THE SPEED OF SOCIAL 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

Acknowledgments: 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 and seminar audiences at Bonn, UC Berkeley, Microsoft Research New England, Montreal, Pittsburgh, UPenn 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 N00014-15-1-0823 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.


\(^3\)Rather than strategically acquired, as in the bandit literature.
may be thought of as “Bayesian groupthink”: this occurs when initial interactions between 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\textsuperscript{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 independent 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 $2a_p$, 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 17).

\textsuperscript{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 observe 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 24). 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 other agent’s signal, and so still both choose the wrong action (Theorem 25); 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 $\Theta$ that takes values in $\{+1, -1\}$, both of which are a priori equally likely: $P(\Theta = +1) = 1/2$. There are two agents, indexed by $i \in \{1, 2\}$, and $n$ time periods. Each agent observes a sequence of $n$ private signals $\{X^i_k\}_{k \leq n}$, which are i.i.d. conditional on the state of the world $\Theta$. The signal $X^i_k \in \{-1, 1\}$ observed by agent $i$ in period $k$ is equal to the true state of the world $\Theta$ with probability $p$ and equals $-\Theta$ with probability $1 - p$:

$$P[X^i_k = \Theta | \Theta] = p.$$ 

We will use $P^+[\cdot]$ to denote $P[\cdot | \Theta = +1]$, and likewise $P^-[\cdot]$ to denote $P[\cdot | \Theta = -1]$. We also denote by $S^i_n = \sum_{k \leq n} X^i_k$ 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 $\Theta$. 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 $A^i_n$ is given in this case by her best guess as to the value of state of the world $\Theta$ - i.e.

$$A^i_n := \arg \max_{\theta \in \{+1, -1\}} P[\Theta = \theta | \{X^i_k\}_{k \leq n}].$$

If the arg max is not unique, we let $A^i_n$ take the value 0 (and do the same below for $B^1_n$, $C^1_n$ and $D^i_n$ below); 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 $A^2_n$, as above. Agent 1, at time $n$, observes agent 2’s penultimate action $A^2_{n-1}$, in addition to her own private signals. Hence agent 1’s action $B^1_n$ is given in this case by

$$B^1_n := \arg \max_{\theta \in \{+1, -1\}} P[\Theta = \theta | \{X^1_k\}_{k \leq n}, A^2_{n-1}].$$
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 \( C^1_n \) is given by

\[
C^1_n := \arg \max_{\theta \in \{+1, -1\}} \mathbb{P} \left[ \Theta = \theta \mid \{X^1_k\}_{k \leq n}, \{A^2_k\}_{k < n} \right].
\]

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 \( D^i_n \), the action of agent \( i \) at period \( n \), is given recursively by

\[
D^i_n := \arg \max_{\theta \in \{+1, -1\}} \mathbb{P} \left[ \Theta = \theta \mid \{X^i_k\}_{k \leq n}, \{D^{3-i}_k\}_{k < n} \right].
\]

Note that here and below \( 3-i \) is the “other agent”; if \( i = 1 \) then \( 3-i = 2 \) and vice versa. \( D^i_n \) is hence the action agent \( i \) would take when she could observe \( D^{3-i}_k \) at all previous time periods \( k \) in addition to her own signal.

Let

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

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 \( \Theta \). Since this probability vanishes exponentially fast in \( n \), we scale the probabilities to extract the exponential rate of vanishing. Specifically, we define

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

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

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

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

We also define

\[
d_p := \limsup_{n \to \infty} -\frac{1}{n} \log \mathbb{P} \left[ D^i_n \neq \Theta \right]
\]

\[
g_p := \limsup_{n \to \infty} -\frac{1}{n} \log \mathbb{P} \left[ G_n | \Theta = +1 \right].
\]

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 $\sigma$-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 $A_n^1$ to these numbers. As it turns out, we show that $A_n^1 < b_n^1$, 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 $d_p$ to these numbers. As it turns out, we show that $d_p < b_p$; 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 $d_p < b_p$; it follows that $d_p \leq \overline{d}_p$ 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 $\mu$ and $\nu$ 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 $\nu$ is the Bernoulli distribution which assigns probability $p'$ to $+1$ and $q'$ to $-1$ and $\mu$ 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 $D_{KL}(-\| p)$ 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

$$a_p = D_{KL}(\frac{1}{2} \| p).$$

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 $\overline{d}_n$, which in turn is strictly less than $b_p$:

$$\overline{d}_p \leq \overline{d}_p = D_{KL}(1 - \hat{p} \| p) < b_p.$$
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 $\{X^1_k\}_{k\leq n}$, as well as the first $\lfloor \alpha n \rfloor$ private signals of agent 2, $\{X^2_k\}_{k\leq \alpha n}$. 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 $H^\alpha_n$ the best guess of agent 1 when she observes the other agent’s $\lfloor \alpha n \rfloor$ signals in addition to her own signals:

$$H^\alpha_n = \operatorname*{arg \, max}_{\theta \in \{+1,-1\}} \mathbb{P}[\Theta = \theta | \{X^1_k\}_{k\leq n}, \{X^2_k\}_{k\leq \alpha n}] .$$

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, $I_n = \{A^2_k\}_{k\leq n}$ is the relevant information structure. In the bidirectional case where both observe each other’s actions, $I_n = \{D^2_k\}_{k\leq n}$ 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 $n$

$$\max_{\theta \in \{+1,-1\}} \mathbb{P}[\Theta = \theta | \{X^1_k\}_{k\leq n}, I_n] > \mathbb{P}[H^\bar{\alpha}_n = \Theta] .$$

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 \( \beta \) is the minimal asymptotic rate of learning of agent 1 under the information structure \( I \)
\[ \beta = \liminf_{n \to \infty} -\frac{1}{n} \log \left( \max_{\theta \in \{+1,-1\}} P\left[ \theta \neq \Theta \mid \{X_k^1\}_{k \leq n}, I_n \right] \right). \]

Proof of Proposition 3. As signals are i.i.d., an agent is equally likely to be correct if she observed \( \hat{\alpha}n \) of the other agent’s signals, in addition to her own \( n \) signals and in the situation where she observed \((1 + \hat{\alpha})n\) of her own signals. It thus follows from the first result of Theorem 1 that the probability of being wrong when observing a fraction \( \hat{\alpha} \) of the other agents signals is given by
\[ P[H_n^{\hat{\alpha}} \neq \Theta] = e^{-a_p(1+\alpha)n+o(n)}. \]

(1)

By the definition of \( \beta \) we have that
\[ \max_{\theta \in \{+1,-1\}} P[\theta \neq \Theta \mid \{X_k^1\}_{k \leq n}, I_n] \leq e^{-\beta n+o(n)}. \]

(2)

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

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 \( \alpha(I) \leq \frac{\beta}{a_p} - 1 \). \( \square \)

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
\[ \frac{b_p}{a_p} - 1 = \frac{c_p}{a_p} - 1 = \frac{D_{KL}(p^* \mid \mid p)}{D_{KL}(1/2 \mid \mid p)}. \]

