2022 IEEE 61st Conference on Decision and Control (CDC)
December 6-9, 2022. Cancún, Mexico

Social Learning with a Self-Interested Coordinator

Xupeng Wei and Achilleas Anastasopoulos

2022 IEEE 61st Conference on Decision and Control (CDC) | 978-1-6654-6761-2/22/$31.00 ©2022 IEEE | DOI: 10.1109/CDC51059.2022.9992549

Abstract— Social learning refers to the process by which networked strategic agents learn an unknown state of the world by observing state-related private signals as well as other agents' actions. In the classic study of social learning by Bikhchandani, Hirshleifer, and Welch, it was shown that in this setting, information cascades occur, in which agents blindly imitate others' behavior and as a result learning stops for the whole community. Several proposals have been forwarded to mitigate this detrimental phenomenon.

In this paper we consider the introduction of an information coordinator to mitigate information cascades. The coordinator commits to a contract and agents choose to enter the mechanism or not. If they enter they pay a fee and inform the coordinator of their private information (not necessarily truthfully). The coordinator, in turn, suggests an action to the agents based on his cumulative knowledge. We study a class of mechanisms that possess properties such as individual rationality for agents (i.e., agents want to enter the mechanism), truth telling, and profit maximization for the coordinator. We show the existence of such a mechanism which strictly improves social welfare, and results in strictly positive profit for the coordinator, so that agents and the coordinator are willing to adopt this approach. Furthermore, we analyze the performance of this mechanism and show significant gains on both aforementioned metrics.

I. Introduction

Learning in general social networks is characterized by the following salient features: there is an unknown state of the world (e.g., the value of a new product or a new technology) that agents want to estimate in order to improve their well being. Agents themselves do not observe this state of the world directly, but only indirectly through some private signals (e.g., recommendations from friends, etc). Since agents are selfish they may not want to share this private information. Nevertheless, they take actions (e.g., buy the product or adopt the new technology) and these actions are publicly observed in the network by all agents. As a result, the action of each agent is based on the public information available (e.g., buying actions of previous agents) and their own private information. Through this process, information about the state of the world, which is beneficial for the entire community, is only partially revealed through the actions of the agents. This partial revelation of the private information may lead to catastrophic behavior in social networks: although the community as a whole has sufficient information to accurately estimate the state of the world, because of the partial revelation of this information through the agents' actions, the actual information revealed in the network is minimal. This scenario was studied in the seminal works [1]–[3], with the key result being the existence of a phenomenon called *information cascade* which is shown to occur almost surely in sequential social learning. When a cascade occurs, each agent makes decisions based only on the observed common history of the previous actions and disregards her own private information. As a result there is no signaling and learning stops completely in the network. This results in a cascading event (i.e., a herding behaviour). The results of these seminal works raise the following question: if strategic agents are not allowed to have direct communication with each other, are there any approaches to disseminating information more efficiently throughout the network?

A number of ideas have been developed to address this question. In [4], [5] noisy observations (action error) and extra observations (a review from previous agents) are introduced, while in [6] agents are allowed to ask binary questions thus achieving asymptotic learning. In [7] it is shown that information cascades can be avoided by non-myopic agents.

In this paper we consider the introduction of a third party, an information coordinator, to enable efficient information dissemination. The coordinator commits to a *mechanism*, which is a public contract specifying what the coordinator will do during the social learning process. According to this contract, each agent first decides whether to join the mechanism and pay an entrance fee, or not to join. If the agent joins, she privately shares her private information with the coordinator, and in turn receives a recommendation from the coordinator as to whether she should buy the product or not. This process resembles a consultation in real life. The client (agent) pays to meet the consultant (coordinator), tells the consultant about her own situation, and asks for advice. The consultant is experienced because he learns from past consultations. Although he is not allowed to disclose others' private information due to professional ethics, he can still make recommendations based on what he knows. For this approach to work, a number of concerns have to be taken into account in the design process. On the agents' side, due to privacy concerns, they may not be willing to share their information with the coordinator truthfully. Thus appropriate incentives have to be introduced. Furthermore, for agents to join the mechanism, an improvement in their utility has to be guaranteed by design. On the coordinator's side, the contract has to be designed in such a way that it results in some positive profit in order for him to have incentive to maintain the mechanism. Moreover, as the information accumulates, storing and sharing the information can be costly. Therefore, the recommendation decision has to be done in an economical way.

The key contributions of this work are as follows:

This research is funded by the National Science Foundation grant ECCS 2015191.

- We introduce a class of mechanisms that we call "feasible and profit maximizing" (FPM) that have the following properties: they are individually rational (IR), truth-telling (TT), have non-negative taxes (NT), and are profit maximizing (PM).
- For FPM mechanisms we analyze the equilibrium behavior of agents and the coordinator.
- A specific FPM recommendation mechanism is constructed, which we call "no switch if indifferent" (NSII). This mechanism improves the social welfare of the agent's community, and is strictly profitable for the coordinator.
- An exact performance analysis of the NSII is performed, based on which we provide numerical results that show a substantial improvement on net social welfare comparing to the case without a coordinator and a substantial coordinator profit.

Our approach borrows ideas from "information design" and "mechanism design". In information design [8]–[11], the designer, known as a sender, is allowed to reveal state-relevant signals to agents (receivers) to influence agents' actions. The designer commits to an information mechanism, which specifies how the signals are generated conditioned on the state. The agents then form a belief with signals and take actions. In mechanism design [12], information flows in a reverse direction. The designer (as a receiver), commits to a mechanism specifying the allocation and taxes as functions of agents' reported messages. Our work contains both flows of information. Agents possess private information, and the designer has another set of information that agents don't have. In similar works [13], [14], the designer knows the state. However, the designer in this paper is not aware of the state of the world. Instead, he accumulates his private information by learning from agents.

The remainder of the paper is organized as follows. The model of social learning with a coordinator is described in Section II. Section III characterizes the equilibrium behavior of agents and the coordinator. Section IV presents results regarding the increased social welfare and profitability of the proposed mechanism, as well as an exact analysis of the NSII mechanism. Numerical results are presented in Section V. Section VI concludes the paper.

II. Model

The model in this work introduces a self-interested coordinator into the basic observational learning model of [1]. In the basic model, there is an unknown *state of the world* $W \in \mathcal{W} = \{-1, 1\}$, which models for instance the (unobserved) quality of a good. A sequence of agents $t = 1, 2, \ldots$ are coming to the marketplace, each at time t, take an action $A_t \in \mathcal{A} = \{-1, 1\}$ and then leave the marketplace forever. Every agent's objective is to match her action with the *state of the world*, i.e., agent t's utility is $u(w, a_t) = 1_w(a_t)$[1]. The prior belief $P(W = w) = 1/2$ for $w = \pm 1$ is a

[1] $1_a(x)$ is an indicator function of x such that it returns 1 if $x = a$, else 0. For simplicity of exposition we are assuming a slightly different utility from the one considered in [1] which was $u(w, a_t) = w(a_t + 1)/2$.

