUM REDUCTION

The SUBSET-SUM problem can be described as follows and it is know to be $\mathcal{NP}$-complete (cf. (Garey and Johnson, 1979, A3.2, SP13)):

PROBLEM 2 (SUBSET-SUM) *Given a set of n positive integers $\{\ell_1, \ldots, \ell_n\}$ and a positive integer $\mathcal{L}$, determine if any non-empty subset of $\{\ell_1, \ldots, \ell_n\}$ sum to $\mathcal{L}$.*

We now describe the reduction to an arbitrary instance of SUBSET-SUM from a particular instance of GROUP-DECISION. Consider the problem of determining the belief of agent i at time 2, $\boldsymbol{\mu}_{i,2}$, in the graph $\mathcal{G}$ with $n+5$ nodes and $2n+4$ edges as in the left graph in Fig. 5: agent i have two neighbors j_1 and j_2, who themselves have n neighbors in common $w_1, \ldots, w_n$. Furthermore, j_1 and j_2 each has one additional neighbor, k_1 and k_2 respectively, whom they do not share. We take the signal spaces of i, j_1, and j_2 to be singletons $\mathcal{S}_i = \mathcal{S}_{j_1} = \mathcal{S}_{j_2} = \{\mathring{s}\}$, so that their private signals reveal no information and as such $\phi_{i,0} = \phi_{j_1,0} = \phi_{j_2,0} = \phi_{i,1} = 0$, following the simplifying assumptions of the binary state space with common uniform priors. We assume that each of the remaining agents $w_1, w_2, \ldots, w_n$ have a binary signal space $\mathbf{s}_{w_h} \in \{0,1\}$, with the probabilities that are set such that

$$(\text{E.1}) \quad \underline{\ell}_h := \log\left(\frac{\mathbb{P}_{\theta_2}(\mathbf{s}_{w_h} = 0)}{\mathbb{P}_{\theta_1}(\mathbf{s}_{w_h} = 0)}\right), \overline{\ell}_h := \log\left(\frac{\mathbb{P}_{\theta_2}(\mathbf{s}_{w_h} = 0)}{\mathbb{P}_{\theta_1}(\mathbf{s}_{w_h} = 0)}\right) , \ell_h = \overline{\ell}_h - \underline{\ell}_h, \text{ for all } h \in [n].$$

As for the agents k_1 and k_2, they also receive binary signals but with probabilities that are set such that for $r = 1, 2$:

$$(\text{E.2}) \quad \underline{\ell}^\star := \log\left(\frac{\mathbb{P}_{\theta_2}(\mathbf{s}_{k_r} = 0)}{\mathbb{P}_{\theta_1}(\mathbf{s}_{k_r} = 0)}\right) , \overline{\ell}^\star := \log\left(\frac{\mathbb{P}_{\theta_2}(\mathbf{s}_{k_r} = 1)}{\mathbb{P}_{\theta_1}(\mathbf{s}_{k_r} = 1)}\right) , -\mathcal{L} = \overline{\ell}^\star - \underline{\ell}^\star.$$

Suppose further that at time 2, agent i observes the beliefs of both agents j_1 and j_2 to be as follows: $\phi_{j_1,1} = \phi_{j_2,1} = \sum_{h=1}^{m} \underline{\ell}_h + \underline{\ell}^\star$. Note that in the above notation we have

$$(\text{E.3}) \qquad \boldsymbol{\lambda}_{k_r} = \mathbf{s}_{k_r} \underbrace{(\overline{\ell}^\star - \underline{\ell}^\star)}_{-\mathcal{L}} + \underline{\ell}^\star, \text{ and } \boldsymbol{\lambda}_{w_h} = \mathbf{s}_{w_h} \underbrace{(\overline{\ell}_{w_h} - \underline{\ell}_{w_h})}_{\ell_h} + \underline{\ell}_{w_h}, r = 1, 2, h \in [n].$$

These quantities are important as they determine the beliefs of agents j_1 and j_2 at time one, which are reported to agent i for processing her belief update at time 2. In particular, at time 2, and from the fact that $\phi_{j_1,1} = \phi_{j_2,1} = \sum_{h=1}^{m} \underline{\ell}_h + \underline{\ell}^\star$ agent i infers the following information:

$$\phi_{j_1,1} = \sum_{h=1}^{n} \boldsymbol{\lambda}_{w_h} + \boldsymbol{\lambda}_{k_1} = \sum_{h=1}^{m} \underline{\ell}_h + \underline{\ell}^\star, \text{ and } \phi_{j_2,1} = \sum_{h=1}^{n} \boldsymbol{\lambda}_{w_h} + \boldsymbol{\lambda}_{k_2} = \sum_{h=1}^{m} \underline{\ell}_h + \underline{\ell}^\star.$$