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 \( \lim_{n \to \infty} \frac{f(n)}{n} = 0 \).
2.5.2. Information Transmission with Bidirectional Information. Our bound on $d_p$ implies that the efficiency in the bidirectional observation case is bounded from above by

$$\frac{d_p}{a_p} - 1.$$ 

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 least 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

$$A^i_n := \arg \max_{\theta \in \{+1,-1\}} \mathbb{P} [\Theta = \theta \mid \{X^i_k\}_{k \leq n}] .$$

Her exponential rate of learning in this setting is

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

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 \( S_n^i / n = \frac{1}{n} \sum X_i \) tends to its expectation \( 2p - 1 \), by the law of large numbers. Because \( \mathbb{E}[\exp(X_k^i)] < \infty \), 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 \( p < \bar{p} \) we have that

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

We are interested in calculating

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

where \( A_n^1 \) is agent 1’s best guess of \( \Theta \), given her private signals \( \{X_k^1\}_{k \leq n} \). The following claim is an easy consequence of the definition of \( A_n^1 \):

Claim 6.

\[
A_n^1 = \text{sgn } S_n^i.
\]

Proof. The log-likelihood ratio of \( \Theta \), given \( \{X_k^1\}_{k \leq n} \), is

\[
L_n^A := \log \frac{\mathbb{P}[\Theta = +1 | \{X_k^1\}]}{\mathbb{P}[\Theta = -1 | \{X_k^1\}]} \quad \text{is} \quad \log \frac{\mathbb{P}^+([X_k^1])}{\mathbb{P}^-([X_k^1])},
\]

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 \( \mathbb{P}[\Theta = +1 | \{X_k^1\}] = \frac{1}{2} \), and furthermore the latter is greater than half (resp. less then half) if the former is positive (resp. negative). Hence \( A_n^1 = \text{sgn } L_n^A \), and so \( A_n^1 = \text{sgn } S_n^i \). \( \square \)
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 \( e^{-a_p n + o(n)} \), 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

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

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

Since the \( X_k^1 \)'s are independent conditional on \( \Theta \), we can now apply Sanov’s Theorem, substituting \( \underline{p} = 0 \) and \( \bar{p} = \frac{1}{2} \):

\[
a_p = \lim_{n \to \infty} -\frac{1}{n} \log \mathbb{P}[A_n^1 \neq \Theta] = \lim_{n \to \infty} -\frac{1}{n} \log \mathbb{P}^+ [-1 \leq S_n^1 / n \leq 0] = D_{KL}(1/2 || p),
\]

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 unidirectionally

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

\[
A_n^2 := \arg \max_{\theta \in \{+1,-1\}} \mathbb{P}[\Theta = \theta | \{X_k^2\}_{k \leq n}],
\]
and agent 1’s actions are given by

\[ B_n^1 := \arg \max_{\theta \in \{+1, -1\}} \mathbb{P} \left[ \Theta = \theta \mid \{X^1_k\}_{k \leq n}, A^2_{n-1} \right]. \]

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 \( \{X^1_k\} \) and \( \{A^2_{n-1}\} \), the relevant log-likelihood ratio is

\[ L_n^B := \log \frac{\mathbb{P}^+\{X^1_k\}, A^2_{n-1}}{\mathbb{P}^-\{X^1_k\}, A^2_{n-1}}. \]

Analogously to \( A^1_n \) 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

\[ L_n^B = \log \frac{\mathbb{P}^+\{X^1_k\}}{\mathbb{P}^-\{X^1_k\}} + \frac{\mathbb{P}^+[A^2_{n-1}]}{\mathbb{P}^-[A^2_{n-1}]} \]

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

\[ \log \frac{\mathbb{P}^+[A^2_{n-1}]}{\mathbb{P}^-[A^2_{n-1}]} = \mathbf{1}_{\{A^2_{n-1} = +1\}} \log \frac{\mathbb{P}[A^2_{n-1} = \Theta]}{\mathbb{P}[A^2_{n-1} \neq \Theta]} + \mathbf{1}_{\{A^2_{n-1} = -1\}} \log \frac{\mathbb{P}[A^2_{n-1} = \Theta]}{\mathbb{P}[A^2_{n-1} \neq \Theta]}
\]

\[ = A^2_{n-1} \log \frac{\mathbb{P}[A^2_{n-1} = \Theta]}{\mathbb{P}[A^2_{n-1} \neq \Theta]} \]

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

\[ L_n^B = \text{sgn}(S_n^1) S_n^1 \log (p/q) + A^2_{n-1} \log \frac{\mathbb{P}[A^2_{n-1} = \Theta]}{\mathbb{P}[A^2_{n-1} \neq \Theta]}. \]

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. Let the threshold signal \( t_n^* \) be given by

\[ t_n^* = \frac{1}{n} \log \frac{\mathbb{P}[A^2_{n-1} = \Theta]}{\mathbb{P}[A^2_{n-1} \neq \Theta]} \log (p/q)^{-1}. \]

**Proposition 7.** Agent 1’s action \( B_n^1 \) is given by:

\[ B_n^1 = \begin{cases} 
\text{sgn } S_n^1 & \text{if } |S_n^1/n| > t_n^* \text{ or } A^2_{n-1} = 0 \\
A^2_{n-1} & \text{if } |S_n^1/n| < t_n^* \text{ and } A^2_{n-1} \neq 0 \\
\text{sgn } (A^2_{n-1} + \text{sgn } S_n^1) & \text{if } |S_n^1/n| = t_n^* 
\end{cases} \]

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 \( B_n^1 = \text{sgn } S_n \). In the second case agent
2 has the weakest possible private signal (since \( S_{n-1}^2 = 0 \)), and so agent 1 again follows her own signal and chooses \( B_1^n = \text{sgn} S_n \). 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 \( B_1^n = A_{n-1}^2 \). 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. \( -t_n^* \leq S_n/n \leq t_n^* \) and \( A_{n-1}^2 = -1 \).
3. \( -t_n^* \leq S_n/n \leq 0 \) and \( A_{n-1}^2 = 0 \).

The threshold signal \( t_n^* \) defined in (7) 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., \( \lim_n \mathbb{P}[A_{n-1}^2 = \Theta] = 1 \)), and using our characterization of the asymptotic rate of learning in the single agent case derived in (4), we have that

\[
\lim_{n \to \infty} t_n^* = \frac{1}{\log(p/q)} \lim_{n \to \infty} \frac{1}{n} \log \mathbb{P}[A_{n-1}^2 = \Theta] - \log \mathbb{P}[A_{n-1}^2 \neq \Theta] - \log \mathbb{P}[A_{n-1}^2 \neq \Theta] = \frac{1}{\log(p/q)} \lim_{n \to \infty} \frac{1}{n} \log \mathbb{P}[A_{n-1}^2 \neq \Theta] \cdot \lim_{n \to \infty} \frac{n-1}{n} = \frac{a_p}{\log(p/q)}. \]