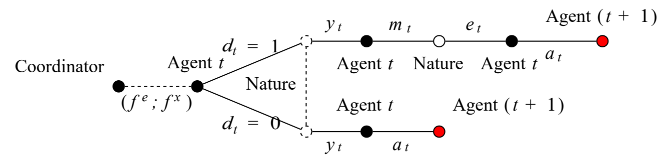

Fig. 1. Part of the game tree for the mechanism-induced game. The dashed line for "Nature" indicates that the realization of y_t is independent of d_t. The red node represents the next subgame for agent $(t+1)$.

common knowledge. In the original setting, each agent t privately observes W through a binary symmetric channel with a crossover probability $p \in (0, 1/2)$, and we denote the private observation by Y_t. We define $Q^y(y|w) := P(Y = y|W = w)$, so that $Q^y(-w|w) = p$. We also denote $q := Q^y(w|w) = 1 - p$. At each time t, only agent t takes action a_t based on her observation on the history $a_{1:t-1}$ (here $x_{j:k}$ represents $x_j, \ldots, x_k$, and sometimes we also use the notation $x^k = x_{1:k}$) and her private observation y_t. Agents are Bayesian learners, so agent t's rationality dictates that

$$a_t \in \arg\max_a \mathrm{P}(W = a \mid a_{1:t-1}, y_t). \tag{1}$$

We now augment the original model of [1] with a *coordinator*. Before the decision making process, the coordinator first commits to a *mechanism*. A mechanism is a contract between the coordinator and the agents, the contents of which are known to all. The agents who sign this contract pay a nonnegative service fee (tax) to the coordinator, report necessary information, and receive recommended actions. The overall process is depicted in Fig. 1 and is detailed below. With this mechanism, before agent t takes an action, she decides whether to join the mechanism or not. This decision is denoted by d_t, with $d_t = 1$ meaning "join", and $d_t = 0$ meaning "not join", based on agent t's observed history d^{t-1}, a^{t-1}. Subsequently, agent t receives her private signal y_t. If she chooses not to join the mechanism, she needs to take action a_t based only on d^{t-1}, a^{t-1}, y_t. If she chooses to join, she pays an entrance fee $f^x(d^{t-1}, a^{t-1})$ and reports a message $m_t \in \mathcal{W}$ (confidential to other agents) directly to the coordinator. After that, agent t receives a recommendation e_t generated by the conditional distribution $f^e(\cdot|d^{t-1}, a^{t-1}, n_t, m_t) \in \Delta(\mathcal{A})$ based on the public history d^{t-1}, a^{t-1}, private information m_t and a summary $n_t = \sum_{i=1}^{t-1} m_i$ of $m_{1:t-1}$, which is the difference between the number of 1's and -1's in $m_{1:t-1}$. Then, agent t takes action $a_t = e_t$[2]. The two functions f^e and f^x constitute the contract between the coordinator and the agents and in our setting the coordinator commits to this contract (so that both functions are know to the agents).

The mechanism induces a sequential game where both the coordinator and agents are the players. As a result, the timeline for the realizations of random variables can be written as

$$W, \ldots, D_t, Y_t, M_t, E_t, A_t, D_{t+1}, \ldots,$$

[2] This type of behavior is called obedience. In our setting, it can be achieved by imposing a large penalty term on tax for the disobedient agents. This can be achieved since the coordinator knows e_t and also observes a_t.

where $M_t; E_t$ are not realized if $D_t = 0$.

With the introduction of the coordinator, if agent t joins the mechanism, her net utility is $u(w; a_t) - f_t^x(d^{t-1}; a^{t-1})$, and the coordinator's profit is a sum of the taxes received from each agent t (discounted by β^{t-1}, $\beta \in (0; 1)$).

In this problem, we restrict attention to mechanisms $f = (f^x; f^e)$ with the following four properties:

Individual Rationality (IR). For a Bayesian rational agent, joining the mechanism brings an expected net utility which is no worse than not to join.

Truth-telling (TT). When joining the mechanism, agents are willing to report $m_t = y_t$.

Nonnegative Tax (NT). Tax $f^x(d^{t-1}; a^{t-1})$ is nonnegative for any $d^{t-1}; a^{t-1}$.

Profit Maximization (PM). Suppose the mechanism is IR and TT, then for a given f^e, there is no other tax scheme $f^{x'}$ such that $(f^{x'}; f^e)$ satisfies IR, TT, and $f^{x'} \geq f^x$.

The nonnegativity restriction for tax is set for easing the analysis. We will show later that even with imposing this additional restriction on taxes, we can find a profitable mechanism for the coordinator which results in a positive social welfare improvement as well. The fourth property, PM, refers only to the tax part f^x for a given f^e. Mechanisms $(f^x; f^e)$ satisfying IR, TT, NT and PM are called feasible and profit maximizing (FPM) mechanisms. In some special cases, an agent may be indifferent between joining and not joining, or indifferent between telling the truth and not doing so. For these scenarios, we further assume that ties are broken in such a way that agents always choose to join and tell the truth. This behavior can be justified by the positive effect to the social welfare brought by the mechanism, which will be shown in subsequent sections.

III. Perfect Bayesian Equilibrium

We adopt perfect Bayesian equilibrium (PBE) as the solution concept. To complete the definition of PBE, we first introduce the agents' strategy profile and information sets ("infosets" for short). Let $g_t = (g_t^d; g_t^m; g_t^a)$ denote the strategy of agent t, and $g = g_{1:I}$. Agent t would see the following possible infosets:

1) Infoset $h_t^d = (f; d^{t-1}; a^{t-1})$ relating to decision $D_t \sim g_t^d(\cdot|h_t^d)$.
2) Infoset $h_t^m = (f; d^{t-1}; a^{t-1}; d_t = 1; y_t)$ relating to decision $M_t \sim g_t^m(\cdot|h_t^m)$;
3) Infoset $h_t^a = (f; d^{t-1}; a^{t-1}; d_t = 0; y_t)$ relating to decision $A_t \sim g_t^a(\cdot|h_t^a)$.