Replacing from (E.1), (E.2) and (E.3), the preceding relations can be written in terms of the private signals $\mathbf{s}_{w_h}$, $h \in [n]$ and $\mathbf{s}_{k_1}, \mathbf{s}_{k_1}$ as follows:

$$(\text{E.4}) \qquad \sum_{h=1}^{n} \mathbf{s}_{w_h} \ell_h - \mathbf{s}_{k_1} \mathcal{L} = 0, \text{ and } \sum_{h=1}^{n} \mathbf{s}_{w_h} \ell_h - \mathbf{s}_{k_2} \mathcal{L} = 0.$$

Note that the constant term $\sum_{h=1}^{m} \underline{\ell}_h + \underline{\ell}^\star$ is canceled out from both sides of the two equations leading to the homogeneous system in (E.4). To compute her Bayesian posterior belief $\boldsymbol{\mu}_{i,2}$ or equivalently $\phi_{i,2}$, agent i should first solve the arbitrary instance of SUBSET-SUM for the given parameters: SUBSET-SUM$(\ell_1, \ldots, \ell_n; \mathcal{L})$. If she determines that the answer to SUBSET-SUM$(\ell_1, \ldots, \ell_n; \mathcal{L})$ is negative then she concludes that all agents must have received zero signals and she sets her belief accordingly: $\phi_{i,2} = \sum_{h=1}^{n} \mathbf{s}_{w_h} \ell_h - \mathbf{s}_{k_1} \mathcal{L} - \mathbf{s}_{k_1} \mathcal{L} + \sum_{h=1}^{m} \underline{\ell}_h + 2\underline{\ell}^\star = \sum_{h=1}^{m} \underline{\ell}_h + 2\underline{\ell}^\star$; in particular, we have:

$$\text{If SUBSET-SUM}(\ell_1, \ldots, \ell_n; \mathcal{L}) = \text{FALSE, then } \phi_{i,2} = \sum_{h=1}^{m} \underline{\ell}_h + 2\underline{\ell}^\star. \qquad (\text{I})$$

It is also worth highlighting that when SUBSET-SUM$(\ell_1, \ldots, \ell_n; \mathcal{L}) = $ FALSE the belief of agent i at time two is in fact an efficient belief but that does not imply the transparency of the graph structure because the latter is a structural property that should hold true for all choices of the signal structure parameters; in fact, the graph structure in Fig. 5, on the left, is not transparent. On the other hand, if the answer to SUBSET-SUM$(\ell_1, \ldots, \ell_n; \mathcal{L})$ is positive, then agent i concludes that in addition to the case of all zero signals, there are additional cases (i.e. feasible signal profiles) in which some agents receive a one signal. In any such cases, we should necessarily have that $\mathbf{s}_{k_1} = \mathbf{s}_{k_1} = 1$, in order for (E.4) to remain satisfied. Subsequently, for all such nontrivial signal profiles we have that:

$$(\cancel{-\mathcal{L}\mathbf{s}_{k_1}} + \underline{\ell}^\star) + (-\mathcal{L}\mathbf{s}_{k_2} + \underline{\ell}^\star) + \sum_{h=1}^{n} \left(\cancel{\ell_h \mathbf{s}_{w_h}} + \underline{\ell}_h\right) = \underline{\ell}^\star + \underbrace{(-\mathcal{L} + \underline{\ell}^\star)}_{\overline{\ell}^\star} + \sum_{h=1}^{n} \underline{\ell}_h$$

$$= \sum_{h=1}^{n} \underline{\ell}_h + \underline{\ell}^\star + \overline{\ell}^\star < \sum_{h=1}^{n} \underline{\ell}_h + 2\underline{\ell}^\star,$$

where in the first equality we use (E.4) to cancel out the indicated terms and in the last inequality we use the fact that $\overline{\ell}^\star = (-\mathcal{L}) + \underline{\ell}^\star < \underline{\ell}^\star$. Agent i thus needs to find all these feasible signal profiles and set her belief at time two based on the the set of all feasible signal profiles. In particular, since in all the non-trivial cases (feasible signal profiles that are not all zero signals), $\sum_{k=1}^{n} \boldsymbol{\lambda}_k + \boldsymbol{\lambda}_{k_1} + \boldsymbol{\lambda}_{k_2} = \sum_{h=1}^{n} \underline{\ell}_h + \underline{\ell}^\star + \overline{\ell}^\star < \sum_{h=1}^{n} \underline{\ell}_h + 2\underline{\ell}^\star$ we have that:

If SUBSET-SUM$(\ell_1, \ldots, \ell_n; \mathcal{L}) = $ TRUE, then $\phi_{i,2} < \sum_{h=1}^{n} \underline{\ell}_h + 2\underline{\ell}^\star$.   (II)

This concludes the reduction because if an algorithm is available that solves any instances of GROUP-DECISION in polynomial time then by inspecting the output of that algorithm according to (I) and (II) for the particular instance of GROUP-DECISION described above, agent i can decide the feasibility of any instance of the SUBSET-SUM problem in polynomial time.   Q.E.D.