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

\[
\mathbb{P}[B_1^n \neq \Theta] = \mathbb{P}^+[B_1^n \neq +1] = \mathbb{P}^+[S_n/n < -t_n^*]
\]

\[
+ \mathbb{P}^+[-t_n^* \leq S_n/n \leq t_n^*] \cdot \mathbb{P}^+[A_{n-1}^2 = -1]
\]

\[
+ \mathbb{P}^+[-t_n^* \leq S_n/n \leq 0] \cdot \mathbb{P}^+[A_{n-1}^2 = 0].
\]
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 
\[ p^* = \frac{1}{2}(1 + t^*) \]
as earlier, and likewise denote \( q^* = 1 - p^* \), then it follows from a generalized version of Sanov’s Theorem (Corollary 26) that 
\[ \lim_{n \to \infty} -\frac{1}{n} \log \mathbb{P}[S_1^n/n < -t_n^*] = \lim_{n \to \infty} -\frac{1}{n} \log \mathbb{P}[S_1^n/n < -t^*] = D_{KL}(q^* || p). \]
For the second event it follows from the generalized version of Sanov’s Theorem that the asymptotic rate is given by 
\[ \lim_{n \to \infty} -\frac{1}{n} \log \left( \mathbb{P}[-t_n^* < S_1^n/n < t_n^*] \cdot \mathbb{P}[A_n^2 = -1] \right) = \lim_{n \to \infty} -\frac{1}{n} \log \mathbb{P}[S_1^n/n < t^*] + \lim_{n \to \infty} -\frac{1}{n} \log \mathbb{P}[S_2^n < 0] = D_{KL}(p^* || p) + D_{KL}(\frac{1}{2} || p). \]
For the third event the asymptotic rate satisfies 
\[ \liminf_{n \to \infty} -\frac{1}{n} \log \left( \mathbb{P}[-t_n^* < S_1^n/n < 0] \cdot \mathbb{P}[A_n^2 = 0] \right) = \lim_{n \to \infty} -\frac{1}{n} \log \mathbb{P}[S_1^n/n < 0] + \lim_{n \to \infty} -\frac{1}{n} \log \mathbb{P}[S_2^n = 0] = D_{KL}(\frac{1}{2} || p). \]
The first limit is equal to \( D_{KL}(\frac{1}{2} || p) \), by Sanov’s Theorem. The second is likewise equal to \( D_{KL}(\frac{1}{2} || p) \); this follows from the fact that using an elementary combinatorial argument, \( \mathbb{P}[S_n^2 = 0] \) can be calculated explicitly:
\[ \mathbb{P}[S_n^2 = 0] = \binom{2n}{n} p^n q^n. \]
Hence the rate of the last summand in (8) is \( 2D_{KL}(\frac{1}{2} || p) \). Gathering the rates of the three summands and taking the minimum, we have that
\[ b_p = \lim_{n \to \infty} -\frac{1}{n} \log \mathbb{P}[B_n^1 \neq \Theta] = \min\{ D_{KL}(q^* || p) , D_{KL}(p^* || p) + D_{KL}(\frac{1}{2} || p) , 2D_{KL}(\frac{1}{2} || p) \}. \]
Now, \( 2D_{KL}(\frac{1}{2} || p) > D_{KL}(p^* || p) + D_{KL}(\frac{1}{2} || p) \), since \( \frac{1}{2} < p^* < p \), and by the monotonicity of \( D_{KL}(\cdot || p) \). Hence
\[ b_p = \min\{ D_{KL}(q^* || p) , D_{KL}(p^* || p) + D_{KL}(\frac{1}{2} || p) \}. \]
Perhaps surprisingly, it so happens that these two numbers are equal: \( D_{KL}(q^* || p) = D_{KL}(p^* || p) + D_{KL}(\frac{1}{2} || p) \); this can be easily shown by substituting the definitions of \( p^* \), \( q^* \) and the Kullback-Leibler divergence.
To understand the intuitive reason for this, it is 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., \( S_n/n \leq \bar{p} < p \) - her mistake will be the smallest possible: i.e., \( S_n/n \approx \bar{p} \). Formally,

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

\[
\lim_{n \to \infty} \mathbb{P}^+ \left[ \bar{p} - \epsilon \leq \frac{S^1_n}{n} \mid \frac{S^1_n}{n} \leq \bar{p} \right] = 1.
\]

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

\[
A_n^2 := \arg \max_{\theta \in \{+1, -1\}} \mathbb{P} \left[ \Theta = \theta \mid \{X^2_k\}_{k \leq n} \right],
\]

and agent 1’s actions are given by

\[
C_n^1 := \arg \max_{\theta \in \{+1, -1\}} \mathbb{P} \left[ \Theta = \theta \mid \{X^1_k\}_{k \leq n}, \{A_k^2\}_{k < n} \right].
\]

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.
To understand the behavior of agent 1 in this setting, we analyze $R^2_n$, the last run length of agent 2 at period $n$. A run is a sequence of time periods in which agent 2 has taken the same action; the last run length is the length of the last run:

$$R^2_n = \max \{k : A^2_\ell = A^2_n \text{ for all } n - k + 1 \leq \ell \leq n\}.$$

The importance of the last run length is that it is - together with agent 2’s last action - a sufficient statistic for the posterior when observing only agent 2’s actions:

**Claim 12.** For all $n > 0$,

$$\mathbb{P} [\Theta = +1 | \{A^2_k\}_{k<n}] = \mathbb{P} [\Theta = +1 | R^2_{n-1}, A^2_{n-1}]$$

We prove this claim in the appendix. The reason for this is that before the beginning of a run it is necessarily the case that agent 2’s action is equal to zero, or, equivalently, that she has observed an equal number of $+1$ and $-1$ private signals. Therefore the prior history is irrelevant for estimating the strength of agent 2’s private signals $|S^2_{n-1}|$. The history since the run is captured by the length of the run and the last action, which is equal to all of the actions in the run.

A calculation similar to the one of the previous section and the independence of the agents signals yields the following expression for agent 1’s likelihood ratio in this setting:

$$L^n_C = \log \frac{\mathbb{P}^+[\{X^1_k\}_{k<n}, \{A^2_k\}_{k<n}]}{\mathbb{P}^-[\{X^1_k\}_{k<n}, \{A^2_k\}_{k<n}]} = \log \frac{\mathbb{P}^+[\{X^1_k\}_{k<n}]}{\mathbb{P}^-[\{X^1_k\}_{k<n}]} + \log \frac{\mathbb{P}^+[\{A^2_k\}_{k<n}]}{\mathbb{P}^-[\{A^2_k\}_{k<n}]}$$

$$= S^1_n \log(p/q) + \log \frac{\mathbb{P} [\Theta = +1 | \{A^2_k\}_{k<n}]}{\mathbb{P} [\Theta = -1 | \{A^2_k\}_{k<n}]}$$

Using Claim 12 in the first step, the symmetry of the states in the second and Bayes rule in the third yields,

$$= S^1_n \log(p/q) + \log \frac{\mathbb{P} [\Theta = +1 | A^2_{n-1}, R^2_{n-1}]}{\mathbb{P} [\Theta = -1 | A^2_{n-1}, R^2_{n-1}]}$$

$$= S^1_n \log(p/q) + A^2_{n-1} \log \frac{\mathbb{P} [\Theta = A^2_{n-1} | R^2_{n-1}]}{\mathbb{P} [\Theta \neq A^2_{n-1} | R^2_{n-1}]}$$

$$= \text{sgn}(S^1_n) |S^1_n| \log(p/q) + A^2_{n-1} \log \frac{\mathbb{P} [R^2_{n-1}, A^2_{n-1} = \Theta]}{\mathbb{P} [R^2_{n-1}, A^2_{n-1} \neq \Theta]}.$$

Comparing the likelihood $L^n_C$ to $L^n_B$ the likelihood when only the last action is observed given in (6), reveals that the only difference is the last term, which depends on the last run length $R^2_{n-1}$.
Figure 5. The estimated precision of agent 2’s action as a function of the number of periods she has been taking her final action for \( p = 0.8 \). The (slightly lower) odd periods are in red, and the even periods are in blue.