Observe that the coordinator's strategy f is included in the infosets of agents. This is equivalent to saying that the coordinator commits to the contract f. Further define the belief system $\mu : H \to \Delta(W \times Z)$, where H is the collection of infosets. The motivation for such beliefs is that the coordinator's suggestion $E_t \sim f^e(\cdot|d^{t-1}; a^{t-1}; n_t; m_t)$ is generated based on information $d^{t-1}; a^{t-1}, m_t$, which is known to agent t, as well as the summary information n_t which is unknown to her. As a result, agent t puts a belief on the unknown variable n_t, as well as the state of the world w in order to evaluate expected future rewards. We define the PBE as follows.

Definition 1: A PBE is an assessment consisting of a strategy profile $(f; g)$ and a belief system μ satisfying

1) *Sequential Rationality.* On each infoset h, the strategy given by the profile $(f; g)$ is a best response toward the belief specified by μ, i.e., for every agent t (resp. the coordinator) and infoset h of this player, the strategy g_t (resp. f) maximizes her expected payoff under the belief $\mu(\cdot|h)$.
2) *Consistency on Path.* μ is consistent with the probability distribution induced by the strategy profile $(f; g)$ at equilibrium. Consistency requires the belief update on the equilirbium path satisfies Bayes' rule.
3) *Plausible Off-Path Belief.* The conditional belief for the off-path infoset should also be specified by μ, in a reasonable way (specified later).

For the third requirement in Definition 1, there are various interpretations on "a reasonable way" in the literature (see [15, Chap. 8], [16, Chap. 5], and [17]). In this work, if the coordinator deviates to some FPM strategy f', since f' is observed by agents, Bayes' rule still works for the update of μ. For agents' deviation, if f is FPM, then d_t reveals no information, and so the deviation on d_t won't play any role in the update. The deviation on m_t is private, so it has no influence on the belief update. Finally, although agents' off-path deviation on a_t may influence the successive beliefs, since the purpose for the deviation analysis is for agent t to check if a profitable deviation exists, and the further belief changes do not affect her payoff (since she only acts once), the influence on belief brought by an off-path deviation on a_t does not matter. For the rest of this section, we further elaborate on the first two requirements of PBE.

A. Consistency

Denote by $P^{f;g}(\cdot)$ the probability distribution induced by strategy profile $(f; g)$. On the equilibrium path, a consistent belief system μ satisfies

$$\mu(w; n_t|h_t) = \frac{P^{f;g}(w; n_t; h_t)}{P^{f;g}(h_t)}. \tag{2}$$

These beliefs can be updated in an iterative fashion. By Bayes' rule, the belief at infosets $h_t^m = (h_t^d; d_t = 1; y_t)$ can be calculated by that of h_t^d as

$$\mu(w; n_t|h_t^m) = \frac{Q^y(y_t|w)\mu(w; n_t|h_t^d)}{\sum_{w,n} Q^y(y_t|w)\mu(w; n|h_t^d)}; \tag{3}$$

and since d_t is agent's own decision, it won't influence agent's belief over $(w; n_t)$, so on the off-path infosets $h_t^a = (h_t^d; d_t = 0; y_t)$ we have

$$\mu(w; n_t|h_t^a) = \mu(w; n_t|h_t^m). \tag{4}$$

On the equilibrium path, the belief at infoset h_t^d can be evaluated recursively as

$$
\begin{aligned}
&(w; n_{t+1} | h^d_{t+1}) \\
&= \frac{\sum_{m_t} (w; n_{t+1} \; m_t | h^d_t) Q^y(m_t | w) \; f^e_t(a_t | d^{t\;1} a^{t\;1}; n_{t+1} \; m_t; m_t)}{\sum_{w; n; m_t} (w; n | h^d_t) Q^y(m_t | w) f^e(a_t | d^{t\;1}_t a^{t\;1}; n \; m_t)} : \quad (5)
\end{aligned}
$$

Therefore, the belief update of $(\,|h^d)_t$ on the equilibrium path can be described by an update function $\phi^{f^e}_t()$ as

$$(| h^d_{t+1}) = \phi_t^{f^e}((| h^d_t); d_t = 1; a_t):$$

The belief at off-path h^d_t can also be evaluated, but since for deviated agent $t' < t$, the future a_t does not influence her profit at all, in the analysis of unilateral deviation in equilibria, we don't have to consider the future off-path belief. Thus, for off-equilibrium paths, the beliefs can be defined in an arbitrary way, which won't affect the subsequent analysis.

As it turns out, $(w; n_t | h^d_t)$ has a special structure. The following Lemma 1 shows that one can recover the joint belief $(w; n_t | h^d_t)$ from the marginal $(n_t | h^d_t)$, and so it suffices to track the belief $(n_t | h^d_t)$.

Lemma 1: $(w; n_t | h^d_t) = \hat{h}(w | n_t)(n_t | h^d_t)$, where

$$\hat{h}(1 | n_t) = \frac{(q=p)^{n_t}}{(q=p)^{n_t} + 1}; \quad \hat{h}(\;1 | n_t) = \frac{=p\,1}{(q\;\;)^{n_t} + 1}:$$

Proof: The proof is a direct consequence of the fact

$$(1; n_t | h^d_t) = (\;1; n_t | h^d_t) \; (q=p)^{n_t}: \quad (6)$$

Now, define $_t(n_t) := (n_t | h^d_t)$, which is a variable of public belief on n_t based on public hstory h^d_t up to time t. ■ can be updated recursively as $_{t+1} = T(_t; d_t = 1; a_t)$:

$$
\begin{aligned}
&_{t+1}(n_{t+1}) = \\
&\frac{\sum_{y_t} {}_t(n_{t+1} \; y_t) \sum_w \hat{h}(w | n_{t+1} \; y_t) Q^y(y_t | w) \; f^e(a_t | d^{t\;1} a^{t\;1}; n_{t+1} \; y_t; y_t)}{\mathrm{P}^{f;g}(d_t = 1; a_t | h^d_t)} ; \quad (7)
\end{aligned}
$$

where the denominator is the same as (5) but replacing with $_t(n)\hat{h}(w | n)$. In the following we use the shorthand notation $_{t+1} = T(_t; a_t)$.

B. Rationality

Fix an FPM mechanism $f = (f^x_t; f^e_t)$ and a belief system. Since only the on-path infosets h_t's are of our interest, we depict agent t's rationality in the subgame starting from a on-path h^d_t. The analysis of this subsection serves two purposes: it unravels agents' rationality, and also uncovers the constraints for IR, TT, NT and PM requirements.

