WHEN MORE INFORMATION REDUCES THE SPEED OF LEARNING

MATAN HAREL, ELCHANAN MOSSEL, PHILIPP STRACK, AND OMER TAMUZ

Abstract. We consider two Bayesian agents who learn from exogenously provided private signals, as well as the actions of the other. Our main finding is that increased interaction between the agents can lower the speed of learning: when both agents observe each other, learning is significantly slower than it is when one only observes the other. This slowdown is driven by a process in which a consensus on the wrong action causes the agents to discount new contrary evidence.

JEL classifications: C73, D82, D83

1. Introduction

Social learning, or learning from the actions of others, is an integral part of human behavior. Children learn by imitating their parents and peers, and firms copy successful business models and products[1]. Following Aumann [1976], a large literature[2] studies two or more agents who are informed by private signals which are set exogenously[3]. The agents learn by repeatedly observing each other's actions, which reveal information on their private signals.

The basic question that we tackle is the following: how does the amount of information flowing between the agents affect the rate at which they learn? The Bayesian calculations involved in such interactions are difficult to analyze, and pose significant technical obstacles to answering this question, even when the structure of the private signals is simple (see, e.g., Gale and Kariv [2003], Kanoria and Tamuz [2013]). We show the counter-intuitive result that *increased* interaction (i.e., allowing more observations of actions) can lead to significantly *slower* learning. We identify a mechanism contributing to this effect which may be thought of as "Bayesian groupthink": this occurs when initial interactions between

Acknowledgements: The authors would like to thank Amir Dembo for insightful discussions of the large deviation problems arising in this model. We likewise would like to thank Deniz Dizdar, Motty Perry for helpful discussions. A large part of this research was conducted at Microsoft Research New England. E.M. acknowledges the support of NSF grants DMS 1106999 and CCF 1320105, ONR grant number N00014-14-1-0823 and grant 328025 from the Simons Foundation.

[1] See for example Bandura [1965], who demonstrates in a seminal contribution to psychology that 4 to 5 year old kids imitate aggressive behavior towards a doll previously observed by an adult. Social learning is not exclusive to humans; animals learn from the observing the behavior of their peers. See for example Hoppitt [2013] for a study on cockroaches or Auersperg et al. [2014] for Goffin cockatoos.

[2] See for example Geanakoplos and Polemarchakis [1982], Sebenius and Geanakoplos [1983], Parikh and Krasucki [1990], Gale and Kariv [2003], Rosenberg et al. [2009], and more below.

[3] Rather than strategically acquired, as in the bandit literature.

members of a group create a consensus on a wrong belief, causing each member to discount new, contrary evidence, and thus re-enforcing the wrong consensus. From a social planner's perspective, it might therefore be advisable to let agents collect information independently before allowing them to learn from each other.

Our model is a game of incomplete information and purely informational externalities, where two agents repeatedly decide which of two actions to take. There is a state of nature which determines which of the two alternatives is "correct", but it is unknown to the agents which one it is. The stage utility for choosing the correct action is one, and is zero otherwise. Following the literature[4] we consider agents who are myopic (i.e., fully discount the future), so that at each period each one chooses the action which she thinks is more likely to be correct. Every period each agent privately observes a binary, noisy signal regarding the correct choice. Conditioned on the state of nature, the private signals are indepedent and identically distributed.

When an agent observes only her own signals, the probability that she takes the wrong action decays exponentially in the number of signals she observed, with some rate a_p that depends on p, the strength of the signals. In a complete information setting, where each agent can observe the signals of both agents, the probability that one of them takes the wrong action again decays exponentially, but with rate $2\,a_p$, since she sees twice as many signals per time period.

We study two intermediate informational settings, in which information is exchanged between the agents, but they do not observe each other's private signals.

(1) *Unidirectional observation.* Agent 1 can see her own signals, and in addition sees agent 2's actions. Agent 2 can only see her own signals.

(2) *Bidirectional observation.* Each agent can see her own signals and the other agent's actions, but again not the other agent's signals.

In the first case, we find that agent 1's rate of learning is between $3/2\,a_p$ and $25/16\,a_p$; the exact value depends on p, and tends towards the former for stronger signals and towards the latter for weaker signals (Theorem 11). Regardless of the value of p, the learning rate is strictly between a_p and $2a_p$, so that agent 1 learns exponentially faster than she would based solely on her own private signals, but exponentially slower than she would if she could also observe agent 2's private signals. The latter inefficiency stems from the fact that the action is only a coarse signal about the belief. Surprisingly, we find that the rate of learning stays the same if only the last action of agent 2 is observed by agent 1, rather than all of her actions (Theorem 14).

[4]E.g., Geanakoplos and Polemarchakis [1982], Sebenius and Geanakoplos [1983], Gale and Kariv [2003].

Our main result, and the one that requires the most significant technical effort, applies to the case where both agents observes each other's actions. There, we show an upper bound on the rate of learning, and, in particular, we show that agent 1's rate of learning is here strictly *slower* than it is in the previous case (Theorem 21). Naïvely, one may have guessed that in this case the rate should have been higher, since the only difference is that now agent 2 can also see the actions of agent 1, so "more information is exchanged". However, as we show, adding this information to agent 2 makes her actions less informative for agent 1, and thus lowers agent 1's rate of learning.

An alternative way of understanding this result is by analysis of what we call *information transmission efficiency*; this is a way of quantifying how much of an agent's private information - measured as an informationally equivalent fraction of her private signals - is revealed by her actions. Actions that carry no information have 0% efficiency, and actions that reveal all the private signals have 100% efficiency. An agent's actions have (say) 25% efficiency if they carry as much information as do 25% of the agent's private signals. We find that in the case of unidirectional observations the efficiency is between 50% and 57%, depending on the strength of the private signals. In the case of bidirectional observations the efficiency is at most 38%.

To obtain an upper bound on the speed of learning in the bidirectional observation case, we consider the event that both agents make the wrong choice in *every* period, and show that its likelihood is exponentially higher than the probability of agent 1 making a mistake in the unidirectional setting. Typically, such an event occurs when the initial signals are wrong for both agents, causing them both to choose the wrong action. Later, the private signals of each agent indicate the *correct* action, but both agents misestimate the the other agent's signal, and so still both choose the wrong action (Theorem 22); this can be thought of as a model of "Bayesian groupthink".

The analysis of this process requires the overcoming of significant technical difficulties. These are due to the fact that an agent's myopic actions are complex functions of the complete private history - namely the other's actions and the agent's own actions and signals - which in particular admit no simple closed-form expressions (see, e.g., Gale and Kariv [2003], Kanoria and Tamuz [2013]). As the agent's behavior cannot be described explicitly, we focus on the event that both agents have been wrong at all time periods. We show that the probability of this event can asymptotically be characterized by a recursive equation which determines its exponential rate. Our main mathematical tools come from *large deviation theory*, and in particular we use *Sanov's Theorem*, to which we provide a short introduction.

Relation to the literature. Starting with the seminal "Agreeing to disagree" paper of Aumann [1976], a large literature has been devoted to the study of the evolution of opinions and actions of interacting Bayesian agents, with notable contributions by Geanakoplos and Polemarchakis [1982], Sebenius and Geanakoplos [1983], McKelvey and Page [1986], Parikh and Krasucki [1990], Gale and Kariv [2003] and others. By-and-large, it has been shown that barring pathological cases, agents who exchange enough information will eventually agree. For example, Rosenberg, Solan, and Vieille [2009] study games with pure informational externalities where agents observe private signals and learn from the actions of others. They show that agents will eventually act myopically and only disagree when indifferent. More recently, some authors have considered the question of whether or not the agents agree on the *correct* action. Mossel, Sly, and Tamuz [2012a,b, 2014] provide conditions under which agreement implies that the correct action is taken in a setting with infinitely many agents, and Arieli and Mueller-Frank [2013] explore the question of when beliefs can be inferred from actions, which also leads to learning the correct action. Ostrovsky [2012] provides sufficient conditions for information aggregation in financial markets with privately informed traders.

Whereas this literature focuses on whether agents agree in the long-run and whether they learn the correct action, we study the speed at which agents learn the correct action. Specifically, as both agents in our model observe an equally informative signal every period, they learn the state with probably one from their own signals and consequently agree in the long-run. The interesting remaining question, then, is the rate at which this happens.

Closely related to our setup, Ellison and Fudenberg [1995] study a setting of social learning where agents observe the signals of N other random agents during each period, and use simple heuristic decision rules to choose their actions.

Also related is a series of papers on information percolation: Duffie and Manso [2007], Duffie, Malamud, and Manso [2009], Duffie, Giroux, and Manso [2010], Duffie, Malamud, and Manso [2014] study the exponential rate of learning in a continuum of agents who learn about each other's private signals. Recently, Jadbabaie, Molavi, and Tahbaz-Salehi [2013] study a model very similar to ours, but on a general social network, and with boundedly-rational agents. They too use exponential rates as a natural way to quantify the speed of learning. A model with two Bayesian agents who learn an underlying binary state from private signals is studied by Cripps, Ely, Mailath, and Samuelson [2008, 2013]. In their model the agents do not observe each other, but have correlated signals.

In contrast, the literature on social learning in bandit problems focuses on aspects of information acquisition (e.g., Bolton and Harris [1999], Keller, Rady, and Cripps [2005], Keller and Rady [2010], Heidhues, Rady, and Strack [2014]). Signals in this literature are usually publicly observable, and different actions lead to signals of different informativeness. This leads to an inefficiency which arises from a decrease in information acquisition, as agents can free-ride on the information of others. We, on the other hand, study the inefficiency arising from the fact that only actions are observable, while signals are private information.

Furthermore, we abstract away any strategic experimentation considerations, by assuming that information arrives independently of the actions taken.

2. Definitions and results

In this section we formally define our model and give an overview of our results.

2.1. **The probability space.**

We consider a state of nature Θ that takes values in $\{+1, -1\}$, both of which are a priori equally likely: $\mathbb{P}[\Theta = +1] = 1/2$. There are two agents, indexed by $i \in \{1, 2\}$, and n time periods. Each agent observes a sequence of n private signals $\{X_k^i\}_{k \le n}$, which are i.i.d. conditional on the state of the world Θ. The signal $X_k^i \in \{-1, 1\}$ observed by agent i in period k is equal to the true state of the world Θ with probability p and equals $-\Theta$ with probability $1 - p$:

$$\mathbb{P}\left[X_k^i = \Theta \middle| \Theta\right] = p\,.$$

We will use $\mathbb{P}^+[\cdot]$ to denote $\mathbb{P}[\cdot|\Theta = +1]$, and likewise $\mathbb{P}^-[\cdot]$ to denote $\mathbb{P}[\cdot|\Theta = -1]$. We also denote by $S_n^i = \sum_{k \le n} X_k^i$ the difference between the number of $+1$ and -1 signals agent i observed.

2.2. **The agent's actions.**

We define four scenarios, which differ by the set of periods in which agents are allowed to observe the other's actions. We assume throughout that the agents share the uniform prior regarding the state of nature Θ. This is not necessary for our results, but simplifies the proofs, some of which are already laborious.