To understand this new term, note that by the Markov property of \( S^2_k \),

\[
\mathbb{P} \left[ R^2_n = k, A^2_n = \Theta \right] = \mathbb{P}^+ \left[ S^2_{n-k} = 0, S^2_l > 0 \text{ for } l = n - k + 1, \ldots, n \right]
= \mathbb{P}^+ \left[ S^2_{n-k} = 0 \right] \cdot \mathbb{P}^+ \left[ S^2_l > 0 \text{ for } l = 1, \ldots, k \right].
\]

Using a symmetric argument for the case where agent 2 takes the wrong action and the observation that \( \mathbb{P}^+ \left[ S^2_{n-k-1} = 0 \right] = \mathbb{P}^- \left[ S^2_{n-k-1} = 0 \right] \) yields that the likelihood ratio induced by a run of length \( R^2_n = k \) in period \( n \) is independent of \( n \). We denote it by \( h(k) \):

\[
h(k) := \log \frac{\mathbb{P} \left[ R^2_n = k, A^2_n = \Theta \right]}{\mathbb{P} \left[ R^2_n = k, A^2_n \neq \Theta \right]} = \log \frac{\mathbb{P}^+ \left[ S^2_l > 0 \text{ for } l = 1, \ldots, k \right]}{\mathbb{P}^- \left[ S^2_l > 0 \text{ for } l = 1, \ldots, k \right]}.
\]

The following lemma proven in the appendix characterizes \( h \) explicitly.

**Lemma 13.** The likelihood ratio induced by a run of length \( R^2_n = k \) is given by

\[
h(k) = \log \frac{\sum_{i=[k/2]+1}^{k} q^k p^{k-j} \binom{k}{j} 2^j - k}{\sum_{i=[k/2]+1}^{k} q^k p^{k-j} \binom{k}{j} 2^j - k}.
\]

Furthermore, \( h \) is asymptotically linear with slope \( \lim_{k \to \infty} h(k)/k = a_p \).

Figure 5 illustrates the precision of agent 2’s action depending on the number of periods she has been taking her final action.
Then, the log-likelihood ratio equals

$$L^n_C = \text{sgn}(S^n_1) |S^n_1| \log(p/q) + A^2_{n-1} h(R^n_{n-1}).$$

Given this, we can again characterize agent 1’s actions. Intuitively, as in the previous section, there is a threshold for the strength of agent 1’s private signal, such that agent 1 follows her own private signal when it is above this threshold, and otherwise follows agent 2’s signal. In this setting, as opposed to the previous, this threshold will depend on the last run length $R^n_{n-1}$.

**Proposition 14.** Let the threshold signal $t^*_n : \mathbb{N} \rightarrow \mathbb{R}$ be given by

$$t^*_n(k) = \frac{1}{n} h(k) \log(p/q)^{-1}. \tag{9}$$

Then agent 1’s action $C^n_1$ is given by:

$$C^n_1 = \begin{cases} 
\text{sgn} S^n_1 & \text{if } |S^n_1/n| > t^*_n(R^n_{n-1}) \text{ or } A^2_{n-1} = 0 \\
A^2_{n-1} & \text{if } |S^n_1/n| < t^*_n(R^n_{n-1}) \text{ and } A^2_{n-1} \neq 0 \\
\text{sgn} (A^2_{n-1} + \text{sgn} S^n_1) & \text{if } |S^n_1/n| = t^*_n(R^n_{n-1})
\end{cases}$$

The proof of this Proposition is identical to that of the analogous Proposition 7, and so we omit it.

*Calculating the rate of mistakes.* We wish to understand the probability that agent 1 chooses the wrong action $C^n_1 \neq \Theta$, and in particular show that $c_p = b_p$, or that errors happen at the same rate as if only the last action was observable. Since $B^n_1$ is measurable with respect to the $\sigma$-algebra generated by $\{X^n_k\}$ and $\{A^n_k\}$, we know that $C^n_1$ is more likely to correctly estimate $\Theta$ than $B^n_1$. Therefore, the error probability decreases, and $b_p \leq c_p$. The proof will be complete if we knew that $c_p \leq b_p$. We are therefore looking for a lower bound on the error probability $P[C^n_1 \neq \Theta]$.

The obstacle to understanding the error probabilities in the case where all actions are observable is that agent 1’s behaviour depends on $R^n_{n-1}$, the random last run length of agent 2. While it is hard to calculate the distribution of $R^n_{n-1}$ explicitly, we show that the event that agent 2 has never switched her action - i.e., that $R^n_{n-1} = n - 1$ - is sufficiently likely, so that it suffices to focus on this event.

To bound the error probability we will consider the probability of the event that $|S^n_1/n| < t^*_n$ (i.e., agent 1 has a weak private signal - here $t^*_n$ is as defined in (7)), that $A^n_2 \neq \Theta$, and that $R^n_2 = n$ (i.e., that agent 2 has taken the wrong action in the first $n$ periods). We show that this indeed implies that $C^n_2 \neq \Theta$, and that furthermore the probability of this event has exponential rate $b_p$. Interestingly, we do not lose much by assuming that $R^n_2 = n$; as we
show below, the probability of this event is significant, in the sense that it only diminishes polynomially with $n$, and therefore does not affect the exponential rate.

Formally, we condition on $\Theta = +1$ and let

$$F_n = \{ |S^1_n/n| < t^*_n, R^2_n = n, A^2_n = -1\}$$

be the event that agent 1’s signal is weak and that agent 2 has taken the wrong action in all periods up to $n$. We show that $F_n$ implies that $C^1_n = -1$ (Proposition 16), from which it follows that

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

We also show that

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

thus completing the proof.

We first prove (10), by calculating the probability of $F_n$, conditioned on $\Theta = +1$. Note that the second part of the definition of $F_n$ can be written as $\{ S^2_k < 0 \}_{k \leq n}$, by Claim 6. Thus $F_n$ can be decomposed into a product, because of the conditional independence of the two agents’ private signals:

$$\mathbb{P}^+[F_n] = \mathbb{P}^+[|S^1_n/n| < t^*_n] \cdot \mathbb{P}^+[\{ S^2_k < 0 \}_{k \leq n}].$$

The first probability is easy to calculate via Sanov’s Theorem, as we do in the previous section. 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 15** (Bertrand’s Ballot Theorem). For any negative integer $x < 0$ and any $\theta \in \{+1, -1\}$

$$\mathbb{P} [\{ S^2_k < 0 \}_{k \leq n}, S_n = x \mid \Theta = \theta] = \frac{|x|}{n} \mathbb{P} [S_n = x \mid \Theta = \theta].$$

As a consequence of the Ballot Theorem,

$$\mathbb{P}^+[\{ S^2_k < 0 \}_{k \leq n}] \geq \frac{1}{n} \mathbb{P}^+[S^2_n < 0].$$

Substituting this into (11), the expression for the probability of $F_n$, we find that

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

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

$$D_{KL}(p^* - \frac{1}{2}t^* \| p) + D_{KL}(\frac{1}{2} \| p).$$

Furthermore, we have shown before that

$$D_{KL}(p^* \| p) + D_{KL}(\frac{1}{2} \| p) = b_p.$$

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

We prove Proposition 16 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 periods $k < n$ is less likely than that she takes the wrong action in period $k = n - 1$.