The analysis is done in a backward recursive manner. First consider the decision m_t at infoset $h^m_t = (h^d_t; d_t = 1; y_t)$. The expected utility for choosing a certain m_t is

$$U_t(m_t | h^m_t) = \mathrm{P}^{f;g}(A_t = W \,|\, h^m_t; m_t) \quad f^x_t(d^{t\;1}; a^{t\;1}); \quad (8)$$

where the above probability can be derived by analyzing the corresponding joint probability

$$
\begin{aligned}
\mathrm{P}^{f;g}(a_t; w \,|\, h^m_t; m_t) &= \frac{\sum_{n_t} \mathrm{P}^{f;g}(a_t; m_t; y_t; n_t; w \,|\, h^d_t)}{\mathrm{P}^{f;g}(y_t; m_t \,|\, h^d_t)} \\
&= \frac{\sum_{n_t} (w; n_t | h^d_t) Q^y(y_t | w) f^e(a_t | d^{t\;1}; a^{t\;1}; n_t; m_t)}{\sum_{w; n_t} (w; n_t | h^d_t) Q^y(y_t | w)} : \quad (9)
\end{aligned}
$$

The expected utility is therefore

$$
\begin{aligned}
&U_t(m_t | h^m_t) \\
&= \frac{\sum_w \sum_{n_t} (w; n_t | h^d_t) Q^y(y_t | w) f^e(w | d^{t\;1}; a^{t\;1}; n_t; m_t)}{\sum_{w; n_t} (w; n_t | h^d_t) Q^y(y_t | w)} \\
&\quad f^x_t(d^{t\;1}; a^{t\;1}): \quad (10)
\end{aligned}
$$

Sequential rationality implies that if $g^m_t(m_t | h^m_t) > 0$ for some m_t, then

$$m_t \; 2 \; \arg\max_m U_t(m | h^m_t): \quad (11)$$

Since f is FPM, TT requires $U_t(y_t | h^m_t) \;\; U_t(\;\; y_t | h^m_t)$. This indeed provides TT constraints for all on-path h^m_t:

$$
\begin{aligned}
&\sum_w \sum_{n_t} (w; n_t | h^d_t) Q^y(y_t | w) \Big(f^e(w | d^{t\;1} a^{t\;1} n_t; y_t) \\
&\qquad f^e_t(w | d^{t\;1} a^{t\;1} n_t; \;\; y_t) \Big) \;\; 0: \quad (12)
\end{aligned}
$$

We now consider the infoset at the decision for a_t in the case that the user did not join the mechanism, $h^a_t = (h^d_t; d_t = 0; y_t)$. The expected utility for choosing a_t at h^a_t is

$$U_t(a_t | h^a_t) = (w | h^a_t) = \frac{\sum_{n_t} (w = a_t; n_t | h^d_t) Q^y(y_t | a_t)}{\sum_{w n_t} (w n_t | h^d_t) Q^y(y_t | w)} : \quad (13)$$

Therefore, if $g^a_t(a_t | h^a_t) > 0$ for some a_t, then

$$a_t \; 2 \; \arg\max_a U_t(a | h^a_t): \quad (14)$$

Since $U_t(a_t | h^a_t)$ depends on h^d_t through $_t() = (| h^d_t)$, $g_t(a_t | h^a_t)$ can be written as $g_t(a_t | _t; y_t)$.

We now consider the decision to join the mechanism or not. For this decision, agent t compares the expected payoff for joining with the expected payoff for not joining. Since f is FPM, IR requires $g^d_t(1 | h^d_t) = 1$, which requires the following for all on-path h^d_t,

$$
\begin{aligned}
&\mathrm{E}^{f;g}\left\{ U_t(Y_t | h^d_t; d_t = 1; Y_t) | h^d_t \right\} \\
&\mathrm{E}^{f;g}\left\{ \sum_{a_t} g^a_t(a_t | _t; Y_t) U_t(a_t | h^d_t; d_t = 0; Y_t) | h^d_t \right\} ; \quad (15)
\end{aligned}
$$

which is simplified to

$$
\begin{aligned}
&\sum_{y_t} \sum_w \sum_{n_t} (w; n_t | h^d_t) Q^y(y_t | w) \\
&\Big(f^e_t(w | d^{t\;1}; a^{t\;1}; n_t; y_t) \quad g^a_t(w | _t; y_t) \Big) \; f^x(d^{t\;1}_t; a^{t\;1}): \quad (16)
\end{aligned}
$$

Notice that the tax function f^x_t is not involved in TT constraints (12), and has an upper bound given by IR (16). By

IR and PM, for a fixed f_t^e, this upper bound must be reached by $f_t^x(d^{t-1}, a^{t-1})$. At the PBE of interest, the coordinator's expected discounted revenue is defined by

$$\mathrm{E}^{f^e, g}\left\{\sum_{t=1}^{\infty} \delta^{t-1} f_t^x(D^{t-1}, A^{t-1})\right\}, \tag{17}$$

and by equation (16), f_t^x is determined by f_t^e at equilibrium. Once f_t^x is fixed to satisfy PM, the other two constraints left are TT and NT. TT is (12), and NT implies the LHS of (16) is nonnegative.

IV. Main Results

This section presents the results that justify the introduction of a coordinator in sequential social learning from different viewpoints. For the agents' community, they will have a strong motivation to accept a mechanism if there are individual and social welfare improvements. For the coordinator, he has incentive to operate the mechanism only if the mechanism brings a positive revenue. In this section, we will explicitly propose an FPM mechanism that meets the above requirements. For the proposed mechanism we will provide an exact analysis for both social welfare and coordinator profit.

To characterize the improvement of the overall welfare of the community, we define the gross social welfare (GSW) as agents' total income before tax as follows:

$$\mathrm{GSW} = \sum_{t=1}^{\infty} \delta^{t-1} u_t(w, a_t), \tag{18}$$

and the net social welfare (NSW) as the GSW minus taxes.

In this paper, we want to compare the social welfare with a coordinator to that of a non-coordinator system (we refer to the baseline non-coordinator system of [1] as the Bikhchandani, Hirshleifer, and Welch (BHW) system). In BHW environments, agents' decisions are based on a public belief and private information. However, the public belief dominates the decision making process after a certain number of consecutive same actions, and the actions no longer reveal new information. When it happens, the estimation of state of the world is never reverted even if it is incorrect, which halts social learning. This is known as an *information cascade*.

Definition 2 (Information cascade without a coordinator): In the sequential Bayesian learning without a coordinator, a belief $\pi_t(w) = \pi(w|h^d)$ is said to be an information cascade if at equilibrium agent's response toward (π_t, y_t) does not depend on the private information y_t.