2.2.1. *No observation.* Our first setting is the baseline scenario, in which each agent chooses an action based on her own private signals, and does not observe the other agent's actions. Since the agents are myopic, each agent's action A_n^i is given in this case by her best guess as to the value of state of the world Θ - i.e.

$$A_n^i := \operatorname*{arg\,max}_{\theta \in \{+1,-1\}} \mathbb{P}\left[\Theta = \theta \,\middle|\, \{X_k^i\}_{k \le n}\right]\,.$$

If the $\arg\max$ is not unique, we let A_n^i take the value 0 (and do the same below for B_n^i, C_n^i and D_n^i); thus we assume that the stage utility of choosing 0 is 1/2. We do this to make the notation and analysis simpler - the results would be identical for any other choice one could make (e.g., always choose $+1$, or choose either $+1$ or -1 at random).

2.2.2. *Observing the final action unidirectionally.* Our second setting is one in which agent 2 still only observes her own signals, and therefore her actions are A_n^2, as above. Agent 1, at time n, observes agent 2's penultimate action A_{n-1}^2, in addition to her own private signals. Hence agent 1's action B_n^1 is given in this case by

$$B_n^1 := \operatorname*{arg\,max}_{\theta \in \{+1,-1\}} \mathbb{P}\left[\Theta = \theta \,\middle|\, \{X_k^1\}_{k \le n}, A_{n-1}^2\right]\,.$$

2.2.3. *Observing all actions unidirectionally.* This setting is a small modification of the previous: the difference is that agent 1 now gets to observe *all* of agent 2's past actions. Hence agent 1's action C_n^1 is given by

$$C_n^1 := \operatorname*{arg\,max}_{\theta \in \{+1,-1\}} \mathbb{P}\left[\Theta = \theta \,\middle|\, \{X_k^1\}_{k \le n}, \{A_k^2\}_{k<n}\right] .$$

2.2.4. *Observing all actions bidirectionally.* Finally, we consider the case in which both agents, at each time period, observe their own private signals as well as the other's action. Hence D_n^i, the action of agent i at period n, is given recursively by

$$D_n^i := \operatorname*{arg\,max}_{\theta \in \{+1,-1\}} \mathbb{P}\left[\Theta = \theta \,\middle|\, \{X_k^i\}_{k \le n}, \{D_k^{3-i}\}_{k<n}\right] .$$

Note that here and below $3-i$ is the "other agent"; if $i = 1$ then $3 - i = 2$ and vice versa. D_n^i is hence the action agent i would take when she could observe D_k^{3-i} *at all previous time periods* k in addition to her own signal.

Let

$$G_n = \bigcap_{k \le n, i \in \{1,2\}} \{D_k^i \neq +1\} .$$

This is the event that both agents do not take the action $+1$, in all time periods. Conditioned on $\Theta = +1$, G_n implies that they both choose the wrong action in all periods up to n.

2.3. **Asymptotics.** We are interested in the probability that an agent's best estimate does not equal the state of the world Θ. Since this probability vanishes exponentially fast in n, we scale the probabilities to extract the exponential rate of vanishing. Specifically, we define

$$a_p := \lim_{n\to\infty} -\frac{1}{n} \log \mathbb{P}\left[A_n^i \neq \Theta\right] .$$

Note that by symmetry, a_p is not a function of i, and thus depends only on p. Continuing the convention of using lowercase letters for rate functions, we let

$$b_p := \lim_{n\to\infty} -\frac{1}{n} \log \mathbb{P}\left[B_n^1 \neq \Theta\right]$$
$$c_p := \lim_{n\to\infty} -\frac{1}{n} \log \mathbb{P}\left[C_n^1 \neq \Theta\right] .$$

We also define

$$\overline{d_p} := \limsup_{n\to\infty} -\frac{1}{n} \log \mathbb{P}\left[D_n^i \neq \Theta\right]$$
$$\overline{g_p} := \limsup_{n\to\infty} -\frac{1}{n} \log \mathbb{P}\left[G_n | \Theta = +1\right] .$$

Note that it is not immediate that the limits a_p, b_p and c_p exist; we prove that they indeed do. We were not able to do the same in the case of d_p and g_p, and hence chose to consider the limits superior, which always exist.

Since A_n^1 is measurable with respect to the σ-algebra that defines B_n^1, while B_n^1 requires less information than C_n^1, we immediately conclude C_n^1 is more likely to be correct than B_n^1, which is better than A_n^1. Due to the negative sign in the definition, we see that $a_p \leq b_p \leq c_p$. We show that the first inequality is strict, while the second is actually an equality: $a_p < b_p = c_p$. It seems *a priori* difficult to guess the relation of $\overline{d_p}$ to these numbers. As it turns out, we show that $\overline{d_p} < b_p$; agent 1 learns more slowly if both agents observe each other's actions, as compared to the case in which only she observes the other's last action. We show this by showing that $\overline{g_p} < b_p$; it follows that $\overline{d_p} \leq \overline{g_p}$ from the definition of G_n, as the event that, conditioned on $\Theta = +1$, both agents are wrong at all periods.

2.4. **Main results.** For the reader's convenience, we gather all of our main results in Theorem 1 below. To state this theorem we will need to define Kullback-Leibler Divergence. For notational convenience, let $q := 1 - p$ and likewise $q' = 1 - p'$ and $\hat{q} = 1 - \hat{p}$. Letting μ and ν be two measures with the same, finite support, we recall that the Kullback-Leibler divergence is defined as

$$D_{KL}(\nu \,||\, \mu) := \sum_i \nu(i) \log \frac{\nu(i)}{\mu(i)} \,.$$

If ν is the Bernoulli distribution which assigns probability p' to $+1$ and q' to -1 and μ is the Bernoulli distribution which assigns probability p to 1, we will slightly abuse notation, and refer to the divergence as $D_{KL}(p' \,||\, p)$. Expanding this out explicitly, we see that

$$D_{KL}(p' \,||\, p) = p' \log \frac{p'}{p} + q' \log \frac{q'}{q} \,.$$

Fixing p, the function $D_{KL}(\cdot||p)$ is nonnegative, has a unique zero at p, and is continuous and strictly convex (see, e.g., Cover and Thomas [2012]).

Theorem 1 (The Asymptotic Rate of Learning)**.**

(1) a_p*, the rate of learning in the no observation case is given by*

$$a_p = D_{KL}\left(1/2 \,||\, p\right) \,.$$

(2) b_p*, the rate of learning when unilaterally observing the last action of the other agent equals* c_p*, the rate when observing all her past signals, and is given by*

$$b_p = c_p = D_{KL}(p^* \,||\, p) + a_p \,,$$

where $t^* = a_p / \log(p/q)$ *and* $p^* = \frac{1}{2}(1 + t^*)$*. Thus* $b_p = c_p > a_p$.

(3) $\overline{d_p}$*, the rate of learning when observing actions bidirectionally is bounded by* $\overline{g_n}$*, which in turn is strictly less than* b_p*:*

$$\overline{d_p} \leq \overline{g_p} = D_{KL}(1 - \hat{p} \,||\, p) < b_p \,.$$

Here $\hat{p}$ is the unique solution to $2D_{KL}(\hat{p}\,||\,p) = D_{KL}(1-\hat{p}\,||\,p)$ satisfying $\hat{p} < p$.

We note that while (1) follows almost immediately from Sanov's Theorem, (2) requires more detailed analysis, and (3) constitutes the main technical effort of this paper. The results in this theorem appear as separate theorems in the following sections.

2.5. **Information transmission efficiency.** An alternative way of quantifying the amount of information revealed by an agent's actions is what we call the *information transmission efficiency*. Consider agent 2 observing agent 1's actions, and imagine that agent 2 was instead offered to observe some fraction of agent 1's private signals. What fraction of the signals would agent 2 need to observe in order to have, in the long run, the same or better probability of choosing the correct actions?

To define this term formally, consider the case that agent 1, instead of observing agent 2's actions, observes some fraction $\alpha \in [0,1]$ of agent 2's private signals. Specifically, let agent 1, at period n, observe her own private signals $\{X_k^1\}_{k\leq n}$, as well as the first $\lfloor \alpha n \rfloor$ private signals of agent 2, $\{X_k^2\}_{k\leq \alpha n}$. Note that it does not matter which $\lfloor \alpha n \rfloor$ signals the agent is allowed to observe, since they are i.i.d., and thus equally informative. We denote by H_n^α the best guess of agent 1 when she observes the other agent's $\lfloor \alpha n \rfloor$ signals in addition to her own signals:

$$H_n^\alpha = \operatorname*{arg\,max}_{\theta \in \{+1,-1\}} \mathbb{P}\left[\Theta = \theta \,\middle|\, \{X_k^1\}_{k\leq n}, \{X_k^2\}_{k\leq \alpha n}\right] .$$

We will call a sequence of random variables $\{I_n\}_n$ an *information structure* if I_k is a function of the private signals observed (by both agents) until time k. For example, in the unidirectional case where agent 1 observes agent 2's actions, $I_n = \{A_k^2\}_{k<n}$ is the relevant information structure. In the bidirectinal case where both observe each other's actions, $I_n = \{D_k^2\}_{k<n}$ is the relevant informational structure.

Definition 2 (Information Transmission Efficiency)**.** *The transmission efficiency $\alpha(I)$ of an information structure $\{I_n\}_n$ is the supremum over $\bar{\alpha} \in [0,1]$ such that for all sufficiently large n*

$$\max_{\theta \in \{+1,-1\}} \mathbb{P}\left[\theta = \Theta \middle| \{X_k^1\}_{k\leq n}, I_n\right] > \mathbb{P}\left[H_n^{\bar{\alpha}} = \Theta\right] .$$

That is, $\alpha(I)$ is the largest fraction of private signals of agent 1 such that observing this fraction of signals, and agent 2's signals, results in a lower probability of choosing the correct action, as compared to choosing it given agent 2's signals and I_n.

A particular reason to consider asymptotic rates is the following connection to information transmission efficiency.

Proposition 3. *The information transmission efficiency is given by*

$$\alpha(I) = \frac{\beta}{a_p} - 1\,,$$

where β is the minimal asymptotic rate of learning of agent 1 under the information structure I

$$\beta = \liminf_{n\to\infty} -\frac{1}{n}\log\left(\max_{\theta\in\{+1,-1\}} \mathbb{P}\left[\theta \neq \Theta \middle| \{X_k^1\}_{k\le n}, I_n\right]\right)\,.$$

Proof of Proposition 3. As signals are i.i.d., an agent is equally likely to be correct if she observed $\bar{\alpha}n$ of the other agent's signals, in addition to her own n signals and in the situation where she observed $(1+\bar{\alpha})n$ of her own signals. It thus follows from the first result of Theorem 1 that the probability of being wrong when observing a fraction $\bar{\alpha}$ of the other agents signals is given by[5]

$$\mathbb{P}\left[H_n^{\bar{\alpha}} \neq \Theta\right] = e^{-a_p(1+\alpha)n+o(n)}. \tag{1}$$

By the definition of β we have that

$$\max_{\theta\in\{+1,-1\}} \mathbb{P}\left[\theta \neq \Theta \middle| \{X_k^1\}_{k\le n}, I_n\right] \le e^{-\beta\, n+o(n)}. \tag{2}$$

Thus, for all sufficiently large n it follows that (1) $>$ (2) if $a_p(1+\bar{\alpha}) < \beta$. Rearranging for $\bar{\alpha}$ gives that (1) $>$ (2) for all $\bar{\alpha} < \frac{\beta}{a_p} - 1$ and taking the supremum over this set yields

$$\alpha(I) \ge \frac{\beta}{a_p} - 1\,.$$

Conversely, by the definition of the limit inferior, we have that (1) $<$ (2) if $a_p(1+\alpha) > \beta$ for some arbitrary large n and thus $\alpha(I) \le \frac{\beta}{a_p} - 1\,.$ ☐

2.5.1. *Information Transmission with Unidirectional Information.* Given our results on the asymptotic rate b_p, c_p from Theorem 1 it follows from Proposition 3 that the information transmission efficiency if either the last or all actions of the other agent are observed is given by

$$\frac{b_p}{a_p} - 1 = \frac{c_p}{a_p} - 1 = \frac{D_{KL}(p^*\,||\,p)}{D_{KL}(1/2\,||\,p)}.$$

As shown in Figure 1, the information transmission efficiency depends on p and ranges between $1/2 = 50\%$ and $9/16 \approx 57\%$. Thus, an agent is more likely to be correct if she observes a fraction higher than $9/16$ of the other agent signals compared to observing all the other agents actions. Conversely, she is less likely to be correct if she observes less than half of the other agent's signals.