We thus conclude that

**Theorem 17.** $c_p = b_p$.

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

$$D^i_n := \arg\max_{\theta \in \{+1, -1\}} \mathbb{P}[\Theta = \theta \mid \{X^i_k\}_{k \leq n}, \{D^{3-i}_k\}_{k < n}].$$

We are interested in calculating

$$d_p := \lim_{n \to \infty} -\frac{1}{n} \log \mathbb{P}[D^i_n \neq \Theta].$$

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 $D^i_n$.

In lieu, we resort to studying the limit superior

$$\overline{d}_p := \limsup_{n \to \infty} -\frac{1}{n} \log \mathbb{P}[D^i_n \neq \Theta].$$
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 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 \( 1, 2, \ldots, n \). 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

\[
\overline{g}_p := \limsup_n -\frac{1}{n} \log^+ [G_n] \geq \overline{d}_p.
\]

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 \( G_n^1 \cap G_n^2 \), where \( G_n^i \) 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 \( \Theta \). However, due to the fact that the agents’ actions are strongly intertwined, \( G_n^1 \) and \( G_n^2 \) 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, \( W_n^1 \cap W_n^2 \), where \( W_n^i \) depends only on the private signals of agent \( i \). To define these events, we consider the perspective of an outside observer, who observes \( \{ D_l^i \} \) 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 \( S_{1k}^1 \) would imply this. We define that value as \( t_k \), the “threshold” that \( S_{1k}^1 \) must be under to imply the event \( G_k \). By symmetry, the same threshold applies to agent 2, and so \( S_{2k}^2 \) must also be under \( t_k \). Thus \( W_n^i \) is the event that agent \( i \)’s private signals are such that \( S_{ik}^i < t_k \) 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
\[
W^i_k = W^i_{k-1} \cap \{ S_k^i \leq t_k \},
\]
and for \( k > 1 \) let
\[
(12) \quad t_k = -\frac{1}{\log(p/q)} \cdot \log \frac{\Pr^+[W^2_{k-1}]}{\Pr^-[W^2_{k-1}]}.
\]

We prove the following claim in the Appendix.

**Claim 18.**
\[
W^1_n \cap W^2_n = G_n.
\]

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 \( W^i_n \), we start by studying \( t_n \), and in particular its asymptotic behavior, as described by
\[
\hat{t} = \lim \inf_n t_n/n.
\]
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 \( \Pr^+[W^2_n] \) and \( \Pr^-[W^2_n] \). We start by analyzing the latter, and (by an easy exercise) show

**Claim 19.** For every \( \frac{1}{2} < p < 1 \) there is a constant \( C > 0 \) such that, for all \( k > 0 \),
\[
C < \Pr^-[W^2_n].
\]

We prove this in the appendix.

We next turn to understanding \( \Pr^+[W^2_n] \). Now, conditioned on \( \Theta = +1 \), \( S^2_n/n \) will, with high probability, be approximately equal to \( 2p - 1 \). \( W^2_n \) 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 \( W^2_n \), conditioned on \( \Theta = +1 \) should be approximately \( e^{-D_{KL}(\hat{p}||p)n} \). Formally, we show that indeed

**Proposition 20.** \( \hat{p} > p \).

and that

**Proposition 21.** Let \( \{n_i\}_i \) be a sequence such that \( \lim_i t_{n_i}/n_i = \hat{t} \). Then
\[
\limsup_n -\frac{1}{k} \cdot \log \Pr^+[W^2_n] = \lim_i -\frac{1}{n_i} \cdot \log \Pr^+[W^2_{n_i}] = D_{KL}(\hat{p}||p).
\]
Both proofs appear in the appendix. We henceforth let \( \{n_i\}_i \) be such that \( \lim_i t_{n_i}/n_i = \hat{t} \).

Joining the latter proposition with Claim 19 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 \( t_n \) (12) by \( k \) and take limits along \( \{n_i\}_i \), we arrive at

\[
\hat{t} = \frac{D_{KL}(\hat{p}||p)}{\log(p/q)}.
\]

(13)

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 \([0, p)\)

\[
\hat{t} = \frac{D_{KL}(\hat{p}||p)}{\log(p/q)} < \frac{D_{KL}(1/2||p)}{\log(p/q)} = \bar{t}^*,
\]

recalling the definition of \( \bar{t}^* \). Hence \( q^* < \hat{q} \). Another interesting consequence of (13) is that

Lemma 22. \( 2D_{KL}(\hat{p}||p) = D_{KL}(\hat{q}||p) \).

This follows by elementary algebraic manipulations of (13), which we omit. Returning to \( G_n \),

\[
\mathbb{P}^+[G_n] = \mathbb{P}^+[W_{n_1}^1, W_{n_2}^2] = \mathbb{P}^+[W_{n_1}^1]^2.
\]

Hence, by Proposition 21,

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

Thus, applying Lemma 22 yields

Theorem 23.

\[
\overline{\tau}_p = D_{KL}(\hat{q}||p).
\]

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

\[
\overline{d}_p \leq D_{KL}(\hat{q}||p) = D_{KL}(1 - \hat{p} || p).
\]

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

Theorem 24.

\[
\overline{d}_p \leq D_{KL}(1 - \hat{p} || p) < b_p.
\]
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 6, which compares $\overline{g}_p$ to $b_p$).

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 25.** For every $\epsilon > 0$

$$\lim_{n \to \infty} \mathbb{P} \left[ \frac{S_1^1}{n} > \hat{t} - \epsilon, \frac{S_2^2}{n} > \hat{t} - \epsilon \mid G_n, \Theta = +1 \right] = 1.$$

That is, $S_1^1$ and $S_2^2$ are typically not less than about $n \cdot \hat{t}$. We prove this in the appendix.