From the result of [1], regardless of the p value, agent t chooses a_t based on y_t as long as the absolute difference between actions 1 and -1 among a^{t-1} is strictly less than 2. Once the absolute difference reaches 2, the information cascade occurs, and the successive agents follow the herd. A mechanism $f^{B,e}$ ("B" represents BHW) can be constructed to simulate agents' behavior in BHW scenario[3]:

$$f^{B,e}(e_t|a^{t-1}, m_t) = \begin{cases} 1_{m_t}(e_t), & |\sum_{\tau=1}^{t-1} a_\tau| \le 1, \\ 1_{\mathrm{sign}(\sum_{\tau=1}^{t-1} a_\tau)}(e_t), & \text{otherwise}, \end{cases} \tag{19}$$

and correspondingly, $f^{B,x}(d^{t-1}, a^{t-1}) \equiv 0$. f^B is an FPM mechanism because for every agent, f^B recommends the optimal action based only on the observable information for this agent with no charge, which means it is both harmless and profitless for an agent to join the mechanism and meaningless to lie to the mechanism.

We further define a special type of public histories for mechanism construction.

Definition 3: The public history d^{t-1}, a^{t-1} is said to be a transparent action sequence (TAS) w.r.t. BHW if $\forall\ t$, $\exists y^{t-1}$, $P^{f^B}(y^{t-1}|d^{t-1}, a^{t-1}) = 1_{y^{t-1}}(y^{t-1})$.

The meaning of this definition is that when we are faced with a TAS, we can infer the agents' private information a.s., i.e., no information cascade has occured up to time t.

Next, we propose a mechanism which has strict social welfare improvement and positive expected revenue.

A. NSII Mechanism

We construct an FPM mechanism $f^{N,e}$ ("N" is the first letter of "NSII") by appropriately modifying the BHW mechanism. In order to ensure no loss in social welfare before a cascade occurs, we set $f^{N,e} = f^{B,e}$ for TAS histories. Then, once the information cascade occurs, the mechanism switches to the following recommendations dependent only on n_t and m_t:

$$f_t^{N,e}(e_t|n_t, m_t) = \begin{cases} 1_1(e_t), & n_t + m_t > 0 \text{ or } n_t > 0, \\ 1_{-1}(e_t), & \text{otherwise}. \end{cases} \tag{20}$$

Note that given y^t and $m_t = y_t$, the likelihood ratio of the state is $(q/p)^{n_t+m_t}$. If $n_t + m_t > 0$, the best estimation is $w = 1$. If $n_t + m_t = 0$, it is indifferent to choose 1. In this case, if $n_t > 0$ (i.e., $n_t = n_{t-1} + m_{t-1} > 0$), then $f^{N,e}$ recommends 1, which is the same as the recommendation at $t-1$. Similar analysis can be done for the cases with $n_t + m_t < 0$. Thus, we call the mechanism $f^{N,e}$ as "no switch if indifferent" (NSII) mechanism.

The NSII mechanism receives no taxes when $f_t^{N,e} = f_t^{B,e}$. Since the NSII mechanism $f^{N,e}$ depends on h^d through π_t after switching to (20), the tax function in (16) simplifies to

$$\begin{aligned} f_t^{N,x}(d^{t-1}, a^{t-1}) = \sum_w \sum_{m_t} \sum_{n_t} & Q^y(m_t|w)\pi(w, n_t|h^d)_t \\ & (f^{N,e}(w|n_t, m_t) - g^a(w|\pi_t, m_t)). \end{aligned} \tag{21}$$

Theorem 1 shows that the NSII mechanism is FPM.

Theorem 1: f^N is an FPM mechanism.

Proof: Given $\pi(w, m^{t-1}|h^d)$, the correct probability for an estimator $\hat{f}(\cdot|n_t, y_t) \in \Delta(W)$ with inputs n_t, y_t is

$$P^f_{\text{correct}} = \sum_w \sum_{m_t} \sum_{n_t} Q^y(m_t|w)\pi(w, n_t|h^d)\hat{f}(w|n_t, y_t).$$

[3]This version of the BHW resolves ties in a way that maximizes the expected social welfare over all BHW-like systems.

From *Maximum A Posteriori*(MAP) rule, if $n_t; y_t$ are known, the optimal estimate of W is determined by the likelihood $(q=p)^{n_t+y_t}$, so we declare 1 if $n_t+y_t > 0$, declare 0 if $n_t+y_t < 0$, and both are fine if $n_t+y_t = 0$. Thus, $f^{N;e}$ is an optimal estimator if $m_t = y_t$. The case with $m_t = \; y_t$ is another estimator. Comparing the correct probabilities of these two estimators shows TT constraints hold. Then, for NT, $g^a_t(jh^d_t; d_t = 0; y_t)$ is also an estimator without utilizing n_t. Note that the tax $f^{N;x}$ is indeed the difference of correct probabilities, by comparing the correct probabilities of $f^{N;e}$ and g_{t}, the nonnegativity of f^x_t is shown. ■

In the following we show the existence of an FPM mechanism with strict improvement on welfares and positive profit by checking whether NSII possesses these properties.

B. Welfare Improvement and Positive Profit

Theorem 2 shows the motivation for an agent community to accept the NSII mechanism.

Theorem 2: The following are true for NSII:

1) NSII improves expected individual welfares.
2) NSII strictly improves the expected net social welfare.

Proof: Fix a realization of $y_{1:t}$. In BHW, if $d^{t\ 1}a^{t\ 1}$ 2 TAS, then $a = y$ for $t \; 1$. Under the mechanism f^e, since we assume $f^{N;e} = f^{B;e}$ for TAS histories, one can verify by induction that $a = y$ still holds. In this case, $f^{N;e}$ does not change agent t's utilities. If $d^{t\ 1}a^{t\ 1}$ 2̸ TAS, information cascade occurs at some $t' < t$. In BHW, agent t knows $y_{1:t'\ 1}$ from public $a_{1:t'\ 1}$, and y_t as private information, but no further information about $y_{t'\ 1:t\ 1}$. In contrast, agent t under $f^{N;e}$ may infer further information about $y_{t'\ 1:t\ 1}$. They choose from (i) using their own rational g^a_t, or (ii) receiving an optimal estimation from $f^{N;e}$ and pay a service fee. From IR, (ii) is at least as good as (i), and even if (i) brings a weak improvement in individual welfare since under $f^{N;e}$ the estimation is based on a weakly larger set of information than BHW, so the first statement is proved.