[5]Recall that a function $f : \mathbb{N} \to \mathbb{R}$ satisfies $f(n) = o(n)$ if and only if $\lim_{n\to\infty} \frac{f(n)}{n} = 0\,.$

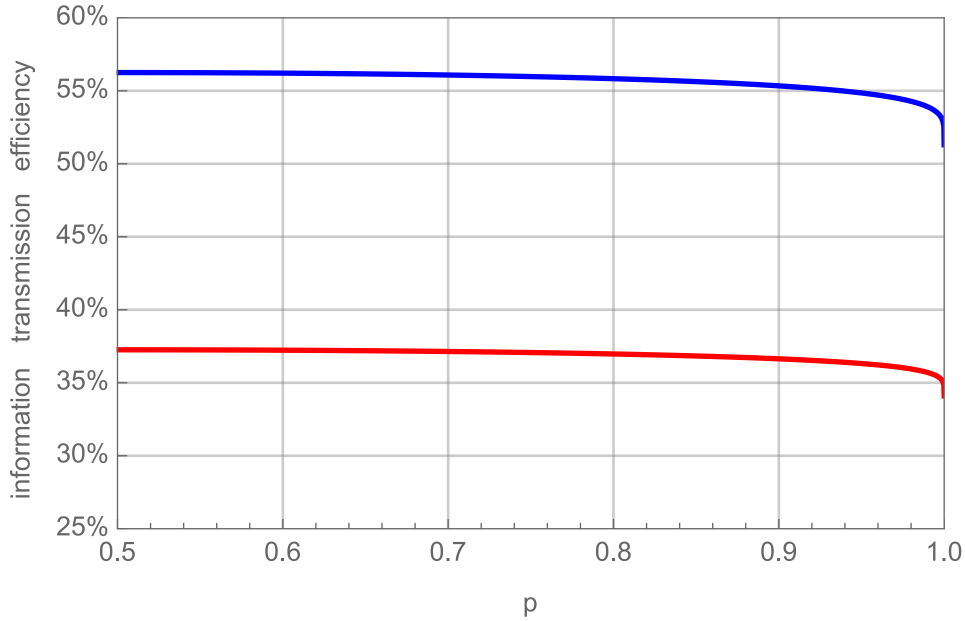

FIGURE 1. The information transmission efficiency, as a function of p, in the case of unidirectional observations (blue, higher line) and the upper bound for the case of bidirectional observations (red, lower line).

2.5.2. *Information Transmission with Bidirectional Information.* Our bound on $\overline{d_p}$ implies that the efficiency in the bidirectional observation case is bounded from above by

$$\frac{\overline{d}_p}{a_p} - 1 \,.$$

Especially, the information transmission efficiency is lower than in the unidirectional case, and is in particular at most 37%, as can be shown by a simple numerical calculation. Figure 1 illustrates the information transmission efficiency in the unidirectional and the bidirectional case and shows that the bidirectional information exchange leads to an additional loss in information transmission efficiency of at lease 18%.

In the following sections we derive the results about asymptotic rates stated in Theorem 1.

3. No observations

In this section we study the case that each agent observes only her own actions

$$A_n^i := \operatorname*{arg\,max}_{\theta \in \{+1,-1\}} \mathbb{P}\left[\Theta = \theta \,\middle|\, \{X_k^i\}_{k \le n}\right] \,.$$

Her exponential rate of learning in this setting is

$$a_p := \lim_{n \to \infty} -\frac{1}{n} \log \mathbb{P}\left[A_n^i \neq \Theta\right] \,.$$

We show that

Theorem 4. $a_p = D_{KL}(1/2\,||\,p)$.

The proof is completely standard, but we provide it here in its entirety in order to introduce some tools that will be useful later. An important tool in our proofs is Sanov's Theorem (Sanov [1957], or see, e.g., Dembo and Zeitouni, 1998, Theorem 2.1.10). To motivate it, note that when $\Theta = +1$, the average signal $S_n^i/n = \frac{1}{n}\sum X_k^i$ tends to its expectation $2p-1$, by the law of large numbers. Because $\mathbb{E}\left[\exp(X_i^k)\right] < \infty$, Markov's inequality implies that S_n^i/n will significantly deviate from $2p-1$ with exponentially small probability in n.

By the Central Limit Theorem, S_n^i/n will significantly deviate from $2p-1$ with exponentially small probability. Sanov's Theorem is a calculation of this exponential vanishing rate.

Theorem 5 (Sanov's Theorem)**.** *For any $\underline{p} < \overline{p}$ we have that*

$$\lim_{n\to\infty} -\log\frac{1}{n}\mathbb{P}^+[2\underline{p}-1 \le S_n^i/n \le 2\overline{p}-1] = \begin{cases} D_{KL}(\underline{p}\,||\,p) & \text{when } p < \underline{p} \\ 0 & \text{when } \underline{p} \le p \le \overline{p} \\ D_{KL}(\overline{p}\,||\,p) & \text{when } \overline{p} < p \end{cases}.$$

We are interested in calculating

$$a_p := \lim_{n\to\infty} -\frac{1}{n}\log\mathbb{P}\left[A_n^1 \neq \Theta\right].$$

where A_n^1 is agent 1's best guess of Θ, given her private signals $\{X_k^1\}_{k\le n}$. The following claim is an easy consequence of the definition of A_n^1:

Claim 6.

$$A_n^i = \operatorname{sgn} S_n^i.$$

Proof. The log-likelihood ratio of Θ, given $\{X_k^1\}_{k\le n}$, is

$$L_n^A := \log\frac{\mathbb{P}\left[\Theta = +1\,|\,\{X_k^1\}\right]}{\mathbb{P}\left[\Theta = -1\,|\,\{X_k^1\}\right]} = \log\frac{\mathbb{P}^+[\{X_k^1\}]}{\mathbb{P}^-[\{X_k^1\}]},$$

where the second equality is Bayes' Rule.

Since the variables $\{X_k^1\}$ are conditionally independent, we can explicitly compute the right hand side to be

$$L_n^A = S_n^1 \cdot \log(p/q).$$

Note that $L_n^A = 0$ if and only if $\mathbb{P}\left[\Theta = +1|\{X_k^1\}\right] = \frac{1}{2}$, and furthermore the latter is greater then half (resp. less then half) if the former is positive (resp. negative). Hence $A_n^1 = \operatorname{sgn} L_n^A$, and so $A_n^1 = \operatorname{sgn} S_n$. $\square$

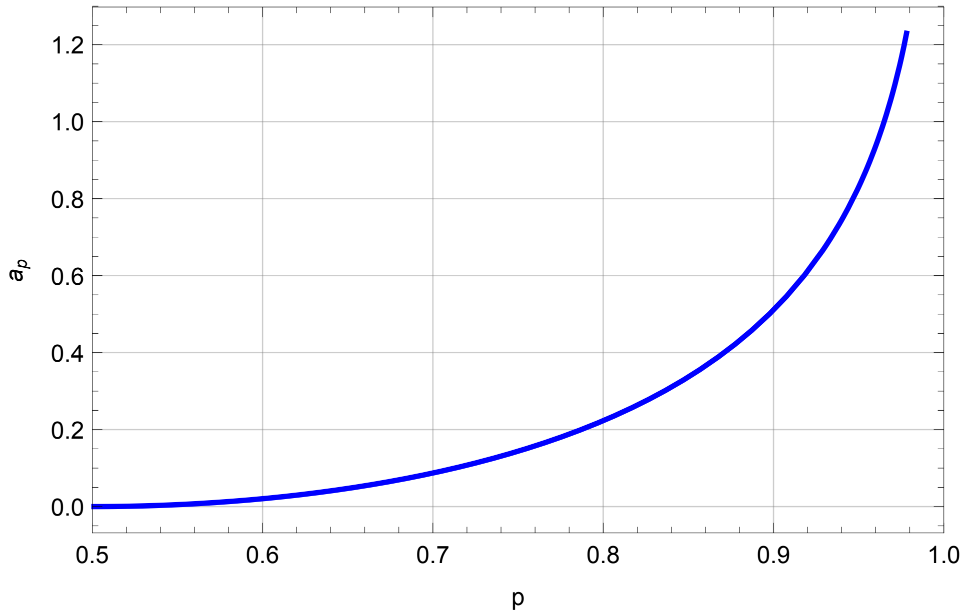

FIGURE 2. When given signals with strength p, an agent who does not observe the other agent chooses the wrong action at time n with probability $e^{-a_p \cdot n + o(n)}$, where $a_p = D_{KL}({}^1\!/\!{}_2 \,||\, p)$.

By symmetry, $\{A_n^1 \neq \Theta\}$ has twice the probability of $\{A_n^1 \neq +1, \Theta = +1\} = \{S_n^1 \leq 0, \Theta = +1\}$. Hence, by Claim 6 and by conditioning on $\Theta = +1$, we can conclude that

$$\mathbb{P}\left[A_n^1 \neq \Theta\right] = \mathbb{P}^+[S_n^1 \leq 0] = \mathbb{P}^+[-1 \leq S_n^1/n \leq 0]\,, \tag{3}$$

where the second equality follows from the fact that $\mathbb{P}\left[-1 \leq S_n^1/n\right] = 1$. We remind the reader that $\mathbb{P}^+[\cdot] = \mathbb{P}\left[\cdot|\Theta = +1\right]$.

Since the X_k^i's are independent conditional on Θ, we can now apply Sanov's Theorem, substituting $\underline{p} = 0$ and $\overline{p} = \frac{1}{2}$:

$$a_p = \lim_{n\to\infty} -\frac{1}{n} \log \mathbb{P}\left[A_n^1 \neq \Theta\right] = \lim_{n\to\infty} -\frac{1}{n} \log \mathbb{P}^+[-1 \leq S_n^1/n \leq 0] = D_{KL}({}^1\!/\!{}_2 \,||\, p)\,, \tag{4}$$

where the last equality is Sanov's Theorem (for $p > \frac{1}{2}$), and the second equality is a consequence of (3). This completes the proof of Theorem 4; this is also part (1) of Theorem 1. Figure 2 illustrates how a_p varies with p.