Theorem 25 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 26** (Generalized Sanov). Let \( 0 < p_n < \overline{p}_n \leq 1 \) be a converging sequence of intervals such that \( \overline{p} = \lim_{n \to \infty} p_n \) and \( \overline{p} = \lim_{n \to \infty} \overline{p}_n \). Then

\[
\lim_{n \to \infty} -\log \frac{1}{n} \mathbb{P}^+ [2p_n - 1 \leq S^*_n/n \leq 2\overline{p}_n - 1] = \begin{cases} 
D_{KL}(\overline{p} \mid \mid p) & \text{when } p < \overline{p} \\
0 & \text{when } p \leq p \leq \overline{p} \\
D_{KL}(\overline{p} \mid \mid p) & \text{when } \overline{p} < p
\end{cases}
\]

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

Since \( \lim \mathbb{P}_n = \overline{p} \), for every \( \epsilon > 0 \) it holds for all \( n \) sufficiently large that \( \overline{p} - \epsilon < \mathbb{P}_n < \overline{p} + \epsilon \). Additionally, when \( \epsilon < p - \overline{p} \), \( \overline{p}_n < p \), again for large \( n \). Hence, by Sanov’s Theorem,

\[
D_{KL}(\overline{p} - \epsilon \mid \mid p) \leq \lim_{n \to \infty} -\log \frac{1}{n} \mathbb{P}^+ [2p_n - 1 \leq S^*_n/n \leq 2\overline{p}_n - 1] \leq D_{KL}(\overline{p} - \epsilon \mid \mid p).
\]

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

**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, and consider the expression for \( L_n^B \) given in (6). If \( |S_n/n| > t_n^* \) or \( A_{n-1}^2 = 0 \) then the first addend dominates the second, and therefore the sign of \( L_n^B \) will equal the sign of \( S_n^1 \); this covers the first case.

If \( |S_n/n| < t_n^* \) and \( A_{n-1}^2 \neq 0 \) then the second addend is dominant, and hence \( \text{sgn}(L_n^B) = A_{n-1}^2 \); this covers the second case. Finally, if \( |S_n/n| = t_n^* \) then the two addends are equal (in absolute value, and when \( A_{n-1}^2 \neq 0 \)) and so the sign of their sum will equal the sign of the average of their signs. This covers the third case. 

**Proof of Claim 12.** We prove the claim by showing that

\[
\frac{\mathbb{P}^+ [\{A_k^2\}_{k \leq n}]}{\mathbb{P}^- [\{A_k^2\}_{k \leq n}]} = \frac{\mathbb{P}^+ [R_n^2, A_n^2]}{\mathbb{P}^- [R_n^2, A_n^2]}.
\]

The event that \( R_n^2 = \ell \) and that \( A_n^2 = a \) is the same as the event\(^6\) that \( A_{n-\ell}^2 = 0 \) (or, equivalently, \( S_{n-\ell}^2 = 0 \)) and that \( \{A_k^2 = a\}_{n-\ell+1 \leq k \leq n} \).

Hence knowing \( \{A_k^2\}_{k \leq n} \) is the same as knowing \( \{A_k^2\}_{k < n-R_n^2} \), \( R_n^2 \) and \( A_n^2 \), and therefore

\[
\frac{\mathbb{P}^+ [\{A_k^2\}_{k \leq n}]}{\mathbb{P}^- [\{A_k^2\}_{k \leq n}]} = \frac{\mathbb{P}^+ [\{A_k^2\}_{k < n-R_n^2} \mid R_n^2, A_n^2]}{\mathbb{P}^- [\{A_k^2\}_{k < n-R_n^2} \mid R_n^2, A_n^2]} \cdot \frac{\mathbb{P}^+ [R_n^2, A_n^2]}{\mathbb{P}^- [R_n^2, A_n^2]}.
\]

Now, \( R_n^2 = \ell, A_n^2 = a \) is the event that \( S_{n-\ell}^2 = 0 \) and \( \{A_k^2 = a\}_{n-\ell+1 \leq k \leq n} \). Hence conditioning \( \{A_k^2\}_{k \leq n-\ell} \) on this event is the same as conditioning it on \( S_{n-\ell}^2 = 0 \); this is because all these\(^6\) using the convention \( A_0^2 = S_0^2 = 0 \).

\[\text{28}\]
events are functions of \( \{S^2_k\}_{k \leq n} \), and because the \( S^2_k \)'s have the Markov property. Hence

\[
\begin{align*}
\mathbb{P}^+ \left[ \{A^2_k\}_{k \leq n} \right] &= \mathbb{P}^+ \left[ \{A^2_k\}_{k \leq n-R^2_n} \mid S_{n-R^2_n} = 0 \right] \cdot \mathbb{P}^+ \left[ R^2_n, A^2_n \right] \\
\mathbb{P}^- \left[ \{A^2_k\}_{k \leq n} \right] &= \mathbb{P}^- \left[ \{A^2_k\}_{k \leq n-R^2_n} \mid S_{n-R^2_n} = 0 \right] \cdot \mathbb{P}^- \left[ R^2_n, A^2_n \right].
\end{align*}
\]

To finish the proof, we note that

\[
\mathbb{P}^+ \left[ \{A^2_k\}_{k \leq n-R^2_n} \mid S_{n-R^2_n} = 0 \right] = \mathbb{P}^- \left[ \{A^2_k\}_{k \leq n-R^2_n} \mid S_{n-R^2_n} = 0 \right].
\]

This follows from the fact that conditioned on \( S_{n-R^2_n} = 0 \), it holds that the distribution of \( \{S_t\}_{t < n-R^2_n} \) is the same whether \( \Theta = +1 \) or \( \Theta = -1 \); in either case it is the uniform distribution over \( (s_1, s_2, \ldots, s_{n-R^2_n}) \in \mathbb{Z}^{n-R^2_n} \) such that \( \mid s_t - s_{el+1} \mid = 1 \) and \( s_{n-R^2_n} = 0 \). \( \square \)

Before proving Lemma 13 we will state and prove an additional lemma. Its proof is a standard calculation of the transience probability of a random walk with a drift, and we provide it in its entirety for the reader’s convenience.

**Lemma 27.** \( p - \frac{1}{2} + \sqrt{\frac{1}{4} - pq} = \mathbb{P}^+ [S^2_l > 0 \text{ for all } l \geq 1] \)

**Proof.** Define \( w_i = \mathbb{P}^+ [S^2_\ell = -i \text{ for some } \ell \geq 1] \) as the probability that at some point in time there are \( i \) more down-ticks (i.e., periods \( k \) for which \( X^2_k = -1 \)) than up-ticks. We have that the probability that there ever is one more down-tick than up-tick equals the probability \( q \) that there is a down-tick in the first period, plus the probability \( p \) that there is first an up-tick, and later two more down-ticks than up-ticks \( w_2 \). As those are disjoint events we have

\[
w_1 = q + w_2.
\]

Now, \( w_2 = w_1^2 \), since

\[
w_2 = \mathbb{P}^+ [S^2_\ell = -2 \text{ for some } \ell \geq 1] = \sum_{\ell} \mathbb{P}^+ [S^2_\ell = -2] = \sum_{\ell} \sum_{k < \ell} \mathbb{P}^+ [S^2_\ell = -2 | S^2_k = -1] \mathbb{P}^+ [S^2_k = -1] = \sum_{\ell} \sum_{k < \ell} \mathbb{P}^+ [S^2_\ell-k = -1] \mathbb{P}^+ [S^2_k = -1],
\]

where the last equality follows from the Markov property of \( S^2_k \). A change of variable yields

\[
w_2 = \sum_{\ell} \sum_{k} \mathbb{P}^+ [S^2_\ell = -1] \mathbb{P}^+ [S^2_k = -1] = w_1^2.
\]