APPENDIX F: AN EXACT-COVER REDUCTION

EXACT-COVER is the fourteenth on Karp's list of 21 $\mathcal{NP}$-complete problems. It is described as follows Karp (1972):

PROBLEM 3 (EXACT-COVER)   *A set of n items $\{j_1, \ldots, j_n\} = \{j_h : h \in [n]\}$ and a family $\{\mathcal{W}_1, \ldots, \mathcal{W}_m\}$ of m subsets: $\mathcal{W}_h \subset \{j_1, \ldots, j_n\}$ for all $h \in [m]$, are given. Determine if there is a subfamily of disjoint subsets belonging to $\{\mathcal{W}_1, \ldots, \mathcal{W}_m\}$ such that their union is $\{j_1, \ldots, j_n\}$: $\{\mathcal{W}_{h_1}, \ldots, \mathcal{W}_{h_p}\} \subset \{\mathcal{W}_1, \ldots, \mathcal{W}_m\}$, $\mathcal{W}_{h_q} \cap \mathcal{W}_{h_{q'}} = \varnothing$ for all $q, q' \in [p]$, and $\cup_{q=1}^{p} \mathcal{W}_{h_q} = \{j_1, \ldots, j_n\}$.*

The input to EXACT-COVER can be represent by a graph $\hat{\mathcal{G}}_{m,n}$ on the $m + n$ nodes $\{w_1, \ldots, w_m ; j_1, \ldots, j_n\}$ which is bipartite between $\{w_1, \ldots, w_m\}$ and $\{j_1, \ldots, j_n\}$ and the only edges are from nodes w_h to $j_{h'}$ whenever subset $\mathcal{W}_h$ contains the element $j_{h'}$ for some $h \in [m]$ and $h' \in [n]$ in the description of EXACT-COVER. Henceforth, we use the notation EXACT-COVER($\hat{\mathcal{G}}_{m,n}$) to denote the output of EXACT-COVER for an arbitrary input $\mathcal{G}_{m,n}$: EXACT-COVER($\hat{\mathcal{G}}_{m,n}$) $\in$ {TRUE,FALSE}. If there is a subset $\mathcal{W}_h$, $h \in [m]$ that alone covers all the items $\{j_1, \ldots, j_n\}$, then the answer to EXACT-COVER($\hat{\mathcal{G}}_{m,n}$) is (trivially) true, and we can thus check for and remove this case in our polynomial reduction.

To construct the reduction from an arbitrary instance of EXACT-COVER to a particular instance of GROUP-DECISION, we consider the decision problem of agent i in a graph $\mathcal{G}$ that is derived from $\hat{\mathcal{G}}_{m,n}$ by adding two additional nodes i and k and $2n$ additional edges: n edges that are directed from node k to each of $\{j_1, \ldots, j_n\}$ and another n edges from each of $\{j_1, \ldots, j_n\}$ to node i (cf. the right graph in Fig. 5). We assume that agents i and $j_1, \ldots, j_n$ can only receive the non-informative signal $\mathring{s}$: $\mathcal{S}_i = \mathcal{S}_{j_1} = \mathcal{S}_{j_2} = \ldots = \mathcal{S}_{j_n} = \{\mathring{s}\}$; hence, $\phi_{i,0} = \phi_{j_1,0} = \phi_{j_2,0} = \ldots = \phi_{j_n,0} = \phi_{i,1} = 0$.

We assume that agents $w_1, \ldots, w_m$ observe initial i.i.d. binary signals: $\mathbf{s}_{w_h} \in \{0, 1\}$ with the probabilities set such that for all $h \in [m]$:

$$\log\left(\frac{\mathbb{P}_{\theta_2}(\mathbf{s}_{w_h} = 1)}{\mathbb{P}_{\theta_1}(\mathbf{s}_{w_h} = 1)}\right) = \bar{\ell}, \log\left(\frac{\mathbb{P}_{\theta_2}(\mathbf{s}_{w_h} = 0)}{\mathbb{P}_{\theta_1}(\mathbf{s}_{w_h} = 0)}\right) = \underline{\ell}, \bar{\ell} - \underline{\ell} = 1.$$

Similarly, agent k receives a binary signal but with probabilities such that

$$\log\left(\frac{\mathbb{P}_{\theta_2}(\mathbf{s}_k = 1)}{\mathbb{P}_{\theta_1}(\mathbf{s}_k = 1)}\right) = \bar{\ell}^\star, \log\left(\frac{\mathbb{P}_{\theta_2}(\mathbf{s}_k = 0)}{\mathbb{P}_{\theta_1}(\mathbf{s}_k = 0)}\right) = \underline{\ell}^\star, \bar{\ell}^\star - \underline{\ell}^\star = -1.$$

Note that with the above setting,

(F.1) $\boldsymbol{\lambda}_k = \mathbf{s}_k \underbrace{(\bar{\ell}^\star - \underline{\ell}^\star)}_{-1} + \underline{\ell}^\star$, and $\boldsymbol{\lambda}_{w_h} = \mathbf{s}_{w_h} \underbrace{(\bar{\ell} - \underline{\ell})}_{1} + \underline{\ell}, h \in [m].$