For the second statement, consider trajectories with $y_{1:2} = 1$, $y_{3:5} = \; 1$. Under BHW $a_{1:5} = 1$, but for NSII, $a_{1:4} = 1$, $a_5 = \; 1$. For agent 6 under NSII, she pays zero tax and plays $a_6 = \; 1$, which is an optimal estimation conditioned on $y_{1:6}$ regardless of y_6, but under BHW she plays 1, which is not the optimal if $y_6 = \; 1$. As a result, her expected utility is strictly improved if $y_6 = \; 1$. Since every one's welfare is weakly improved from statement 1, and there are trajectories leading to strict improvements, statement 2 is proved. ■

Next, we prove that the NSII mechanism is strictly profitable for the coordinator.

Theorem 3 (Profitability): The mechanism $f^{N;e}$ has a positive expected revenue.

Proof: Under the same trajectory $y_{1:4} = 1; 1; \; 1; \; 1$, $a_{1:3} = 1$ for both BHW and NSII systems. Then, BHW system gets into an information cascade, and NSII system switches to $f^{N;e}$ in (20). From the definition, $a_4 = 1$ in both systems. For agent 5, if she joins NSII, the mechanism will guide her to a better estimation $a_5 = \; 1$ when $y_5 = \; 1$, because in BHW, agent 5 doesn't know $y_{3:4} = \; 1; \; 1$ so she would play $a_5 = 1$. From this utility surplus, the coordinator earns a positive profit $f^{N;e}_5(d^4 = 1; a^4 = 1) > 0$. The trajectories with $y_{1:4} = 1; 1; \; 1; \; 1$ has a positive probability, so NSII has a positive expected revenue. ■

Theorem 3 shows strict profitability, but we don't know whether the profit is infinitestimally above 0. In Section IV.C we present an exact analysis of the NSII mechanism. This analysis, together with the numerical results in Section V show that coordinator's profit is substantial as long as private information is neither too informative nor too uninformative.

C. Exact Analysis of NSII Mechanism

In this part, we investigate NSII mechanism in order to characterize the expected coordinator's revenue and social welfare. The NSII mechanism has two stages in implementation. It simply follows $f^{B;e}$ when action sequence is transparent, and then switch to (20). Define indicator z_t

$$z_t = 1fd^{t\ 1}a^{t\ 1}\ 2\ \text{TAS}g: \qquad (22)$$

We will show that $(_t; z_t)$ is a Markov chain. If $z^t = 1$, the Markovianity is guaranteed by BHW. Once $_t = 1_2$ or $1\ _2$, z_{t+1} flips to 0, and $f^{N;e}$ switches to (20). Next, we look into the state transition after $_t = 1_2$ or $1\ _2$.

Lemma 2: For NSII mechanism, starting from $_{t_0} = 1_2$ or $1\ _2$, the following statements are true:

1) For time $t \; t_0$, the support of $_t$ either contains $n_t \; 0$, or $n_t \; 0$, but does not have $n_t > 0$ and $n' < 0$ simultaneously. Accordingly, the states can be catagorized into phase $\; 1$ and 1 (where $_t$ assigns probability to $n_t \; 0$ and $n_t \; 0$ respectively).
2) Once phase flips from $\; i$ to i, collapses to $1_i()$.
3) If for $t \; t_0$, $_t$ is in phase i, then agent t and the coordinator are indifferent to $1_i()$ and any other outside strategy $g^a_t(jh^d_t; y_t)$ of agent t.

Proof: For 1), at $t = t_0$, 1_2 or $1\ _2$ satisfies 1). If $_t$ satisfies 1), assume assigns probability 1 to $n \; 1$ without loss of generality. Then, from (20), $a_t = 1$ implies $n_{t+1} = n_t + m_t \; 0$; $a_t = 0$ implies $n_{t+1} \; 0$, so 1) is true.

For 2), consider a flipping from phase 1 to $\; 1$, the reverse direction is similar. If $_t$ is of phase 1 but $_{t+1}$ is of phase $\; 1$, then $n_t \; 0$ and $n_t + m_t < 0$, which implies $m_t = \; 1$, and $n_t = 0$. Therefore, n_{t+1} has to be $\; 1$.

For 3), suppose $_t = 1_1$, then for agent t,she knows $n_t = 1$, so the likelihood ratio toward w is $(q=p)^{1+y_{t_0+1}} \; 1$. Thus, playing 1 is rational regardless of y_t. Similar proof apply to $_t = 1\ _1$, 1_2 and $1\ _2$.

Next, consider general $_t$. From 2), a sequence $_t$ of phase i can only start from 1_i or 1_{2i} (right after $f^{N;e}$'s switch). Here we consider phase 1 starting from 1_1, the other three cases are similar. If $_{t_0} = 1_1$, then by NSII, $a_{t_0} = 1$. As a result, $_{t_0+1}$ contains the same information as $_{t_0}$, so agent $t_0 + 1$ would play 1 if not to join NSII. If $_{t_0+2}$ is still in phase 1, then $n_{t_0+2} = 1$ or 3. In a couple system with NSII up to $t_0 + 2$ where the coordinator is muted, agent $t > t_0 + 2$ has a likelihood ratio $(q=p)^{n_{t_0+2}+y_t} \quad (q=p)$, because any y_t cannot reverse the action from 1 to $\; 1$, so the actions after $t_0 + 2$ are uninformative (cascade). As a result, agents $t > 3$ make decisions based only on $y_1; y_3$. In

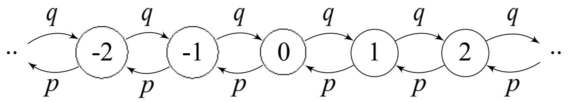

Fig. 2. Markov chain n_t given $W = 1$ related to NSII GSW.

contrast, under NSII mechanism, agents $t > t_0+2$ have extra information from $a_{t_0+2:t-1}$ induced by recommendations. If $a_{t_0+2:t-1} = 1$, it means $n+y \geq 0$ for all $\in [t_0+2; t-1]$, which is more supportive for state $W = 1$. Therefore, agent $t > t_0 + 2$ in the system with a coordinator is more confident about $W = 1$ than that in the couple system. Hence, agent t plays outside action $a_t = 1$ regardless of y_t. ■

Under the NSII mechanism, before it switches to (20), z_t stays 1, and $_t$ jumps among 1_k for $k = -2; \ldots; 2$, depending on the realization of y_t. Once $_t$ reaches 1_2 (or 1_{-2} resp.), due to information cascade, $_{t+1} = 1_3$ (or 1_{-3} resp.) and $z_{t+1} = 0$. From Lemma 2, before phase 1 flips (similar for phase -1 resp.), takes value from ($k = 2; \ldots$)

$$ ^{(+k)}_0 := T^{(k-1)}(^{(+2)}_0 = 1_2; a_{1:k-1} = +1). \quad (23)$$

Once $a_t = -1$, a flip occurs, $_{t+1} = 1_{-1}$, and before next flip to phase 1, takes value from ($k = 1; \ldots$)

$$ ^{(-k)} := T^{(k-1)}(^{(0)} = 1_{-1}; a_{1:k-1} = -1); \quad (24)$$

where $T^{(k-1)}$ means recursively impose T on $^{(0)}$ for $(k-1)$ times. Then, once $a_t = 1$, $_{t+1} = 1_1$. The case for the opposite phase is similar.