4. OBSERVING THE FINAL ACTION

We now move on to analyze the setting in which agent 2 only observes her private signal, but agent 1 also observes agent 2's last action. Agent 2's actions are hence given by

$$A_n^2 := \operatorname*{arg\,max}_{\theta\in\{+1,-1\}} \mathbb{P}\left[\Theta = \theta \,\middle|\, \{X_k^2\}_{k\leq n}\right]\,,$$

and agent 1's actions are given by

$$B_n^1 := \underset{\theta\in\{+1,-1\}}{\arg\max}\ \mathbb{P}\left[\Theta=\theta \,\middle|\, \{X_k^1\}_{k\le n}, A_{n-1}^2\right] .$$

We calculate

$$b_p := \lim_{n\to\infty} -\frac{1}{n}\log\mathbb{P}\left[B_n^1\neq\Theta\right],$$

which is agents 1's rate of learning in this setting. As in the previous section, we can analyze the event $\{B_n^1\neq+1,\Theta=+1\}$ by symmetry, and we use the log-likelihood ratio to find when $B_n^1\neq+1$. Given $\{X_k^1\}$ and $\{A_n^2\}$, the relevant log-likelihood ratio is

$$L_n^B := \log\frac{\mathbb{P}^+[\{X_k^1\},A_{n-1}^2]}{\mathbb{P}^-[\{X_k^1\},A_{n-1}^2]} .$$

Analogously to A_n^1 and L_n^A in the previous section, B_n^1 is equal to the the sign of L_n^B. Using the conditional independence of the agents' signals, we see that

$$L_n^B = \log\frac{\mathbb{P}^+[\{X_k^1\}]}{\mathbb{P}^-[\{X_k^1\}]} + \log\frac{\mathbb{P}^+[A_{n-1}^2]}{\mathbb{P}^-[A_{n-1}^2]} . \tag{5}$$

The first logarithm is equal to $S_n^1\log(p/q)$, as it was in the previous section. As for the second expression, it is easy to see that

$$\begin{aligned}\log\frac{\mathbb{P}^+[A_{n-1}^2]}{\mathbb{P}^-[A_{n-1}^2]} &= \mathbf{1}_{\{A_{n-1}^2=+1\}}\log\frac{\mathbb{P}\left[A_{n-1}^2=\Theta\right]}{\mathbb{P}\left[A_{n-1}^2\neq\Theta\right]} + \mathbf{1}_{\{A_{n-1}^2=-1\}}\log\frac{\mathbb{P}\left[A_{n-1}^2\neq\Theta\right]}{\mathbb{P}\left[A_{n-1}^2=\Theta\right]}\\ &= A_{n-1}^2\log\frac{\mathbb{P}\left[A_{n-1}^2=\Theta\right]}{\mathbb{P}\left[A_{n-1}^2\neq\Theta\right]} .\end{aligned}$$

It thus follows that the log-likelihood ratio L_n^B is given by

$$L_B^n = \operatorname{sgn}(S_n^1)\,|S_n^1|\,\log(p/q) + A_{n-1}^2\log\frac{\mathbb{P}\left[A_{n-1}^2=\Theta\right]}{\mathbb{P}\left[A_{n-1}^2\neq\Theta\right]} .$$

As the action taken by agent 1 is given by the sign of the log-likelihood ratio, we have the following proposition describing agent 1's behavior:

Proposition 7. *Let the threshold signal t_n^* be given by*

$$t_n^* = \frac{1}{n}\log\frac{\mathbb{P}\left[A_{n-1}^2=\Theta\right]}{\mathbb{P}\left[A_{n-1}^2\neq\Theta\right]}\log(p/q)^{-1} . \tag{6}$$

Then agent 1's action B_n^1 is given by:

$$B_n^1 = \begin{cases}\operatorname{sgn} S_n & \text{if } |S_n/n| > t_n^* \text{ or } A_{n-1}^2 = 0\\ A_{n-1}^2 & \text{if } |S_n/n| < t_n^* \text{ and } A_{n-1}^2\neq 0\\ \operatorname{sgn}\left(A_{n-1}^2+\operatorname{sgn} S_n\right) & \text{if } |S_n/n| = t_n^*\end{cases}$$

In the first case, agent 1 has such a strong private signal that she ignores agent 2's observed action, follows her own signal, and thus and chooses $B^1_n = \operatorname{sgn} S_n$. In the second case agent 2 has the weakest possible private signal (since $S^2_{n-1} = 0$), and so agent 1 again follows her own signal and chooses $B^1_n = \operatorname{sgn} S_n$. In the third case, agent 1 has a weak signal but agent 2 does not, and so agent 1 follows agent 2's action and chooses $B^1_n = A^2_{n-1}$. In the fourth case, agent 1's private signal is as strong as agent 2's, and so agent 1's action will depend on both signals. We formally prove this proposition in the appendix.

The following is an immediate corollary.

Corollary 8. *Agent 1's action B^1_n is equal to $+1$,* **unless** *one of the three following disjoint events occur:*

(1) $S_n/n < -t^*_n$.
(2) $-t^*_n \le S_n/n \le t^*_n$ *and* $A^2_{n-1} = -1$.
(3) $-t^*_n \le S_n/n \le 0$ *and* $A^2_{n-1} = 0$.

The threshold signal t^*_n defined in (6) plays a crucial role in the description of the behavior of agent 1. While it is hard to calculate the threshold t^*_n explicitly for small n, the asymptotic behavior follows easily from our results of the previous section.

Lemma 9. *The asymptotic threshold signal $t^* = \lim_{n\to\infty} t^*_n$ is given by*

$$t^* = \frac{a_p}{\log(p/q)}\,.$$

Proof. Since the probability that agent 2 chooses the right action converges to one (i.e., $\lim_n \mathbb{P}\left[A^2_{n-1} = \Theta\right] = 1$), and using our characterization of the asymptotic rate of learning in the single agent case derived in (4), we have that

$$\begin{aligned}
\lim_{n\to\infty} t^*_n &= \frac{1}{\log(p/q)} \lim_{n\to\infty} \frac{1}{n} \log \frac{\mathbb{P}\left[A^2_{n-1} = \Theta\right]}{\mathbb{P}\left[A^2_{n-1} \neq \Theta\right]} \\
&= \frac{1}{\log(p/q)} \lim_{n\to\infty} \frac{1}{n} \left(\log \mathbb{P}\left[A^2_{n-1} = \Theta\right] - \log \mathbb{P}\left[A^2_{n-1} \neq \Theta\right]\right) \\
&= \frac{1}{\log(p/q)} \lim_{n\to\infty} -\frac{1}{n-1} \log \mathbb{P}\left[A^2_{n-1} \neq \Theta\right] \cdot \lim_{n\to\infty} \frac{n-1}{n} = \frac{a_p}{\log(p/q)}\,.
\end{aligned}$$ ∎

As a consequence of Corollary 7 the probability that agent 1 makes an incorrect guess is given by

$$\begin{aligned}
\mathbb{P}\left[B^1_n \neq \Theta\right] &= \mathbb{P}^+[B^1_n \neq +1] \\
&= \mathbb{P}^+[S_n/n < -t^*_n] \\
&\quad + \mathbb{P}^+[-t^*_n \le S_n/n \le t^*_n] \cdot \mathbb{P}^+[A^2_n = -1] \\
&\quad + \mathbb{P}^+[-t^*_n \le S_n/n \le 0] \cdot \mathbb{P}^+[A^2_n = 0]\,.
\end{aligned} \tag{7}$$

As the asymptotic rate of a sum is the minimum of the rates of the summands, we calculate the rate of the above three summands separately. First, if we denote

$$p^* = \tfrac{1}{2}(1 + t^*)\,,$$

as earlier, and $q^* = 1 - p^*$, it follows from a generalized version of Sanov's Theorem we prove in Corollary 23 in the Appendix that

$$\lim_{n\to\infty} -\frac{1}{n}\log \mathbb{P}^+[S_n^1/n < -t_n^*] = \lim_{n\to\infty} -\frac{1}{n}\log \mathbb{P}^+[S_n^1/n < -t^*] = D_{KL}(q^*||\,p)\,.$$

For the second event it follows from the generalized version of Sanov's Theorem that the asymptotic rate is given by

$$\begin{aligned}
&\lim_{n\to\infty} -\frac{1}{n}\log\left(\mathbb{P}^+[-t_n^* < S_n^1/n < t_n^*]\cdot\mathbb{P}^+[A_n^2 = -1]\right)\\
&\quad= \lim_{n\to\infty} -\frac{1}{n}\log \mathbb{P}^+[S_n^1/n < t^*] + \lim_{n\to\infty} -\frac{1}{n}\log \mathbb{P}^+[S_n^2 < 0]\\
&\quad= D_{KL}(p^*||p) + D_{KL}(1/2||p)\,.
\end{aligned}$$

For the third event the asymptotic rate satisfies

$$\begin{aligned}
&\liminf_{n\to\infty} -\frac{1}{n}\log\left(\mathbb{P}^+[-t_n^* < S_n^1/n < 0]\cdot\mathbb{P}^+[A_n^2 = 0]\right)\\
&\quad= \lim_{n\to\infty} -\frac{1}{n}\log \mathbb{P}^+[S_n^1/n \le 0] + \lim_{n\to\infty} -\frac{1}{n}\log \mathbb{P}^+[S_n^2 = 0]\,.
\end{aligned}$$

The first limit is equal to $D_{KL}(1/2||p)$, by Sanov's Theorem. The second is likewise equal to $D_{KL}(1/2||p)$; this follows from the fact that using an elementary combinatorial argument, $\mathbb{P}^+[S_n^2 = 0]$ can be calculated explicitely:

$$\mathbb{P}^+[S_n^2 = 0] = \binom{2n}{n} p^n q^n.$$

Hence the rate of the last summand in (7) is $2D_{KL}(1/2||p)$. Gathering the rates of the three summands and taking the minimum, we have that

$$\begin{aligned}
b_p &= \lim_{n\to\infty} -\frac{1}{n}\log \mathbb{P}\left[B_n^1 \neq \Theta\right]\\
&= \min\{\, D_{KL}(q^*||p)\,,\ D_{KL}(p^*||p) + D_{KL}(1/2||p)\,,\ 2D_{KL}(1/2||p)\}.
\end{aligned}$$

Now, $2D_{KL}(1/2||p) > D_{KL}(p^*||p) + D_{KL}(1/2||p)$, since $1/2 < p^* < p$, and by the monotonicity of $D_{KL}(\cdot||p)$. Hence

$$b_p = \min\{\, D_{KL}(q^*||p)\,,\ D_{KL}(p^*||p) + D_{KL}(1/2||p)\}.$$

Perhaps surprisingly, it so happens that these two numbers are equal: $D_{KL}(q^*||p) = D_{KL}(p^*||p) + D_{KL}(1/2||p)$; this can be easily shown by substituting the definitions of p^*, q^* and the Kullback-Leibler divergence.