Thus we have that \( w_1 = q + w_1^2 \), and solving for \( w_1 \) yields \( w_1 = \frac{1}{2p} \pm \sqrt{\frac{1}{4p^2} - \frac{q}{p}} \). It follows from \( w_1 \leq 1 \) that \( w_1 = \frac{1}{2p} - \sqrt{\frac{1}{4p^2} - \frac{q}{p}} \). To derive the probability that the random walk
never returns to zero observe that this probability is equal to the probability that the first
tick is an up-tick times the probability that the random walk never has one more down that
up-tick 1 − w1

\[ \mathbb{P}^+[S_l^2 > 0 \text{ for all } l \geq 1] = p(1 - w_1) = p - \frac{1}{2} + \sqrt{\frac{1}{4} - pq}. \]

Proof of Lemma 13. The result follows by applying the Ballot theorem

\[ h(k) = \log \frac{\mathbb{P}^+[S_k^2 = +1 \text{ for } l = 1, \ldots, k]}{\mathbb{P}^-[S_k^2 = +1 \text{ for } l = 1, \ldots, k]} \]

\[ = \log \frac{\sum_{j=[k/2]+1}^k \mathbb{P}^+[S_k^2 = 2j - k, S_l^2 > 0 \text{ for } l = 1, \ldots, k]}{\sum_{j=[k/2]+1}^k \mathbb{P}^-[S_k^2 = 2j - k, S_l^2 > 0 \text{ for } l = 1, \ldots, k]} \]

\[ = \log \frac{\sum_{j=[k/2]+1}^k \frac{2j-k}{k} \mathbb{P}^+[S_k^2 = 2j - k]}{\sum_{j=[k/2]+1}^k \frac{2j-k}{k} \mathbb{P}^-[S_k^2 = 2j - k]}. \]

To derive the asymptote of h, first observe that the probability that agent 2 chooses the
action +1 conditional on the state Θ = 1 is bounded from above and below by a constant

\[ p - \frac{1}{2} + \sqrt{\frac{1}{4} - pq} = \mathbb{P}^+[S_l^2 > 0 \text{ for all } l \geq 1] < \mathbb{P}^+[S_l^2 > 0 \text{ for } l = 1, \ldots, k] < 1. \]

Similarly, the probability that agent 2 chooses the action +1 conditional on the state
Θ = 1 is bounded from above and below by the probability that he chooses the wrong action
in the last period

\[ \mathbb{P}^-[S_k^2 > 0] < \mathbb{P}^-[S_l^2 > 0 \text{ for } l = 1, \ldots, k] \]

\[ = \sum_{j=[k/2]+1}^k \frac{2j-k}{k} \mathbb{P}^-[S_k^2 = 2j - k] < \frac{1}{k} \mathbb{P}^-[S_k^2 > 0]. \]

This leads to the upper bound

\[ \lim_{k \to \infty} \frac{h(k)}{k} = \lim_{k \to \infty} \frac{1}{k} \log \mathbb{P}^+[S_l^2 = +1 \text{ for } l = 1, \ldots, k] - \lim_{k \to \infty} \frac{1}{k} \log \mathbb{P}^-[S_l^2 = +1 \text{ for } l = 1, \ldots, k] \]

\[ \leq \lim_{k \to \infty} \log(1) - \lim_{k \to \infty} \frac{\log(1/k)}{k} - \lim_{k \to \infty} \frac{1}{k} \mathbb{P}^-[S_k^2 > 0] = \lim_{k \to \infty} \frac{1}{k} \mathbb{P}^-[S_k^2 > 0] \cdot \]

and the matching lower bound \( \lim_{k \to \infty} h(k)/k \geq \lim_{k \to \infty} \frac{1}{k} \mathbb{P}^-[S_k^2 > 0] \).

As shown in section 3 \( \lim_{k \to \infty} \frac{1}{k} \mathbb{P}^-[S_k^2 > 0] = 0 \).

Proof of Proposition 16. The event \( F_n \) implies that \( R_{n-1}^2 = n - 1, |S_{n-1}^1/n| < t_n^* \) and \( A_{n-1}^2 = -1 \). For this case, Proposition 14 states that \( C_n^1 = A_{n-1}^2 \), assuming that \( t_n^*(n - 1) \geq t_n^* \). It thus remains to be shown that indeed \( t_n^*(n - 1) \geq t_n^* \). Since

\[ t_n^*(n - 1) = t_n^* + \frac{1}{n} \log h_{n-1}(n - 1) \log(p/q)^{-1}, \]

30
we have to show that \(h(n - 1) \geq 1\), or that, by the definition of \(h(n)\), that

\[
\Pr[R_n^2 = n | A_n^2 = \Theta] \geq \Pr[R_n^2 = n | A_n^2 \neq \Theta].
\]

Conditioning on \(\Theta\), and by symmetry restricting to the case \(\Theta = +1\), this is equivalent to showing that

\[
\Pr^+ \left[ \{ A_k^2 = +1 \}_{k \leq n} \right] \geq \Pr^+ \left[ \{ A_k^2 = -1 \}_{k \leq n} \right]
\]

This is clearly true: when \(\Theta = +1\), the probability of \(\{ A_k^2 = +1 \}_{k \leq n}\), or, equivalently, of \(\{ S_k^2 > 0 \}_{k \leq n}\), is greater than that of \(\{ S_k^2 < 0 \}_{k \leq n}\). Formally, let \(S \subset \mathbb{N}^n\) be the set of sequences \((s_1, s_2, \ldots, s_n)\) such that \(|s_k - s_{k+1}| = 1\) and each \(s_k > 0\). Then

\[
\Pr^+ \left[ \{ S_k^2 > 0 \}_{k \leq n} \right] = \sum_{(s_1, \ldots, s_n) \in S} (pq)^{n/2} \cdot (p/q)^{s_n/2},
\]

whereas

\[
\Pr^+ \left[ \{ S_k^2 < 0 \}_{k \leq n} \right] = \sum_{(s_1, \ldots, s_n) \in S} (pq)^{n/2} \cdot (p/q)^{-s_n/2}.
\]

Since \(p > q > 1\), it follows that indeed the former display is larger than the latter, and so the proof is complete. \(\square\)

**Proof of Claim 18.** 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 \(\Pr[W_{k-1}^1, W_{k-1}^2, S_k^1 = t]\) is non-zero. Then the log-likelihood ratio of the event \(\Theta = +1\) given \(W_{k-1}^1, W_{k-1}^2\) and \(S_k^1 = t\) is

\[
\log \frac{\Pr^+[W_{k-1}^1, W_{k-1}^2, S_k^1 = t]}{\Pr^-[W_{k-1}^1, W_{k-1}^2, S_k^1 = t]}.
\]

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

\[
\log \frac{\Pr^+[W_{k-1}^1, S_k^1 = t]}{\Pr^-[W_{k-1}^1, S_k^1 = t]} + \log \frac{\Pr^+[W_{k-1}^2]}{\Pr^-[W_{k-1}^2]}.
\]

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 \(W_{k-1}^1\) and satisfying \(S_k^1 = t\), times the probability of each path, which is always equal. Hence the log-likelihood ratio is

\[
t \log(p/q) + \log \frac{\Pr^+[W_{k-1}^2]}{\Pr^-[W_{k-1}^2]},
\]

which by the definition of \(t_k\) is non-positive for any \(t \leq t_k\), and positive when \(t > t_k\). Therefore, given \(W_{k-1}^1\) and \(W_{k-1}^2, S_k^1 \leq t_k\) 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 \(D_k^1 \neq +1\). By symmetry, \(W_{k-1}^1\) and \(W_{k-1}^2, S_k^2 \leq t_k\) is equivalent to \(D_k^2 \neq +1\), proving the claim. \(\square\)
Proof of Claim 19. Returning to the definition of $W^2_k$, and since $t_k$ is positive, $\mathbb{P}^- [W^2_k]$ is at least the probability that $S^2_l$ is negative for all $l \leq k$, 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 20. Assume the contrary. By Claim 19, there is a $C > 0$ such that $C < \mathbb{P}^- [W^2_k]$ for all $k$.