At time two, agent i observes that each of her neighbors $j_1, \ldots, j_n$ have changed their beliefs from their initial uniform priors, such that

(F.2) $\phi_{j_r,1} = \text{card}\left(\{h \in [m] : j_r \in \mathcal{W}_h\}\right) \underline{\ell} + \underline{\ell}^\star = \left(\text{card}\left(\mathcal{N}_{j_r}\right) - 1\right) \underline{\ell} + \underline{\ell}^\star.$

Note that $\mathcal{N}_{j_r} = \{k\} \cup \{w_h : h \in [m], j_r \in \mathcal{W}_h\}$, and $\mathrm{card}\,(\mathcal{N}_{j_r}) - 1 = \mathrm{card}\,(\{h \in [m] : j_r \in \mathcal{W}_h\})$ counts the number of subsets $\mathcal{W}_h, h \in [m]$ that cover item j_r in the original description of EXACT-COVER (Problem 3). To make a Bayesian inference about the reported beliefs in (F.2) and to decide her time two belief $\boldsymbol{\mu}_{i,2}$ (or equivalently $\boldsymbol{\phi}_{i,2}$), agent i should first consider the following construction of the reported beliefs, $\boldsymbol{\phi}_{j_r,1}$ for all $r \in [n]$:

$$\boldsymbol{\phi}_{j_r,1} = \boldsymbol{\lambda}_k + \sum_{w_h \in \mathcal{N}_{j_r}\setminus\{k\}} \boldsymbol{\lambda}_{w_h} = (-\mathbf{s}_k + \underline{\ell}^\star) + \sum_{w_h \in \mathcal{N}_{j_r}\setminus\{k\}} (\mathbf{s}_{w_h} + \underline{\ell})$$

$$(\text{F.3}) \qquad = \left(\sum_{w_h \in \mathcal{N}_{j_r}\setminus\{k\}} \mathbf{s}_{w_h} - \mathbf{s}_k \right) + (\mathrm{card}\,(\mathcal{N}_{j_r}) - 1)\,\underline{\ell} + \underline{\ell}^\star,$$

Combining her observations in (F.2) with the construction of the reported beliefs in (F.3), agent i should consider the solutions of the resultant system of n equations in the following $m+1$ binary variables: $\mathbf{s}_{w_1}, \ldots, \mathbf{s}_{w_m}$ and $\mathbf{s}_k$. In particular, she has to decide whether her observations in (F.2) are the result of k and $w_1, \ldots, w_m$ having all received zero signals, or else if it is possible that agent k has received a one signal ($\boldsymbol{\phi}_{k,0} = \boldsymbol{\lambda}_k = -1 + \underline{\ell}^\star$) and a specific subset of the agents $w_1, \ldots, w_m$ have also received one signals ($\boldsymbol{\phi}_{w_h,0} = \boldsymbol{\lambda}_{w_h} = 1 + \underline{\ell}$, for all w_h who see $\mathbf{s}_{w_h} = 1$) enough to exactly balance the net effect, leading to (F.2). The latter is possible only if there is a non-trivial solution to the following system:

$$(\text{F.4}) \qquad \sum_{w_h \in \mathcal{N}_{j_r}\setminus\{k\}} \mathbf{s}_{w_h} - \mathbf{s}_k = 0, \text{ for all } r \in [n]; \ (\mathbf{s}_{w_1}, \ldots, \mathbf{s}_{w_m}, \mathbf{s}_k) \in \{0,1\}^{m+1}.$$

This is equivalent to the feasibility of EXACT-COVER($\hat{\mathcal{G}}_{m,n}$) since the latter can be formulated as the following 0/1-integer program:

$$(\text{F.5}) \qquad \sum_{h \in [m]: j_r \in w_h} \mathbf{s}_{w_h} = 1, \text{ for all } j_r \in \{j_1, \ldots, j_n\}; \ (\mathbf{s}_{w_1}, \ldots, \mathbf{s}_{w_m}) \in \{0,1\}^m.$$

Note that a variable $\mathbf{s}_{w_h}$ in System (F.5) will be one only if the corresponding set w_h is chosen in the solution of the feasible EXACT-COVER; moreover, the constraints in (F.5) express the requirement that the chosen sets do not intersect at any of the elements $\{j_1, \ldots, j_n\}$. In other words, each of the n items are contained in one and exactly one subset: for each $p \in [n]$, there is a unique $h \in [m]$ such that $j_p \in \mathcal{W}_h$ and $\mathbf{s}_{w_h} = 1$. System (F.4) having a non trivial solution is equivalent to the feasibility of System (F.5), because in any non trivial solution of (F.4) we should necessarily have $\mathbf{s}_k = 1$; and furthermore, from our construction of the graph $\mathcal{G}$ based on the EXACT-COVER input $\hat{\mathcal{G}}_{m,n}$ we have that $\mathcal{N}_{j_p} = \{w_h : h \in [m], j_p \in \mathcal{W}_h\} \cup \{k\}$ for all $j_p \in \{j_1, \ldots, j_n\}$.