To understand the intuitive reason for this, is it important to understand a general principle of large deviations, which follows immediately from Sanov's Theorem, and which we call the "smallest possible mistake" principle: conditioned on an agent being wrong - i.e., $S_n/n \leq \overline{p} < p$ - her mistake will be the smallest possible: i.e., $S_n/n \approx \overline{p}$. Formally,

Theorem 10 (Smallest possible mistake)**.** *For all $\overline{p} < p$ and $\epsilon > 0$,*

$$\lim_{n\to\infty} \mathbb{P}^+ \left[\overline{p} - \epsilon \leq S_n^1/n \;\middle|\; S_n^1/n \leq \overline{p}\right] = 1.$$

Now, condition, as above, on $\Theta = +1$, and consider the two events whose rates are $D_{KL}(q^*||p)$ and $D_{KL}(p^*||p) + D_{KL}(1/2||p)$. The first is that $S_n/n < -t^*$, and the second is that $|S_n/n| \leq t^*$ and $A_{n-1}^2 = -1$. By the argument above, the first implies that $S_n/n \approx -t^*$, and that second implies that $S_n/n \approx t^*$. But this is precisely the threshold in which agent 1 is indifferent in choosing between following her own signal and that of agent 2! Clearly this is because both signals have the same probability of being wrong, by the optimality of agent 1's choice. Thus these two events have the same rate.

We have therefore shown that $b_p = D_{KL}(q^*||p) = D_{KL}(p^*||p) + D_{KL}(1/2||p)$. Since $a_p = D_{KL}(1/2||p)$, we have completed the proof of the main theorem of this section; this result also appears in part (2) of Theorem 1.

Theorem 11. $b_p = D_{KL}(p^* \,||\, p) + a_p$.

The relation between b_p and p is illustrated in Figure 3.

As mentioned in the introduction, b_p varies between $1.5a_p = 24/16a_p$ and $1.5626a_p = 25/16a_p$. This is illustrated in Figure 4.

5. Observing all actions unidirectionally

In this section we consider the setting in which agent 1 observes each action of agent 2, rather than just the last one, as in the previous section. Hence agent 2's actions are again given by

$$A_n^2 := \operatorname*{arg\,max}_{\theta\in\{+1,-1\}} \mathbb{P}\left[\Theta = \theta \,\middle|\, \{X_k^2\}_{k\leq n}\right] \,,$$

and agent 1's actions are given by

$$C_n^1 := \operatorname*{arg\,max}_{\theta\in\{+1,-1\}} \mathbb{P}\left[\Theta = \theta \,\middle|\, \{X_k^1\}_{k\leq n}, \{A_k^2\}_{k<n}\right] \,.$$

We show that here, agent 1's learning rate

$$c_p := \lim_{n\to\infty} -\frac{1}{n} \log \mathbb{P}\left[C_n^1 \neq \Theta\right]$$

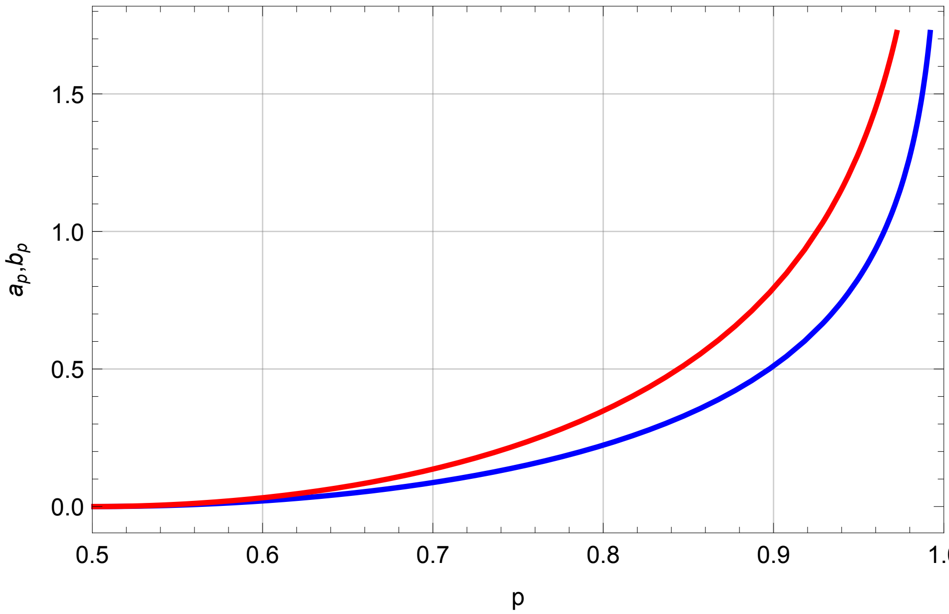

FIGURE 3. When given signals with strength p, an agent who observes the other agent chooses the wrong action at time n with probability $e^{-b_p \cdot n + o(n)}$, where $b_p = D_{KL}(p^* \,||\, p) + a_p$; both b_p (red, higher line) and a_p (blue, lower line) are shown.

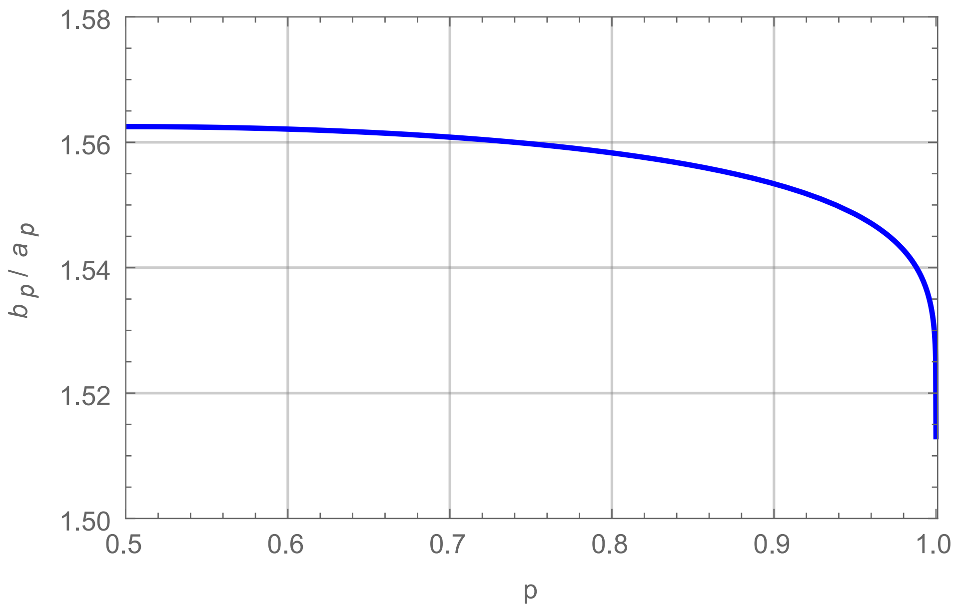

FIGURE 4. b_p varies between $1.5a_p$ and $1.5625a_p$.

is the same as in the previous section, despite the fact that she has strictly more information available to her. Of course, her probability of choosing the wrong action is always smaller than in the previous section. However, the *exponential rate* is identical.

We wish to show that $c_p = b_p$. Since B_n^i is measurable with respect to the σ-algebra generated by $\{X_k^1\}$ and $\{A_k^2\}$, we know that C_n^i is more likely to correctly estimate Θ than B_n^i. Therefore, the error probability decreases, and $b_p \leq c_p$. The proof will be complete if we knew that $c_p \leq b_p$. Probabilistically, this would follow if there was an event F_n that implied that agent 1 chooses an action $C_n^1 \neq +1$, and conditioned on $\Theta = +1$ had the asymptotic rate function b_p - i.e.

$$\liminf_{n\to\infty} -\frac{1}{n}\log \mathbb{P}^+[F_n] \leq b_p + \varepsilon\,,$$

for any $\varepsilon > 0$ sufficiently small. We will show that the event F_n that agent 1's private signal is not too positive, and that agent 2 always chooses the wrong action -1 satisfies these requirements

$$F_n = \{S_n^1 < t^*(1-\varepsilon)n, \{A_k^2 = -1\}_{k\leq n}\}\,.$$

Note that the second part of the definition of F_n can be written as $\{S_k^2 < 0\}_{k\leq n}$, by Claim 6.

First, we calculate the probability of F_n, conditioned on $\Theta = +1$. This can be decomposed into a product, because of the conditional independence of the two agents' private signals:

$$\mathbb{P}^+[F_n] = \mathbb{P}^+[S_n^1 < t^*(1-\varepsilon)n]\cdot \mathbb{P}^+[\{S_k^2 < 0\}_{k\leq n}]\,. \tag{8}$$

The first probability is easy to calculate via Sanov's Theorem. For the second expression, we state a corollary of the reflection principle, sometimes referred to as "Bertrand's Ballot Theorem" (Bertrand [1887], or see, e.g., [Durrett, 1996, pg. 198]):

Theorem 12 (Bertrand's Ballot Theorem)**.** *For any negative integer $x < 0$ and any $\theta \in \{+1,-1\}$*

$$\mathbb{P}\left[\{S_k^2 < 0\}_{k<n}, S_n = x \,\middle|\, \Theta = \theta\right] = \frac{|x|}{n}\mathbb{P}\left[S_n = x \,|\, \Theta = \theta\right].$$

As a consequence of the Ballot Theorem,

$$\mathbb{P}^+[\{S_k^2 < 0\}_{k\leq n}] \geq \frac{1}{n}\mathbb{P}^+[S_n^2 < 0]\,.$$

Substituting this into (8), the expression for the probability of F_n, we find that

$$-\frac{1}{n}\log \mathbb{P}^+[F_n] \leq \frac{\log n}{n} - \frac{1}{n}\left(\log \mathbb{P}^+[S_n^1/n < t^*(1-\varepsilon)] + \log \mathbb{P}^+[S_n^2 < 0]\right)\,.$$

If we take a limit inferior of both sides, the $\log n/n$ term vanishes, and we are left with an expression that is nearly identical to the one we found in the previous section. Hence by the same considerations the rate function of F_n is bounded below by

$$D_{KL}(p^* - \tfrac{1}{2}t^*\varepsilon \,||\, p) + D_{KL}\left({}^1\!/\!_2 \,||\, p\right)\,.$$

Furthermore, we have shown before that

$$D_{KL}(p^* \,||\, p) + D_{KL}\left({}^1\!/\!_2 \,||\, p\right) = b_p\,.$$

It therefore remains to be shown that agent 1 takes the action -1 if she observed agent 2 taking the action -1 in all prior periods and if she would have taken the action -1 if she only observed that agent 2 took the action -1 in the last period (and no information about 2's prior actions).

Proposition 13. *F_n implies $\{C_n^1 \neq +1\}$.*

We prove Proposition 13 in the appendix. Intuitively, seeing agent 2 take the action -1 more often can only be evidence of the fact that agent 2 observed the signal -1 more often, which makes the state -1 more likely from agent 1's perspective.

The proof shows this by first arguing that the event that agent 2 takes the wrong action in all period $k < n$ is less likely than that she takes the wrong action in period $k = n - 1$.

We thus conclude that

Theorem 14. *$c_p = b_p$.*

This also appears in part (2) of Theorem 1.