Now, recalling (15), for any $\varepsilon > 0$, there is an $m$ such that
\[
\mathbb{P}^+ [W^2_k] \geq \mathbb{P}^+ [W^2_m \cap \{ S^2_l < (\hat{t} - \varepsilon)l \}_{m < l \leq k}].
\]
The assumption $\hat{p} \geq p$ implies $\hat{t} > p - q$, and therefore the event $\{ S^2_l < (p - q - \varepsilon)l \}_{m < l \leq k}$ implies $\{ S^2_l < (\hat{t} - \varepsilon)l \}_{m < l \leq k}$. Applying the Ballot Theorem and Sanov’s Theorem again, we deduce that
\[
\mathbb{P}^+ [W^2_k] \geq \frac{q^m}{k} e^{-(k-m)[D_{KL}(p-\varepsilon||p) + \varepsilon]},
\]
after possibly increasing the value of $k$. Substituting this into (12), dividing through by $k$ and taking limits, we see that
\[
\hat{t} < D_{KL}(p-\varepsilon||p) + \varepsilon.
\]
Since $\varepsilon$ 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 $\hat{p} \geq p > 1/2$. □

Proof of Proposition 21. By inclusion, $\mathbb{P}^+ [W^2_k] \leq \mathbb{P}^+ [S^2_k \leq t_k]$, and so, by Sanov’s Theorem
\[
-\frac{1}{k} \log \mathbb{P}^+ [W^2_k] \geq -\frac{1}{k} \log \mathbb{P}^+ [S^2_k \leq t_k] \geq -D_{KL}(1/2(1 + t_k/k)||p) - \varepsilon,
\]
for every positive $\varepsilon$ 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 $\varepsilon$ is arbitrary, we conclude that
\[
\lim \inf_i -\frac{1}{n_i} \cdot \log \mathbb{P}^+ [W^2_{n_i}] \geq D_{KL}(\hat{p}||p),
\]
and that
\[
\lim \sup_k -\frac{1}{k} \cdot \log \mathbb{P}^+ [W^2_k] \geq D_{KL}(\hat{p}||p).
\]
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 $W^2_k$, we see that
\begin{equation}
\mathbb{P}^+ [W^2_k] \geq \mathbb{P}^+ [W^2_m \cap \{S^2_l/l < (\hat{t} - \varepsilon)\}_{m\leq l \leq k}].
\end{equation}

Now, $W^2_m$ includes the event $\{S^1_l = -l\}_{l \leq m}$, which, conditioned on $\Theta = +1$, has probability $q^m$. Hence
\begin{align*}
\mathbb{P}^+ [W^2_k] &\geq q^m \cdot \mathbb{P}^+ [\{S^2_l/l < (\hat{t} - \varepsilon)\}_{m\leq l \leq k} | S^2_m = -m].
\end{align*}

By the Ballot Theorem (which adds a $1/k$ factor) and Sanov’s Theorem, it follows that
\begin{align*}
\mathbb{P}^+ [W^2_k] &\geq \frac{q^m}{k} e^{-(k-m)\var{D}_{KL}(\hat{p}||p)+\varepsilon},
\end{align*}

after possibly increasing $k$. Taking the limit superior and noting that $\varepsilon$ is arbitrary, we deduce
\begin{align*}
\limsup_k -\frac{1}{k} \cdot \log \mathbb{P}^+ [W^2_k] &\leq \var{D}_{KL}(\hat{p}||p),
\end{align*}
proving the claim. \hfill \Box

**Proof of Theorem 25.** By the definition of $G_n$ and by the conditional independence of the private signals,
\begin{align*}
\mathbb{P} [S^1_n/n > \hat{t} - \varepsilon, S^2_n/n > \hat{t} - \varepsilon | G_n, \Theta = +1] &= \mathbb{P} [S^1_n/n > \hat{t} - \varepsilon | S^1_n/n \leq t_n/n, \Theta = +1]^2.
\end{align*}
Let $p_n = \mathbb{P} [S^1_n/n > \hat{t} - \varepsilon | S^1_n/n \leq t_n/n, \Theta = +1]$ be the probability of the event on the right hand side. Then
\begin{align*}
1 - p_n &= \mathbb{P} [S^1_n/n \leq \hat{t} - \varepsilon | S^1_n/n \leq t_n/n, \Theta = +1].
\end{align*}

By Bayes’ Theorem
\begin{align*}
1 - p_n &= \frac{\mathbb{P}^+ [S^1_n/n \leq \hat{t} - \varepsilon, S^1_n/n \leq t_n/n]}{\mathbb{P}^+ [S^1_n/n \leq t_n/n]},
\end{align*}
and for $n$ large enough $t_n/n > \hat{t} - \varepsilon$, and so
\begin{align*}
1 - p_n &= \frac{\mathbb{P}^+ [S^1_n/n \leq \hat{t} - \varepsilon]}{\mathbb{P}^+ [S^1_n/n \leq t_n/n]}.
\end{align*}

Since $t_n/n \leq \hat{t} - \varepsilon/2$ for all $n$ large enough, we have that the denominator
\begin{align*}
\mathbb{P}^+ [S^1_n/n \leq t_n/n] &\geq \mathbb{P}^+ [S^1_n/n \leq \hat{t} - \varepsilon/2],
\end{align*}
and so
\begin{align*}
1 - p_n &\leq \frac{\mathbb{P}^+ [S^1_n/n \leq \hat{t} - \varepsilon]}{\mathbb{P}^+ [S^1_n/n \leq t - \varepsilon/2]}.
\end{align*}
Taking logarithms, dividing by \( n \), and taking and limits we get that
\[
\lim_{n \to \infty} -\frac{1}{n} \log(1 - p_n) \geq \lim_{n \to \infty} -\frac{1}{n} \log \mathbb{P}^+[S^1_n/n \leq \hat{t} - \epsilon] - \lim_{n \to \infty} -\frac{1}{n} \log \mathbb{P}^+[S^1_n/n \leq \hat{t} - \epsilon/2].
\]
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
\[
\lim_{n \to \infty} -\frac{1}{n} \log(1 - p_n) \geq D_{KL}(\hat{p} - \epsilon/2||p) - D_{KL}(\hat{p} - \epsilon/4||p).
\]
Since \( \hat{p} < p \) it follows that that
\[
\lim_{n \to \infty} -\frac{1}{n} \log(1 - p_n) > 0
\]
and so
\[
\lim_{n \to \infty} p_n = 1.
\]
□

References


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


(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