Note that since in our polynomial reduction we have removed the case where all of the items $\{j_1, \ldots, j_n\}$ are covered by one subset $\mathcal{W}_h$ for some $h \in [m]$, in any nontrivial solution of exact cover, we have that $\mathbf{s}_{w_h} = 1$ for at least two distinct values of $h \in [m]$: at least two subsets are needed for all the elements to be covered in a feasible EXACT-COVER. Subsequently, if agent i determines that EXACT-COVER($\hat{\mathcal{G}}_{m,n}$) is FALSE, then she concludes that all agents must have received zero signals and she sets her belief accordingly: $\boldsymbol{\phi}_{i,2} = \boldsymbol{\lambda}_k + \sum_{h=1}^m \boldsymbol{\lambda}_{w_h} + \sum_{r=1}^n \boldsymbol{\lambda}_{j_r} + \boldsymbol{\lambda}_i = \underline{\ell}^\star + m\underline{\ell}$, where we use the facts that $\boldsymbol{\lambda}_{j_r} = \boldsymbol{\lambda}_i = 0$ for all $r \in [n]$, as well as $\boldsymbol{\lambda}_k = \underline{\ell}^\star$ and $\boldsymbol{\lambda}_{w_h} = \underline{\ell}$ for all $w_1, \ldots, w_m$ with zero signals. Put succinctly,

$$\text{If EXACT-COVER}(\hat{\mathcal{G}}_{m,n}) = \text{FALSE, then } \boldsymbol{\phi}_{i,2} = \underline{\ell}^\star + m\underline{\ell}. \qquad (\text{III})$$

However, if the answer to EXACT-COVER($\hat{\mathcal{G}}_{m,n}$) is TRUE, then for any additional feasible signal profile that agent i identifies and determines to satisfy (F.4), it is necessarily true that $\mathbf{s}_k = 1$ and $\mathbf{s}_{w_h} = 1$, for

at list two distinct agents among $\{w_1, \ldots, w_m\}$; hence, for any such additionally feasible signal profiles it is always true that

$$\boldsymbol{\lambda}_k + \sum_{h=1}^m \boldsymbol{\lambda}_{w_h} = -\mathbf{s}_k + \underline{\ell}^\star + \sum_{h=1}^m \mathbf{s}_{w_h} + m\underline{\ell} \geq 1 + \underline{\ell}^\star + m\underline{\ell},$$

where in the latter lower-bound we use the facts that $\mathbf{s}_k$ as well as at least two of $\mathbf{s}_{w_h}$ are one in any non-trivially feasible signal profile, i.e. $-\mathbf{s}_k + \sum_{h=1}^m \mathbf{s}_{w_h} \geq 1$. In particular, we conclude that

If SUBSET-SUM$(\ell_1, \ldots, \ell_n; \mathcal{L}) = $ TRUE, then $\phi_{i,2} < \sum_{h=1}^n \underline{\ell}_h + 2\underline{\ell}^\star$. $\qquad$ (II)

Hence we conclude the $\mathcal{NP}$-hardness of GROUP-DECISION by its reduction to EXACT-COVER. Because if the polynomial time computation of beliefs in GROUP-DECISION was possible, then by inspecting the computed beliefs according to (III) and (IV) for the particular instance of GROUP-DECISION (with i.i.d. binary signals) described above, agent i can decide the feasibility of any instance of the EXACT-COVER problem in polynomial time. $\hfill$ *Q.E.D.*

APPENDIX G: BELIEF CALCULATIONS IN BOUNDED NEIGHBORHOODS WITH I.I.D. SIGNALS

In this example, we consider a variation of the right-hand-side structure in Fig. 5 in which agent k is removed and also n, the number of directly observed neighbors of agent i, is fixed. We show that the belief of agent i at time two can be computed efficiently in the number of indirectly observed neighbors (m). We suppose that the signal structures for agent i, her neighboring agents $j_1, \ldots, j_n$, and the indirectly observed agents $w_1, \ldots, w_m$ are as in Appendix F; subsequently, $\phi_{i,0} = \phi_{j_1,0} = \phi_{j_2,0} = \ldots = \phi_{j_n,0} = \phi_{i,1} = 0$. At time two, agent i has to incorporate the time one beliefs of her neighbors, which are themselves caused by the time zero beliefs of $w_1, \ldots, w_m$: Given $\phi_{j_r,1} = \sum_{w_h \in \mathcal{N}_{j_r}} \boldsymbol{\lambda}_{w_h}$, for $r = 1, \ldots, n$, agent i aims to determine her belief at time two (or equivalently $\phi_{i,2}$). Using (F.1), we can write $\boldsymbol{\psi}_{j_r} = \sum_{w_h \in \mathcal{N}_{j_r}} \mathbf{s}_{w_h}$, where