6. Observing all actions bidirectionally

In this section we consider the case that each agent, at each time period, observes both her private signal and the other agent's action, which, since the agents are myopic, is the other agent's best estimate of the state of nature. That is, agent i's action at time n is given by

$$D_n^i := \operatorname*{arg\,max}_{\theta \in \{+1,-1\}} \mathbb{P}\left[\Theta = \theta \,\middle|\, \{X_k^i\}_{k \leq n}, \{D_k^{3-i}\}_{k<n}\right] .$$

We are interested in calculating

$$d_p := \lim_{n \to \infty} -\frac{1}{n} \log \mathbb{P}\left[D_n^i \neq \Theta\right] .$$

However, in trying to analyze this expression we encountered significant obstacles, and were indeed unable to even prove that the limit exists. These stem from the complexity of the recursive definition of D_n^i.

In lieu, we resort to studying the limit superior

$$\overline{d_p} := \limsup_{n \to \infty} -\frac{1}{n} \log \mathbb{P}\left[D_n^i \neq \Theta\right] ,$$

which is guaranteed to exist; note that by symmetry, $\overline{d_p}$ does not depend on i.

We show that $\overline{d_p} < c_p = b_p$ - that is, the probability of taking the wrong action in this setting is (for n large enough) exponentially higher than in the previous two. Thus we in fact do not lose much by resorting to the study of $\overline{d_p}$ rather than d_p. Still, the interesting questions remain: does the limit d_p exist, and can it be calculated? A natural conjecture is that d_p exists and is equal to $\overline{d_p}$.

As mentioned above, it is difficult to analyze the probability of the event $D^i_n \neq \Theta$, i.e., the event that agent i chooses the wrong action at time n, in the setting of bidirectional observations. We find, however, that an easier probability to calculate (or at least bound from below) is that of the event that *both* agents choose the wrong actions at *all* time periods $1, 2, \dots, n$. Accordingly, let

$$G_n = \bigcap_{k \leq n, i \in \{1,2\}} \{D^i_k \neq +1\}.$$

Conditioned on $\Theta = +1$, this is the event that both agents choose a wrong action in all time periods. Since G_n implies $\{D^1_n \neq +1\}$, it follows that

$$\overline{g_p} := \limsup_n -\frac{1}{n} \log \mathbb{P}^+[G_n] \geq \overline{d_p}.$$

We will therefore prove that $\overline{d_p} < b_p$ by showing that $\overline{g_p} < b_p$; that is, the probability of both agents choosing the wrong action in the bidirectional case is exponentially higher than that of agent 1 choosing the wrong action at time n, in the unidirectional case.

Now, the event G_n can be written as $G^1_n \cap G^2_n$, where G^i_n is the event that agent i chooses the wrong action at every period up to n. To calculate the probability of G_n, it would of course have been convenient if these two events were independent, conditioned on Θ. However, due to the fact that the agents' actions are strongly intertwined, G^1_n and G^2_n are not independent; given that agent 1 played -1 at all time periods, agent 2 is clearly more likely to do the same.

Perhaps surprisingly, it turns out that G_n can never-the-less be written as the intersection of two other independent events, $W^1_n \cap W^2_n$, where W^i_n depends only on the private signals of agent i. To define these events, we consider the perspective of on outside observer, who observes $\{D^i_l\}$ for $i = 1, 2$ and $l < k$, but has no access to the signals of either agents. This outside observer can calculate, given G_{k-1}, whether there exists a trajectory which would cause agent 1 (for example) to guess that $\Theta = +1$ in the k^{th} period, and, if so, what minimal value of S^1_k would imply this. We define that value as t_k, the "threshold" that S^1_k must be under to imply the event G_k. By symmetry, the same threshold applies to agent 2, and so S^2_k must also be under t_k. Thus W^i_n is the event that agent i's private signals are such that $S^i_k < t_k$ for all $k \leq n$. These are clearly conditionally independent events, and their intersection is G_n.

We now formalize this construction, defining t_k, W^1_k and W^2_k inductively. Let $t_1 = 0$ and let W^1_0 and W^2_0 be full measure events. For $k \geq 1$ and $i \in \{1, 2\}$, let

$$W^i_k = W^i_{k-1} \cap \{S^i_k \leq t_k\},$$

and for $k > 1$ let

$$t_k = -\frac{1}{\log(p/q)} \cdot \log \frac{\mathbb{P}^+[W_{k-1}^2]}{\mathbb{P}^-[W_{k-1}^2]}. \tag{9}$$

We prove the following claim in the Appendix.

Claim 15.

$$W_n^1 \cap W_n^2 = G_n.$$

Note that it follows that t_k is a non-negative number: the probability that both agents play -1 at all times conditioned on $\Theta = -1$ is clearly larger than the same probability, conditioned on $\Theta = +1$.

To analyze the events W_n^i, we start by studying t_n, and in particular its asymptotic behavior, as described by

$$\hat{t} = \liminf_n t_n/n.$$

Denote also $\hat{p} = \frac{1}{2}(1 + \hat{t})$ and $\hat{q} = 1 - \hat{p}$. By the definition of t_n we will equivalently need to understand the asymptotic conditional probabilities $\mathbb{P}^+[W_n^2]$ and $\mathbb{P}^-[W_n^2]$. We start by analyzing the latter, and (by an easy exercise) show

Claim 16. *For every $1/2 < p < 1$ there is a constant $C > 0$ such that, for all $k > 0$,*

$$C < \mathbb{P}^-[W_n^2].$$

We prove this in the appendix.

We next turn to understanding $\mathbb{P}^+[W_n^2]$. Now, conditioned on $\Theta = +1$, S_n^2/n will, with high probability, be approximately equal to $2p - 1$. W_n^2 is (roughly) the event that it is less than $2\hat{p} - 1$. Hence, if $\hat{p} < p$, our experience from Sanov's Theorem suggests that the probability of W_n^2, conditioned on $\Theta = +1$ should be approximately $e^{-D_{KL}(\hat{p}||p)n}$. Formally, we show that indeed

Proposition 17. *$\hat{p} > p$.*

and that

Proposition 18. *Let $\{n_i\}_i$ be a sequence such that $\lim_i t_{n_i}/n_i = \hat{t}$. Then*

$$\limsup_n -\frac{1}{k} \cdot \log \mathbb{P}^+[W_n^2] = \lim_i -\frac{1}{n_i} \cdot \log \mathbb{P}^+[W_{n_i}^2] = D_{KL}(\hat{p}||p)\,.$$

Both proofs appear in the appendix. We henceforth let $\{n_i\}_i$ be sucht that $\lim_i t_{n_i}/n_i = \hat{t}$.

Joining the latter proposition with Claim 16 yields that

$$\lim_i -\frac{1}{n_i} \cdot \log \frac{\mathbb{P}^+[W_{n_i}^2]}{\mathbb{P}^-[W_{n_i}^2]} = \lim_i -\frac{1}{n_i} \cdot \log \mathbb{P}^+[W_{n_i}^2] = D_{KL}(\hat{p}||p).$$

Therefore, if we divide both sides of the definition of t_n (9) by k and take limits along $\{n_i\}_i$, we arrive at

$$\hat{t} = \frac{D_{KL}(\hat{p}||p)}{\log(p/q)}\,. \tag{10}$$

Substituting $\hat{p} = (1+\hat{t})/2$ yields an equation for $\hat{t}$:

$$\hat{t} = \frac{D_{KL}({}^1\!/\!{}_2(1+\hat{t})||p)}{\log(p/q)}.$$

Note that $\hat{t} = 0$ is not a solution of this equation, and so $\hat{t} > 0$ and $\hat{p} > 1/2$. Furthermore, we know that $\hat{p} < p$. Therefore, and since $D_{KL}(\cdot||p)$ is decreasing on $[0,p)$

$$\hat{t} = \frac{D_{KL}(\hat{p}||p)}{\log(p/q)} < \frac{D_{KL}({}^1\!/\!{}_2||p)}{\log(p/q)} = t^*,$$

recalling the definition of t^*. Hence $q^* < \hat{q}$. Another interesting consequence of (10) is that

Lemma 19. $2D_{KL}(\hat{p}||p) = D_{KL}(\hat{q}||p)$.

This follows by elementary algebraic manipulations of (10), which we omit. Returning to G_n,

$$\mathbb{P}^+[G_n] = \mathbb{P}^+[W_n^1, W_n^2] = \mathbb{P}^+[W_n^1]^2.$$

Hence, by Proposition 18,

$$\overline{g_p} = \limsup_n -\frac{1}{k}\cdot\log\mathbb{P}^+[G_n] = 2D_{KL}(\hat{p}||p).$$

Thus, applying Lemma 19 yields

Theorem 20.

$$\overline{g_p} = D_{KL}(\hat{q}||p)\,.$$

This appears in part (3) of Theorem 1. It follows that

$$\overline{d_p} \le D_{KL}(\hat{q}||p) = D_{KL}(1-\hat{p}\,||\,p).$$

Finally, since $q^* < \hat{q}$, and since $b_p = D_{KL}(q^*||p)$, it follows that

Theorem 21.

$$\overline{d_p} \le D_{KL}(1-\hat{p}\,||\,p) < b_p\,.$$

This is also in part (3) of Theorem 1.

The event G_n (conditioned on $\Theta = +1$) implies that the agents both choose the wrong action at all times. We have shown that this happens with probability that is exponentially higher than that of agent 1 choosing the wrong action in the unidirectional observation setting (see Figure 5, which compares $\overline{g_p}$ to b_p).

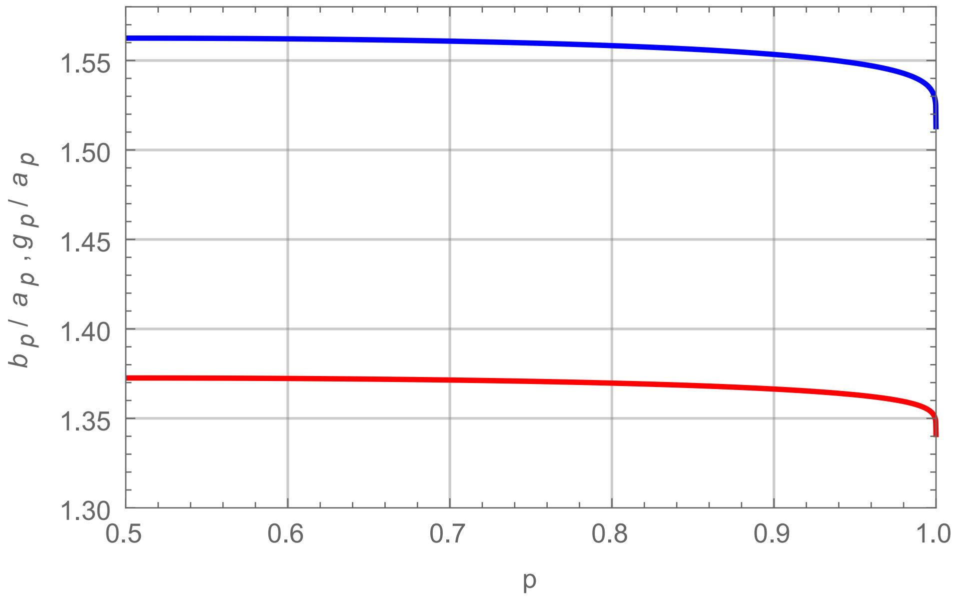

FIGURE 5. $\overline{g_p}$ (red) is lower than $b_p = c_p$ (blue); that is, bidirectional observation causes an exponential increase in the probability of error. The graph is normalized by a_p.