Lemma 3: $(_t; z_t)_t$ induced by $f^{N;e}$ is a Markov chain.

Proof: If $z_t = 1, f^{N;e} = f^{B;e}$, $_t$ takes value among 1_n for $n = -2; \ldots; 2$. Given $_t = 1_n; z_t = 1$, one can verify:

$$\begin{aligned} &\mathrm{P}^{f^{N;e}}(_{t+1}; z_{t+1} | _{1:t}; z_{1:t}) \\ =& \sum_w \hat{h}(w|n) \sum_{y_t} Q^y(y_t|w) 1_{1_{n+y_t}}(_{t+1}) 1_z(z_{t+1}). \end{aligned} \quad (25)$$

where $z = 1$ if $|n + y_t| \leq 2$, and otherwise 0.

If $z_t = 0$, $f^{N;e}$ has switched to (20), $z_{t+1} = 0$. Suppose $_t = ^{ik}_0$, $i = \pm 1$, $k = 2; \ldots$, from Lemma 2, $_{t+1}$ is either $^{(i(k+1))}_0$ or $^{(-i)}$. $_{t+1} = ^{(-i)}$ only if $n_t = 0$ and $y_t = -i$. It turns out that

$$\begin{aligned} &\mathrm{P}^{f^{N;e}}(_{t+1}; z_{t+1} | _{1:t}; z_{1:t}) \\ =& 1_0(z_{t+1}) \Big[1_{(-i)}(_{t+1}) \sum_w {}_t(0)\hat{h}(w|0)Q(-1|w) \\ &+ 1_{ ^{(i(k+1))}_0}(_{t+1}) \Big(1 - \sum_w {}_t(0)\hat{h}(w|0)Q(-1|w) \Big) \Big] \\ =& \mathrm{P}^{f^{N;e}}(_{t+1}; z_{t+1} | _t; z_t). \end{aligned}$$

(26)

Similarly, if $_t = ^{ik}$, one only needs to remove the subscript "0" in "$^{( )}_0$".

The Markov process $(_t; z_t)$ with the tax $f^{N;x}()$ forms a Markov reward process (MRP). If $z_t = 1, f^{N;e}$ is indifferent ■ of $f^{B;e}$ and earns no profit. If $z_t = 0$, suppose $_t = ^{(ik)}$, by (21) the instantaneous reward is

$$ f^x(d^{t-1}; a^{t-1}) = \frac{1}{2}(q - p) ^{(ik)}(0) =: r^{ik}; \quad (27)$$

and we can define r^{ik}_0 for $^{ik}_0$ resp. The MRP has 3 stages in a temporal order due to the structure of Markov chain: (1) $z_t = 1$, (2) $z_t = 0$ but $(_t; z_t)$ is transient, (3) $z_t = 0$, $(_t; z_t)$ is recurrent. Any $(^{(ik)}_0; 0)$ is transient, and $(^{(ik)}; 0)$ is recurrent. Use R^{ik} to denote the expected reward-to-go at $(^{ik}; 0)$, R^{ik}_0 to denote that of $(^{ik}_0; 0)$, and R^{ik}_B to denote that of $(1_{ik}; 1)$. By symmetry, $R^{ik} = R^{-ik}$ (means empty, "0" or "B"), so it suffices to focus on R^k with $k \in \mathbb{N}_0$. These R's can be characterized by recursive equations:

Stage 1: for $k = 0; 1; 2$,

$$R^0_B = R^1_B; \quad (28a)$$

$$R^1_B = 2qpR^0_B + (q^2 + p^2)R^2_B; \quad (28b)$$

Stage 2: for $k = 2; 3; \ldots$,

$$R^k_0 = r^k_0 + (q^{k;k+1}_0 R^{k+1}_0 + (1 - q^{k;k+1}_0)R^1); \quad (28c)$$

Stage 3: for $k = 1; 2; \ldots$,

$$R^k = r^k + (q^{k;k+1}R^{k+1} + (1 - q^{k;k+1})R^1); \quad (28d)$$

where $R^2_0 := R^2_B$, $q^{k;k+1}$ is defined by (26) as a transition probability from $^{(k)}$ to $^{(k+1)}$. R^0_B is the expected revenue.

Now evaluate the gross social welfare through another MRP. Since $f^{B;e}$ and $f^{N;e}$ in (20) would provide indifferent expected welfare before cascade, one may use (20) as $f^{N;e}$ for all stages for analysis of GSW. By symmetry, $\mathrm{E}[GSW|W = w] = \mathrm{E}[GSW]$ for any w. Thus, we consider $w = 1$ only. Conditioned on $w = 1$, n_t is a Markov chain with the state transition: $n_{t+1} = n_t + y_t$, and $Y_{t+1} = 1; -1$ with probability $q; p$ resp. (Fig. 2). Let R^n_s be the expected reward-to-go for $n_t = n$ ("s" represents "society"). Then,

$$R^n_s = 1 + (qR^{n+1}_s + pR^{n-1}_s); \quad n > 0; \quad (29a)$$

$$R^n_s = (qR^{n+1}_s + pR^{n-1}_s); \quad n < 0; \quad (29b)$$

$$R^0_s = q + (qR^1_s + pR^{-1}_s); \quad (29c)$$

The closed form of $\mathrm{E}[GSW]$ can be derived from the difference equations, which is omitted due to limited space. The expected gross social welfare $\mathrm{E}[GSW]$ under NSII will be compared with that under $f^{B;e}$, which is

$$\mathrm{E}[GSW] = \frac{q(2 - p^3)}{2(1 -)(1 - pq^2)}. \quad (30)$$

V. Numerical Analysis

The numerical analysis for the performance of the NSII mechanism is done with discount factor $= 0.9$ and crossover probabilities p ranging from 0.005 to 0.495, normalized by multiplying $(1 -)$. For all p's, the expected revenue of the coordinator and the gross expected social welfare (the amount without taxes) under the NSII mechanism are evaluated by (28), (29) using value iteration; the expected social welfare of [1] ("BHW" for short, but note that the utility function here is slightly different) is calculated

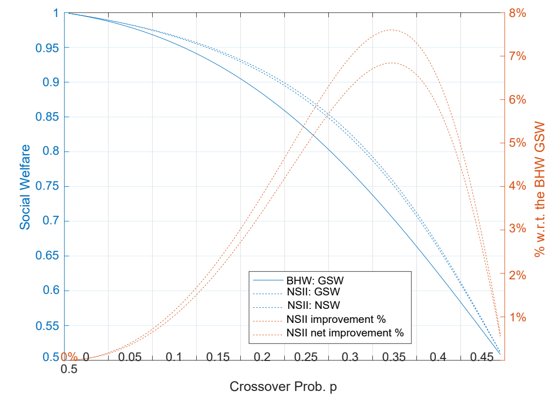

Fig. 3. Social welfares (SW) comparison between [1] and NSII mechanism.