$$\boldsymbol{\psi}_{j_r} = \frac{1}{\overline{\ell} - \underline{\ell}} \left(\phi_{j_r,1} - \text{card}\left(\mathcal{N}_{j_r}\right) \underline{\ell} \right), r \in [n],$$

are necessarily non-negative integers belonging in to $[m]_0 = \{0\} \cup [m]$, due to their generation process, i.e. the fact that they count the number of one signals that are received in the neighborhood $\mathcal{N}_{j_r}$ of each of the neighbors $j_r, r \in [n]$. For all $r \in [n]$ and $r' \in [m]$, let $a_{j_r,w_{r'}} = 1$ if $w_{r'} \in \mathcal{N}_{j_r}$ and $a_{j_r,w_{r'}} = 0$ otherwise. Denoting $\overline{a}_{j_r} = (a_{j_r,w_1}, \ldots, a_{j_r,w_m})$ and using the transpose notation ^{T}, we can rewrite $\boldsymbol{\psi}_{j_r}$ as an inner product $\boldsymbol{\psi}_{j_r} = \overline{a}_{j_r} \overline{\mathbf{s}}^T$, where $\overline{\mathbf{s}} = (\mathbf{s}_{w_1}, \ldots, \mathbf{s}_{w_m})$. To proceed for each $r \in [m]$, let $\overline{a}_{w_r} = (a_{j_1,w_r}, \ldots, a_{j_n,w_r})$. To determine her belief, agent i acts as follows:

1. For each $\overline{\kappa} = (\kappa_1, \ldots, \kappa_n) \in \{0,1\}^n$, let $\Psi_{\overline{\kappa}} = \{w_r : \overline{a}_{w_r} = \overline{\kappa}\}$, note that $\Psi_{\overline{\kappa}}$ are non-intersecting, possibly empty sets, whose union is equal to $\{w_1, \ldots, w_m\}$. Also let $\eta_{\overline{\kappa}}$ be the number of agents belonging to $\Psi_{\overline{\kappa}}$ who have received one signals; the rest having received zero signals, the variables $\eta_{\overline{\kappa}}, \overline{\kappa} \in \{0,1\}^n$ should satisfy:

(G.1) $\displaystyle\sum_{\overline{\kappa} \in \Xi_r} \eta_{\overline{\kappa}} = \boldsymbol{\psi}_{j_r}$, for all $r \in [n]$, where $\Xi_r = \{\overline{\kappa} : \overline{\kappa} = (\kappa_1, \ldots, \kappa_n), \kappa_r = 1\}$.

2. Note that $\eta_{\overline{\kappa}} \in [\text{card}(\Psi_{\overline{\kappa}})]_0$ for each $\overline{k} \in \{0,1\}^n$, and to determine her belief, agent i needs to find the set Γ_i of all such non-negative integer solutions of (G.1):

- Initialize $\Gamma_i = \varnothing$.
- For each $\overline{\eta} := (\eta_{\overline{\kappa}}, \overline{\kappa} \in \{0,1\}^n) \in \prod_{\overline{k} \in \{0,1\}^n} [\text{card}(\Psi_{\overline{\kappa}})]_0$, if all $\eta_{\overline{\kappa}}$, $\overline{k} \in \{0,1\}^n$ satisfy (G.1) for each $r \in [n]$, then set $\Gamma = \Gamma \cup \{\overline{\eta}\}$.

3. Having thus found Γ_i, agent i sets her belief (or equivalently its log-ratio) as follows:

$$
(G.2) \quad \phi_{i,2} = \log \frac{\sum\limits_{\overline{\eta} \in \Gamma_i} \left[\prod\limits_{\overline{\kappa} \in \{0,1\}^n} \binom{\text{card}(\Psi_{\overline{\kappa}})}{\eta_{\overline{\kappa}}} \right] \left[\mathbb{P}_{\theta_2}(\mathbf{s}=1)^{\sum\limits_{\overline{\kappa} \in \{0,1\}^n} \eta_{\overline{\kappa}}} \right] \left[\mathbb{P}_{\theta_2}(\mathbf{s}=0)^{m - \sum\limits_{\overline{\kappa} \in \{0,1\}^n} \eta_{\overline{\kappa}}} \right]}{\sum\limits_{\overline{\eta} \in \Gamma_i} \left[\prod\limits_{\overline{\kappa} \in \{0,1\}^n} \binom{\text{card}(\Psi_{\overline{\kappa}})}{\eta_{\overline{\kappa}}} \right] \left[\mathbb{P}_{\theta_1}(\mathbf{s}=1)^{\sum\limits_{\overline{\kappa} \in \{0,1\}^n} \eta_{\overline{\kappa}}} \right] \left[\mathbb{P}_{\theta_1}(\mathbf{s}=0)^{m - \sum\limits_{\overline{\kappa} \in \{0,1\}^n} \eta_{\overline{\kappa}}} \right]}.
$$