6.0.3. *Bayesian Groupthink.* We next show that conditioned on the state of the world $\Theta = +1$ and the event G_n that both agents guessed incorrectly in all periods, the agents have, with high probability, strong *positive* private signals. Thus both agents have strong evidence indicating the correct action (+1), and yet take the wrong action (-1 or 0).

Theorem 22. *For every* $\epsilon > 0$

$$\lim_{n\to\infty} \mathbb{P}\left[S_n^1/n > \hat{t} - \epsilon, S_n^2/n > \hat{t} - \epsilon \middle| G_n, \Theta = +1\right] = 1.$$

That is, S_n^1 and S_n^2 are typically not less than about $n \cdot \hat{t}$. We prove this in the appendix.

Theorem 22 may seem surprising at first: even though both agents guess the state incorrectly they both have evidence indicating the correct state. This is not the case in other scenarios: for example, when two agents do not observe each other, when they are both wrong it is obviously the case that both have signals indicating the wrong action.

To understand this, note that as each agent has seen the other taking the wrong action repeatedly, she requires strong evidence in favor of the correct action, in order to switch to it. As the action is incorrect by assumption, both agents will be approximately indifferent; this follows from the "smallest possible mistake" principle (Theorem 10). Thus, conditioned on both agents guessing wrong all the time, they both, in the long run, hold significant evidence in favor of the true state.

7. Conclusion

We have shown that increasing the exchange of information between two agents might lead to exponentially slower learning. To prove this result we used tools from large deviation theory, which we think might be useful also in other models of learning.

Of course, many questions remain open. First and foremost, we were not able to overcome the hurdles involved in calculating the speed of learning in the bidirectional case, but only provided an upper bound for it. Indeed, it is possible that the information transmission efficiency is in fact zero in this case - we were not able to prove otherwise!

Another interesting question is to understand the limits of information transmission efficiency. What actions should the agents choose to maximize the information transmission efficiency? Are there actions which would result in 100% efficiency in both directions, so that both agent's learning rate is $2a_p$? If not, what is the maximum possible rate achievable by both agents?

Appendix A. Additional proofs

The following is a Corollary of Sanov's Theorem.

Corollary 23 (Generalized Sanov)**.** *Let $0 \leq \underline{p}_n < \overline{p}_n \leq 1$ be a converging sequence of intervals such that $\underline{p} = \lim_{n\to\infty}$ and $\overline{p} = \lim_{n\to\infty} \overline{p}_n$ we have that*

$$\lim_{n\to\infty} -\log \frac{1}{n}\mathbb{P}^+[2\underline{p}_n - 1 \leq S_n^i/n \leq 2\overline{p}_n - 1] = \begin{cases} D_{KL}(\underline{p}\,||\,p) & \textit{when } p < \underline{p} \\ 0 & \textit{when } \underline{p} \leq p \leq \overline{p} \\ D_{KL}(\overline{p}\,||\,p) & \textit{when } \overline{p} < p \end{cases} \cdot$$

Proof. We consider the case that $\overline{p} < p$; the rest of the cases follow by a similar argument.

Since $\lim_n \overline{p}_n = \overline{p}$, for every $\epsilon > 0$ it holds for all n sufficiently large that $\overline{p} - \epsilon < \overline{p}_n < \overline{p} + \epsilon$. Additionally, when $\epsilon < p - \overline{p}$, $\overline{p}_n < p$, again for large n. Hence, by Sanov's Theorem,

$$D_{KL}(\overline{p} - \epsilon\,||\,p) \leq \lim_{n\to\infty} -\log \frac{1}{n}\mathbb{P}^+[2\underline{p}_n - 1 \leq S_n^i/n \leq 2\overline{p}_n - 1] \leq D_{KL}(\overline{p} - \epsilon\,||\,p).$$

The claim follows by the continuity of the Kullback-Leibler divergence function $D_{KL}(\cdot\,||\,p)$. $\square$

Proof of Proposition 7. Recall that the optimal action is $+1$ if the log-likelihood ratio L_n^B is greater zero and -1 if it is smaller. By definition of t_n^* it follows that $\operatorname{sgn}(L_n^B) = \operatorname{sgn}(S_n^1)$ whenever $|S_n/n| > t_n^*$ or $A_n^2 = 0$; this covers the first and third cases. Conversely, $\operatorname{sgn}(L_n^B) = \operatorname{sgn}(A_n^2)$ whenever $|S_n/n| \leq t_n^*$ and $A_n^2 \neq 0$; this covers the second case. Whenever agent 1 is indifferent $A_{n-1}^2 \neq 0, |S_n/n| = t_n^*$ we let her follow her own signals which is optimal as she was indifferent. $\square$

Proof of Proposition 13. We define

$$L_n^C := \log \frac{\mathbb{P}^+[\{X_k^1\}_{k\leq n}, \{A_k^2\}_{k<n}]}{\mathbb{P}^-[\{X_k^1\}_{k\leq n}, \{A_k^2\}_{k<n}]}\,.$$

Applying Bayes' rule and conditional independence, it follows that

$$L_n^C = S_n^1 \log \frac{p}{q} + \log \frac{\mathbb{P}^+[\{A_k^2\}_{k<n}]}{\mathbb{P}^-[\{A_k^2\}_{k<n}]}\,.$$

This is a function of S_n^1 and $\{A_k^2\}_{k<n}$. If F_n occurs, $S_n/n < p^* - \varepsilon$, and $A_k^2 = -1$ for every $1 \leq k < n$. Since $C_n^1 = \operatorname{sgn} L_n^C$, it is sufficient to show that, for these values, L_n^C is negative.

By Claim 6, we know that the event $\{\{S_k^2 < 0\}_{k<n}\}$ is identical to $\{\{A_k^2 = -1\}_{k<n}\}$. Therefore,

$$\mathbb{P}^-[\{A_k^2 = -1\}_{k<n}] \geq \mathbb{P}^-[\{S_k^2 < 0\}_{k<n}] \geq \frac{1}{n-1}\mathbb{P}^-[S_{n-1}^2 < 0]\,,$$

where the second inequality follows from Theorem 12. Conditioned on $\Theta = -1$, the mean of S_n^2/n is $q - p$, and its variance is $K_p n$ for some K_p independent of n. Thus, by Chebyshev's

Inequality,

$$\mathbb{P}^-[S^2_{n-1} < 0] \geq 1 - \varepsilon$$

for any $\varepsilon > 0$ and n high enough. Meanwhile, the event $\{\{A^2_k = -1\}_{k<n}\}$ implies $\{S^2_{n-1} \leq 0\}$. Therefore,

$$\mathbb{P}^+[\{A^2_k = -1\}_{k<n}] \leq \mathbb{P}^+[S^2_{n-1} \leq 0] \leq \exp[-(a_p - K'_p \varepsilon)n]\,,$$

where the final inequality holds for any fixed $K'_p > 0$ independent of n, any $\varepsilon > 0$ sufficiently small, and some sufficiently large n, by Sanov's Theorem and the definition of a_p.

Substituting this in to the log-likelihood ratio, we find that, conditioned on F_n,

$$L^C_n/n < (p^* - \varepsilon)\log(p/q) + \frac{1}{n}\log[(n-1)/(1-\varepsilon)] - (a_p - K'_p\varepsilon)\,.$$

Noting that $p^* \cdot \log(p/q) = a_p$ by definition, we see that this quantity is bounded above by

$$L^C_n/n \leq \frac{1}{n}\log[(n-1)/(1-\varepsilon)] - (\log(p/q) - K'_p)\varepsilon\,.$$

If we choose $K'_p < \frac{1}{2}\log(p/q)$, the upper bound is negative for n large enough and ε small enough, and, in particular, L^C_n is negative when F_n occurs. Since $C^1_n = \operatorname{sgn} L^C_n$, we conclude that F_n implies $\{C^1_n \neq +1\}$.

□

Proof of Claim 15. The claim holds at time 1, since $D^1_1 \neq +1$ iff $S^1_1 \leq t_1 = 0$. Assume that it holds up to time $k-1$.

Pick integer $t < t_k$ such that $\mathbb{P}[W^1_{k-1}, W^2_{k-1}, S^1_k = t]$ is non-zero. Then the log-likelihood ratio of the event $\Theta = +1$ given W^1_{k-1}, W^2_{k-1} and $S^1_k = t$ is

$$\log\frac{\mathbb{P}^+[W^1_{k-1}, W^2_{k-1}, S^1_k = t]}{\mathbb{P}^-[W^1_{k-1}, W^2_{k-1}, S^1_k = t]}.$$

By the conditional independence of the signals, this can be separated into

$$\log\frac{\mathbb{P}^+[W^1_{k-1}, S^1_k = t]}{\mathbb{P}^-[W^1_{k-1}, S^1_k = t]} + \log\frac{\mathbb{P}^+[W^2_{k-1}]}{\mathbb{P}^-[W^2_{k-1}]}.$$

The term on the left is equal to $t\log(p/q)$, since each of the probabilities is equal to the number of paths satisfying W^1_{k-1} and satisfying $S^1_k = t$, times the probability of each path, which is always equal. Hence the log-likelihood ratio is

$$t\log(p/q) + \log\frac{\mathbb{P}^+[W^2_{k-1}]}{\mathbb{P}^-[W^2_{k-1}]},$$

which by the definition of t_k is non-positive for any $t \leq t_k$, and positive when $t > t_k$. Therefore, given W^1_{k-1} and W^2_{k-1}, $S^1_k \leq t_k$ is equivalent to this ratio being non-positive, and since this is agent 1's log-likelihood ratio for $\Theta = +1$, it is equivalent to $D^1_k \neq +1$. By symmetry, W^1_{k-1} and W^2_{k-1}, $S^2_k \leq t_k$ is equivalent to $D^2_k \neq +1$, proving the claim. □

Proof of Claim 16. Returning to the definition of W_k^2, and since t_k is positive, $\mathbb{P}^-[W_k^2]$ is at least the probability that S_l^2 is negative for all $l \le k$, conditioned on $\Theta = -1$. This, as the probability that a simple random walk with a negative drift is always negative, is well known to be bounded from below by a constant C independent of k (but not of p). □

Proof of Proposition 17. Assume the contrary. By Claim 16, there is a $C > 0$ such that $C < \mathbb{P}^-[W_k^2]$ for all k.