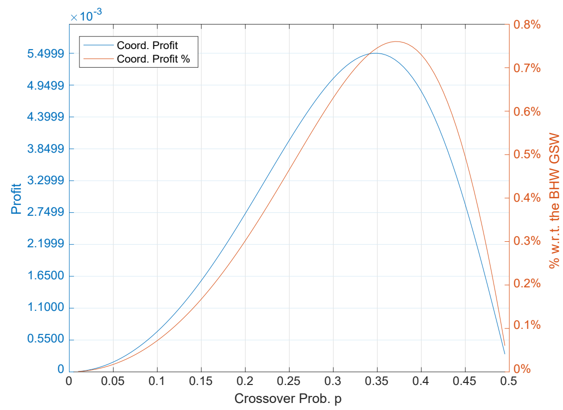

Fig. 4. The coordinator's profit and the percentile of the SW improvement w.r.t. [1].

by (30). The expected net social welfare for NSII is the amount after tax. All the percentage numbers are calculated with respect to the social welfare of BHW. Fig. 3 shows the social welfare for BHW and NSII, and the percentages of the social welfare improvements (gross and net). The results corroborate the social welfare improvement stated in Theorem 2. Both the gross and net social welfares of NSII are better than that of BHW. The improvement is substantial under crossover probabilities away from 0 (fully-informative) and 0.5 (uninformative). At crossover probability $p = 0.37$, the percentage improvement reaches 7.60% for the gross, and 6.84% for the net.

Fig. 4 presents the coordinator's profit (in absolute and relative terms). The results show that the NSII coordinator gains a nontrivial proportion of profit under the crossover probabilities away from $p = 0$ and $p = 0.5$ as well. At $p = 0.37$, the percentage of the profit comparing to BHW social welfare is 0.76%.

By further comparison on Fig. 3 and 4, we find that the percentages of net social welfare improvement and the coordinator's profit share a similar shape, in which both obtain a larger value when private information is neither too informative nor uninformative. This somehow reveals the essence of the recommendation mechanism. By utilizing the mechanism, the coordinator assembles previous private information to build an information asymmetry against future agents. The motivation for agents to join is that this asymmetry improves utilities. Hence, if private information is too informative, agents' private information is close to full information, which leaves less room for asymmetry. On the other side, if private information is close to uninformative case, the coordinator's cumulative information won't grow to an informative one in a short term. As a result, the improvement brought by NSII is more significant under intermediate informative cases rather than the extreme cases.

VI. Conclusion

We introduced a class of recommendation mechanisms to a sequential Bayesian learning system, and investigated a specific mechanism named NSII. The NSII mechanism improves the social welfare, so the community of agents has motivation to join; it is strictly profitable, so the coordinator is willing to deploy it. The performance of NSII was characterized, and we demonstrated the benefits of introducing the mechanism through numerical analysis.

Despite that NSII mechanism is beneficial for agents and the coordinator, it may not be the optimal one from the perspective of the coordinator. Characterization of the optimal mechanism, as well as methods for evaluating the optimal recommendation mechanism are open problems for future research.

References

[1] S. Bikhchandani, D. Hirshleifer, and I. Welch, "A theory of fads, fashion, custom, and cultural change as informational cascades," *Journal of political Economy*, pp. 992–1026, 1992.

[2] A. V. Banerjee, "A simple model of herd behavior," *The quarterly journal of economics*, vol. 107, no. 3, pp. 797–817, 1992.

[3] L. Smith and P. Sørensen, "Pathological outcomes of observational learning," *Econometrica*, vol. 68, no. 2, pp. 371–398, 2000.

[4] T. N. Le, V. G. Subramanian, and R. A. Berry, "The impact of observation and action errors on informational cascades," in *53rd IEEE Conference on Decision and Control*. IEEE, 2014, pp. 1917–1922.

[5] ——, "Are imperfect reviews helpful in social learning?" in *2016 IEEE International Symposium on Information Theory (ISIT)*. IEEE, 2016, pp. 2089–2093.

[6] G. Schoenebeck, S.-T. Su, and V. Subramanian, "Social learning with questions," *arXiv preprint arXiv:1811.00226*, 2018.

[7] I. Bistritz, N. Heydaribeni, and A. Anastasopoulos, "Informational cascades with nonmyopic agents," *IEEE Trans. Automatic Control*, vol. 67, no. 9, pp. 4451–4466, 2022.

[8] E. Kamenica and M. Gentzkow, "Bayesian persuasion," *American Economic Review*, vol. 101, no. 6, pp. 2590–2615, 2011.

[9] J. C. Ely, "Beeps," *American Economic Review*, vol. 107, no. 1, pp. 31–53, 2017.

[10] D. Bergemann and S. Morris, "Information design: A unified perspective," *Journal of Economic Literature*, vol. 57, no. 1, pp. 44–95, 2019.

[11] M. O. Sayin and T. Basar, "Bayesian persuasion with state-dependent quadratic cost measures," *IEEE Transactions on Automatic Control*, 2021.

[12] T. Börgers and D. Krahmer, *An introduction to the theory of mechanism design*. Oxford University Press, USA, 2015.

[13] A. Kolotilin, T. Mylovanov, A. Zapechelnyuk, and M. Li, "Persuasion of a privately informed receiver," *Econometrica*, vol. 85, no. 6, pp. 1949–1964, 2017.

[14] N. Heydaribeni and A. Anastasopoulos, "Joint information and mechanism design for queues with heterogeneous users," *arXiv preprint arXiv:2109.14673*, 2021.

[15] D. Fudenberg and J. Tirole, *Game theory*. MIT press, 1991.

[16] G. J. Mailath, *Modeling Strategic Behavior: A Graduate Introduction to Game Theory and Mechanism Design*. World Scientific, 2018, vol. 6.

[17] J. Watson, "Perfect Bayesian equilibrium: General definitions and illustrations," *Working Paper, University of California at San Diego*, vol. 133, 2016.