Note that with private signals restricted to be i.i.d. binary signals, the set Γ_i in fact represents the set of all private signals profiles that are deemed feasible by agent i at time two, as with $\mathcal{I}_{i,t}$ in (3.1). The symmetry of the binary structure allows for the summation over the feasible signal profiles to be simplified as in (G.2) by counting the number of ways in which the agents would receive one signals within each of subsets $\Psi_{\overline{\kappa}}, \overline{\kappa} \in \{0,1\}^n$; this is achieved by the product of the binomial coefficients in (G.2). The overall complexity of computing the Bayesian posterior belief in (G.2) can now be bounded by a total of $O(n2^n m^{2^n})$ additions and multiplications for computing the set Γ_i and another $O(m^{1+2^n}(2^{2n})(3m+2))$ for computing the beliefs (or their ratio) in (G.2). Note that we made no effort in optimizing these computations beyond the fact that they increase polynomially in m for a fixed neighborhood size (n). Q.E.D.

ACKNOWLEDGEMENTS

REFERENCES

AARONSON, S. (2005): "The complexity of agreement," in *Proceedings of the thirty-seventh annual ACM symposium on Theory of computing*, ACM, 634–643.

ACEMOGLU, D., M. A. DAHLEH, I. LOBEL, AND A. OZDAGLAR (2011): "Bayesian learning in social networks," *The Review of Economic Studies*, 78, 1201–1236.

AGUILAR, J. H. AND A. A. BUCKAREFF (2010): *Causing Human Actions*, The MIT Press.

ALI, S. N. (2014): "Social learning with endogenous information," Tech. rep.

ARORA, S. AND B. BARAK (2009): *Computational complexity: a modern approach*, Cambridge University Press.

AUMANN, R. J. (1976): "Agreeing to disagree," *The annals of statistics*, 1236–1239.

BACHARACH, M. (1985): "Some extensions of a claim of Aumann in an axiomatic model of knowledge," *Journal of Economic Theory*, 37, 167–190.

BALA, V. AND S. GOYAL (1998): "Learning from Neighbours," *The Review of Economic Studies*, 65, 595–621.

BANERJEE, A. V. (1992): "A Simple Model of Herd Behavior," *The Quarterly Journal of Economics*, 107, 797–817.

BANG-JENSEN, J. AND G. Z. GUTIN (2008): *Digraphs: theory, algorithms and applications*, Springer Science & Business Media.

BIKHCHANDANI, S., D. HIRSHLEIFER, AND I. WELCH (1998): "Learning from the Behavior of Others: Conformity, Fads, and Informational Cascades," *The Journal of Economic Perspectives*, 12, pp. 151–170.

BLACKWELL, D. AND L. DUBINS (1962): "Merging of Opinions with Increasing Information," *The Annals of Mathematical Statistics*, 33, 882 – 886.

BOGDANOV, A., L. TREVISAN, ET AL. (2006): "Average-case complexity," *Foundations and Trends® in Theoretical Computer Science*, 2, 1–106.

BOLTON, P. AND M. DEWATRIPONT (1994): "The firm as a communication network," *The Quarterly Journal of Economics*, 109, 809–839.

CAVE, J. (1983): "Learning to agree," *Economics Letters*, 12, 147–152.

CHANDRASEKARAN, V., N. SREBRO, AND P. HARSHA (2012): "Complexity of inference in graphical models," *arXiv preprint arXiv:1206.3240*.

CHEN, Y. AND J. XU (2016): "Statistical-computational tradeoffs in planted problems and submatrix localization with a growing number of clusters and submatrices," *Journal of Machine Learning Research*, 17, 1–57.

CHRISTENSEN, M. AND T. KNUDSEN (2010): "Design of decision-making organizations," *Management Science*, 56, 71–89.

COWELL, R. G., P. DAWID, S. L. LAURITZEN, AND D. J. SPIEGELHALTER (2006): *Probabilistic networks and expert systems: Exact computational methods for Bayesian networks*, Springer.

CSASZAR, F. A. AND J. EGGERS (2013): "Organizational decision making: An information aggregation view," *Management Science*, 59, 2257–2277.

DASKALAKIS, C., P. W. GOLDBERG, AND C. H. PAPADIMITRIOU (2009): "The Complexity of Computing a Nash Equilibrium," *SIAM Journal on Computing*, 39, 195–259.

EYSTER, E. AND M. RABIN (2010): "Naive herding in rich-information settings," *American economic journal: microeconomics*, 2, 221–243.

——— (2014): "Extensive imitation is irrational and harmful," *The Quarterly Journal of Economics*, qju021.

GALE, D. AND S. KARIV (2003): "Bayesian learning in social networks," *Games and Economic Behavior*, 45, 329–346.

GAREY, M. R. AND D. S. JOHNSON (1979): *Computers and Intractability: A Guide to the Theory of NP-Completeness*, New York, NY, USA: W. H. Freeman & Co.

GEANAKOPLOS, J. D. AND H. M. POLEMARCHAKIS (1982): "We can't disagree forever," *Journal of Economic Theory*, 28, 192–200.