Now, recalling (12), for any $\varepsilon > 0$, there is an m such that

$$\mathbb{P}^+[W_k^2] \ge \mathbb{P}^+[W_m^2 \cap \{S_l^2 < (\hat{t} - \varepsilon)l\}_{m<l\le k}].$$

The assumption $\hat{p} \ge p$ implies $\hat{t} > p - q$, and therefore the event $\{S_l^2 < (p - q - \varepsilon)l\}_{m<l\le k}$ implies $\{S_l^2 < (\hat{t} - \varepsilon)l\}_{m<l\le k}$. Applying the Ballot Theorem and Sanov's Theorem again, we deduce that

$$\mathbb{P}^+[W_k^2] \ge \frac{q^m}{k} e^{-(k-m)[D_{KL}(p-\varepsilon||p)+\varepsilon]},$$

after possibly increasing the value of k. Substituting this into (9), dividing through by k and taking limits, we see that

$$\hat{t} < D_{KL}(p - \varepsilon||p) + \varepsilon\,.$$

Since ε is arbitrary and $D_{KL}(\cdot||p)$ is continuous, we conclude that $\hat{t} = 0$ and so $\hat{p} = 1/2$, contradicting the assumption that $\hat{p} \ge p > 1/2$. □

Proof of Proposition 18. By inclusion, $\mathbb{P}^+[W_k^2] \le \mathbb{P}^+[S_k^2 \le t_k]$, and so, by Sanov's Theorem

$$-\frac{1}{k}\log \mathbb{P}^+[W_k^2] \ge -\frac{1}{k}\log \mathbb{P}^+[S_k^2 \le t_k] \ge -D_{KL}(1/2(1 + t_k/k)||p) - \varepsilon,$$

for every positive ε and k sufficiently large, and whenever t_k/k is smaller than $2p-1$. Since $\frac{1}{2}(1+\hat{t}) = \hat{p} < p$, this happens for all n_i large enough.

Noting that ε is arbitrary, we conclude that

$$\liminf_i -\frac{1}{n_i} \cdot \log \mathbb{P}^+[W_{n_i}^2] \ge D_{KL}(\hat{p}||p),$$

and that

$$\limsup_k -\frac{1}{k} \cdot \log \mathbb{P}^+[W_k^2] \ge D_{KL}(\hat{p}||p). \tag{11}$$

It thus remains to be shown that this limit superior is at most $D_{KL}(\hat{p}||p)$.

Moving to the lower bound, we fix $\varepsilon > 0$. Then there exists an $m > 0$ such that for all $l > m$

$$t_l > (\hat{t} - \varepsilon)l.$$

Then, from the definition of W_k^2, we see that

$$\mathbb{P}^+[W_k^2] \geq \mathbb{P}^+[W_m^2 \cap \{S_l^2/l < (\hat{t}-\varepsilon)\}_{m<l\leq k}]. \tag{12}$$

Now, W_m^2 includes the event $\{S_l^1 = -l\}_{l\leq m}$, which, conditioned on $\Theta = +1$, has probability q^m. Hence

$$\mathbb{P}^+[W_k^2] \geq q^m \cdot \mathbb{P}^+[\{S_l^2/l < (\hat{t}-\varepsilon)\}_{m<l\leq k} \mid S_m^2 = -m].$$

By the Ballot Theorem (which adds a $1/k$ factor) and Sanov's Theorem, it follows that

$$\mathbb{P}^+[W_k^2] \geq \frac{q^m}{k} e^{-(k-m)[D_{KL}(\hat{p}-\varepsilon||p)+\varepsilon]},$$

after possibly increasing k. Taking the limit superior and noting that ε is arbitrary, we deduce

$$\limsup_k -\frac{1}{k} \cdot \log \mathbb{P}^+[W_k^2] \leq D_{KL}(\hat{p}||p)\,,$$

proving the claim. □

Proof of Theorem 22. By the definition of G_n and by the conditional independence of the private signals,

$$\mathbb{P}\left[S_n^1/n > \hat{t}-\epsilon, S_n^2/n > \hat{t}-\epsilon \middle| G_n, \Theta = +1\right] = \mathbb{P}\left[S_n^1/n > \hat{t}-\epsilon \middle| S_n^1/n \leq t_n/n, \Theta = +1\right]^2.$$

Let $p_n = \mathbb{P}\left[S_n^1/n > \hat{t}-\epsilon \middle| S_n^1/n \leq t_n/n, \Theta = +1\right]$ be the probability of the event on the right hand side. Then

$$1 - p_n = \mathbb{P}\left[S_n^1/n \leq \hat{t}-\epsilon \middle| S_n^1/n \leq t_n/n, \Theta = +1\right].$$

By Bayes' Theorem

$$1 - p_n = \frac{\mathbb{P}^+[S_n^1/n \leq \hat{t}-\epsilon, S_n^1/n \leq t_n/n]}{\mathbb{P}^+[S_n^1/n \leq t_n/n]},$$

and for n large enough $t_n/n > \hat{t}-\epsilon$, and so

$$1 - p_n = \frac{\mathbb{P}^+[S_n^1/n \leq \hat{t}-\epsilon]}{\mathbb{P}^+[S_n^1/n \leq t_n/n]}.$$

Since $t_n/n \leq \hat{t}-\epsilon/2$ for all n large enough, we have that the denominator

$$\mathbb{P}^+[S_n^1/n \leq t_n/n] \geq \mathbb{P}^+[S_n^1/n \leq \hat{t}-\epsilon/2],$$

and so

$$1 - p_n \leq \frac{\mathbb{P}^+[S_n^1/n \leq \hat{t}-\epsilon]}{\mathbb{P}^+[S_n^1/n \leq \hat{t}-\epsilon/2]}.$$

Taking logarithms, dividing by n, and taking and limits we get that

$$\lim_n -\frac{1}{n}\log(1-p_n) \geq \lim_n -\frac{1}{n}\log \mathbb{P}^+[S^1_n/n \leq \hat{t}-\epsilon] - \lim_n -\frac{1}{n}\log \mathbb{P}^+[S^1_n/n \leq \hat{t}-\epsilon/2].$$

Keeping in mind that $\frac{1}{2}(\hat{t}-1) = \hat{p} < p$, and applying Sanov's Theorem to each of the addends in the right hand side, we get that

$$\lim_n -\frac{1}{n}\log(1-p_n) \geq D_{KL}(\hat{p}-\epsilon/2||p) - D_{KL}(\hat{p}-\epsilon/4||p).$$

Since $\hat{p} < p$ it follows that that

$$\lim_n -\frac{1}{n}\log 1 - p_n > 0$$

and so

$$\lim_n p_n = 1.$$

$\square$

References

Itai Arieli and Manuel Mueller-Frank. Inferring beliefs from actions. *Available at SSRN*, 2013.

AMI Auersperg, AMI von Bayern, S Weber, A Szabadvari, T Bugnyar, and A Kacelnik. Social transmission of tool use and tool manufacture in goffin cockatoos (cacatua goffini). *Proceedings of the Royal Society B: Biological Sciences*, 281(1793):20140972, 2014.

Robert J. Aumann. Agreeing to disagree. *The Annals of Statistics*, 4(6):1236–1239, 1976. ISSN 0090-5364.

Albert Bandura. Influence of models' reinforcement contingencies on the acquisition of imitative responses. *Journal of personality and social psychology*, 1(6):589, 1965.

Joseph Bertrand. Solution d'un probléme. *Comptes Rendus de l'Acadèmie des Sciences, Paris*, 105:369, 1887.

Patrick Bolton and Christopher Harris. Strategic experimentation. *Econometrica*, 67(2): 349–374, 1999.

Thomas M Cover and Joy A Thomas. *Elements of information theory*. John Wiley & Sons, 2012.

Martin Cripps, Jeffrey Ely, George Mailath, and Larry Samuelson. Common learning with intertemporal dependence. *International Journal of Game Theory*, 42(1):55–98, 2013.

Martin W Cripps, Jeffrey C Ely, George J Mailath, and Larry Samuelson. Common learning. *Econometrica*, 76(4):909–933, 2008.

Amir Dembo and Ofer Zeitouni. *Large deviations techniques and applications*. Springer, second edition, 1998.

Darrell Duffie and Gustavo Manso. Information percolation in large markets. *The American economic review*, pages 203–209, 2007.

Darrell Duffie, Semyon Malamud, and Gustavo Manso. Information percolation with equilibrium search dynamics. *Econometrica*, 77(5):1513–1574, 2009.

Darrell Duffie, Gaston Giroux, and Gustavo Manso. Information percolation. *American Economic Journal: Microeconomics*, pages 100–111, 2010.

Darrell Duffie, Semyon Malamud, and Gustavo Manso. Information percolation in segmented markets. *Journal of Economic Theory*, 153:1–32, 2014.

Rick Durrett. *Probability: theory and examples*. Cambridge University Press, 1996.

Glenn Ellison and Drew Fudenberg. Word-of-mouth communication and social learning. *The Quarterly Journal of Economics*, 110(1):pp. 93–125, 1995.

Douglas Gale and Shachar Kariv. Bayesian learning in social networks. *Games and Economic Behavior*, 45(2):329–346, 2003.

John D Geanakoplos and Heraklis M Polemarchakis. We can't disagree forever. *Journal of Economic Theory*, 28(1):192–200, 1982.

Paul Heidhues, Sven Rady, and Philipp Strack. Strategic experimentation with private payoffs. *Available at SSRN 2152117*, 2014.

Kevin N. Hoppitt, William; Laland. *Social Learning: An Introduction to Mechanisms, Methods, and Models*. Princeton University Press, 2013.

Ali Jadbabaie, Pooya Molavi, and Alireza Tahbaz-Salehi. Information heterogeneity and the speed of learning in social networks. *Columbia Business School Research Paper*, (13-28), 2013.

Yashodhan Kanoria and Omer Tamuz. Tractable bayesian social learning on trees. *Selected Areas in Communications, IEEE Journal on*, 31(4):756–765, 2013.

Godfrey Keller and Sven Rady. Strategic experimentation with poisson bandits. *Theoretical Economics*, 5(2):275–311, 2010.

Godfrey Keller, Sven Rady, and Martin Cripps. Strategic experimentation with exponential bandits. *Econometrica*, 73(1):39–68, 2005.

Richard D McKelvey and Talbot Page. Common knowledge, consensus, and aggregate information. *Econometrica: Journal of the Econometric Society*, pages 109–127, 1986.

Elchanan Mossel, Allan Sly, and Omer Tamuz. On agreement and learning. *arXiv preprint arXiv:1207.5895*, 2012a.

Elchanan Mossel, Allan Sly, and Omer Tamuz. Strategic learning and the topology of social networks. Forthcoming in Econometrica, 2012b.

Elchanan Mossel, Allan Sly, and Omer Tamuz. Asymptotic learning on bayesian social networks. *Probability Theory and Related Fields*, 158(1-2):127–157, 2014.

Michael Ostrovsky. Information aggregation in dynamic markets with strategic traders. *Econometrica*, 80(6):2595–2647, 2012.

Rohit Parikh and Paul Krasucki. Communication, consensus, and knowledge. *Journal of Economic Theory*, 52(1):178–189, 1990.

Dinah Rosenberg, Eilon Solan, and Nicolas Vieille. Informational externalities and emergence of consensus. *Games and Economic Behavior*, 66(2):979–994, 2009.

IN Sanov. On the probability of large deviations of random variables. *Matematicheskii Sbornik*, 42(84), 1957.

James K Sebenius and John Geanakoplos. Don't bet on it: Contingent agreements with asymmetric information. *Journal of the American Statistical Association*, 78(382):424–426, 1983.

(M. Harel) University of Geneva

(E. Mossel) University of Pennsylvania and University of California, Berkeley

(P. Strack) U.C. Berkeley

(O. Tamuz) California Institute of Technology