GIBBONS, R. (1998): "Incentives in Organizations," *Journal of Economic Perspectives*, 12, 115–132.

GOLUB, B. AND M. O. JACKSON (2010): "Naïve Learning in Social Networks and the Wisdom of Crowds," *American Economic Journal: Microeconomics*, 2, 112–149.

HAREL, M., E. MOSSEL, P. STRACK, AND O. TAMUZ (2014): "The Speed of Social Learning," ArXiv preprint arXiv:1412.7172.

IOANNIDES, Y. M. (1987): "On the architecture of complex organizations," *Economics Letters*, 25, 201–206.

KANORIA, Y. AND O. TAMUZ (2013): "Tractable Bayesian Social Learning on Trees," *Selected Areas in Communications, IEEE Journal on*, 31, 756–765.

KARP, R. M. (1972): "Reducibility among combinatorial problems," in *Complexity of computer computations*, Springer, 85–103.

KRISHNAMURTHY, V. AND W. HOILES (2014): "Online Reputation and Polling Systems: Data Incest, Social Learning, and Revealed Preferences," *IEEE Transactions on Computational Social Systems*, 1, 164–179.

LEHRER, E. AND R. SMORODINSKY (1996): "Merging and learning," *Lecture Notes-Monograph Series*, 147–168.

MARSCHAK, J. AND R. RADNER (1972): *Economic theory of teams.*, New Haven: Yale University Press.

MOLAVI, P., C. EKSIN, A. RIBEIRO, AND A. JADBABAIE (2015): "Learning to coordinate in social networks," *Operations Research*, 64, 605–621.

MOSSEL, E., A. SLY, AND O. TAMUZ (2014): "Asymptotic learning on Bayesian social networks," *Probability Theory and Related Fields*, 158, 127–157.

——— (2015): "Strategic learning and the topology of social networks," *Econometrica*, 83, 1755–1794.

MOSSEL, E. AND O. TAMUZ (2010): "Efficient bayesian learning in social networks with gaussian estimators," *arXiv preprint arXiv:1002.0747*.

——— (2013): "Making consensus tractable," *ACM Transactions on Economics and Computation*, 1, 20.

MOTWANI, R. AND P. RAGHAVAN (2010): *Randomized algorithms*, Chapman & Hall/CRC.

MUELLER-FRANK, M. (2013): "A general framework for rational learning in social networks," *Theoretical Economics*, 8, 1–40.

MURPHY, K. P., Y. WEISS, AND M. I. JORDAN (1999): "Loopy belief propagation for approximate inference: An empirical study," in *Proceedings of the Fifteenth conference on Uncertainty in artificial intelligence*, Morgan Kaufmann Publishers Inc., 467–475.

PAPADIMITRIOU, C. H. AND J. N. TSITSIKLIS (1987): "The complexity of Markov decision processes," *Mathematics of operations research*, 12, 441–450.

PARIKH, R. AND P. KRASUCKI (1990): "Communication, consensus, and knowledge," *Journal of Economic Theory*, 52, 178–189.

PEARL, J. (2014): *Probabilistic reasoning in intelligent systems: networks of plausible inference*, Morgan Kaufmann.

RADNER, R. (1962): "Team decision problems," *The Annals of Mathematical Statistics*, 33, 857–881.

RAHIMIAN, M. A. AND A. JADBABAIE (2016): "Bayesian learning without recall," *IEEE Transactions on Signal and Information Processing over Networks*.

——— (2017): "Bayesian Heuristics for Group Decisions," Preprint.

ROSENBERG, D., E. SOLAN, AND N. VIEILLE (2009): "Informational externalities and emergence of consensus," *Games and Economic Behavior*, 66, 979–994.

SAH, R. K. AND J. E. STIGLITZ (1985): "Human fallibility and economic organization," *The American Economic Review*, 75, 292–297.

——— (1986): "The architecture of economic systems: Hierarchies and polyarchies," *The American Economic Review*, 716–727.

——— (1988): "Committees, hierarchies and polyarchies," *The Economic Journal*, 98, 451–470.

SIMON, H. (1990): *Reason in Human Affairs*, Stanford University Press.

TSITSIKLIS, J. N. AND M. ATHANS (1985): "On the complexity of decentralized decision making and detection problems," *IEEE Transactions on Automatic Control*, 30, 440–446.

VASSILAKIS, S. (1997): "Accelerating new product development by overcoming complexity constraints," *Journal of Mathematical Economics*, 28, 341–373.

VELUPILLAI, K. (2000): *Computable economics: the Arne Ryde memorial lectures*, vol. 5, Oxford University Press on Demand.

VISSER, B. (2000): "Organizational communication structure and performance," *Journal of economic behavior & organization*, 42, 231–252.

WANG, T., Q. BERTHET, R. J. SAMWORTH, ET AL. (2016): "Statistical and computational trade-offs in estimation of sparse principal components," *The Annals of Statistics*, 44, 1896–1930.