ON THE SPEED OF SOCIAL LEARNING

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

Abstract. We study the speed of social learning, when two players learn from private signals as well as the actions of the other. Our main finding is that increased interaction between the agents can lower the rate of learning: learning is significantly slower when both players observe each other, than when one only observes the other.

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\textsuperscript{1}. A basic premise of modern liberalism is that the open exchange of information and opinions increases the rate of scientific and technological advances, accelerates the spread of progressive ideas and practices, and ultimately benefits society. In this paper we study a model in which unidirectional observations leads to faster learning, but where bilateral opinion exchange results in an exponential increase in the probability of “groupthink”, and in slower learning.

We consider a game of purely informational externalities, where two players repeatedly decide between two alternatives. One of the two alternatives is “correct”, but it is unknown to the players which one it is. Every period each player privately observes a binary, noisy signal regarding the correct choice. Players are myopic, so that each period they choose the action which they think is more likely to be correct. In a situation of no observation, where each player observes only her own signals, the probability that a player 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 situation of complete information, where each player can observe the signals of both players, the probability that a player takes the wrong action decays exponentially, with rate exactly $2a_p$, since she sees twice as many signals per time period.

We study two intermediate informational situations:

\textsuperscript{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.
(1) Player 1 can see player 2’s actions, in addition to her own signals.
(2) Each player can see her own signals and the other player’s actions (but again not the other player’s signals).

In the first case, we find that player 1’s rate of learning is between $\frac{3}{2} a_p$ and $\frac{25}{16} a_p$; the exact value depends on $p$, and tends towards the former for stronger signals and towards the latter for weaker signals. This result has the intuitive interpretation that player 1 is asymptotically more likely to be correct if she can observe $\frac{9}{16} (\approx 56\%)$ of player 2’s signals than if she could observe all of player 2’s actions. This 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 player 2 is observed.

Second, in the case where both players observes each others’ actions, we show an upper bound on the rate of learning. In particular, we show that player 1’s rate of learning is here strictly slower than it is in the previous case. Naively, one may have guessed that in this case the rate should have been higher, since the only difference is that now player 2 can also see the actions of player 1, so “more information is exchanged”. However, as we show, adding this information to player 2 makes her actions less informative for player 1, and thus lowers player 1’s rate of learning.

Relation to the literature. The literature on social learning in bandit problems focuses on aspects of information acquisition (Bolton and Harris [1999], Keller, Rady, and Cripps [2005], Keller and Rady [2010]). 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 players can free-ride on the information of others. In contrast, we 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.

Starting with the seminal “Agreeing to disagree” Aumann [1976], a large literature has been devoted to the study of agreement between Bayesian agents, with notable contributions by Geanakoplos and Polemarchakis [1982], Sebenius and Geanakoplos [1983], McKelvey and Page [1986] and 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 players observe private signals and learn from the actions of others. They show that players 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 [2012] provide conditions under which agreement
implies that the correct action is taken in a setting with infinitely many players, 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.

Whereas this literature focuses on whether players agree in the long-run and whether they learn the correct action, we study the speed at which players learn the correct action. Especially, as both players 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 players each period observe the signals of $N$ other random players and use simple heuristic decision rules. Recently, Jadbabaie et al. [2013] study a model very similar to ours, but on a general social network, and with boundedly-rational players. 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 et al. [2013]. In their model the agents do not observe each other, but have correlated signals.

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 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.

2. Definitions and 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 players, indexed by $i = 1, 2$, and $n$ time periods. Each player observes a sequence of $n$ signals $\{X_k^i\}$, which are i.i.d. conditional on the state of the world $\Theta$. The signal $X_k^i \in \{-1, 1\}$ observed by player $i$ in period $k$ is equal to the true state of the world $\Theta$ with probability $p$ and equals $-\Theta$ with probability $1 - p$:

$$P \left[ X_k^i = \Theta | \Theta \right] = p.$$ 

2.2. The estimation processes. We define the estimation processes in four different informational situations of interest.
2.2.1. No observation. First, let $A_n^i$ be the $i$th player’s best guess as to the value of state of the world $\Theta$ when she observes only her own signals $\{X_k^i\}_{k \leq n}$ - i.e.

$$A_n^i := \arg\max_{\theta \in \{+1, -1\}} \mathbb{P}[\Theta = \theta \mid \{X_k^i\}_{k \leq n}] .$$

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 $D_n^i$).

2.2.2. Observing the final best guess unidirectionally. Next, we define $B_n^i$ to be the best guess of player $i$ given all of her signals and the final guess of the other player. Formally,

$$B_n^i := \arg\max_{\theta \in \{+1, -1\}} \mathbb{P}[\Theta = \theta \mid \{X_k^i\}_{k \leq n}, A_n^{3-i}] .$$

$B_n^i$ is the action player $i$ would take in a situation where she could observe the other player’s last estimate in addition to her own signal. Note that here and below $3 - i$ is the “other player”; if $i = 1$ then $3 - i = 2$ and vice versa.

2.2.3. Observing all best guesses unidirectionally. Let $C_n^i$ be $i$’s best guess, provided all of the guesses of the other player, instead of just the final one:

$$C_n^i := \arg\max_{\theta \in \{+1, -1\}} \mathbb{P}[\Theta = \theta \mid \{X_k^i\}_{k \leq n}, \{A_k^{3-i}\}_{k < n}] .$$

$C_n^i$ describes the action player $i$ would take in a situation where she could observe all past estimates of the other player in addition to her own signal.

2.2.4. Observing all best guesses bidirectionally. Finally, in the following definition of $D_n^i$, we consider a process in which, at each time $k$, each player observes the other’s best guess. Formally we define $\{D_n^i\}$ recursively by

$$D_n^i := \arg\max_{\theta \in \{+1, -1\}} \mathbb{P}[\Theta = \theta \mid \{X_k^i\}_{k \leq n}, \{D_k^{3-i}\}_{k < n}] .$$

$D_n^i$ is hence the action player $i$ would take when she could observe $D_k^{3-i}$ at all previous time periods $k$ in addition to her own signal.

2.3. Asymptotics. We are interested in the probability that a player’s best estimate does not equal the state of the world $\Theta$. Since this probability vanishes exponentially fast in $n$ for all three cases, we scale the probabilities to extract the exponential rate of vanishing. Specifically, we define

$$a_p := \lim_{n \to \infty} -\frac{1}{n} \log \mathbb{P}[A_n^i \neq \Theta] .$$
Note that $a_p$ is not a function of $i$ by symmetry, and thus depends only on $p$. Continuing the convention of using lowercase letters for rate functions, we let

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

as well as

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

and

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

Note that it is not obvious that the first three limits exist. As part of our main result we prove that they indeed do. We were not able to do the same in the last case, and hence chose to define $d_p$ as a limit superior (which always exists) rather than a limit.

Since $A_n^i$ is measurable with respect to the $\sigma$-algebra that defines $B_n^i$, while $B_n^i$ requires less information than $C_n^i$, we immediately conclude $C_n^i$ is more likely to be correct than $B_n^i$, which is better than $A_n^i$. Thanks to the negative sign in the definition, we see that $a_p \leq b_p \leq c_p$; In fact, we show that the first inequality is strict, while the second is actually 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 $d_p < b_p$; the players learn more slowly when they observe each other’s best guesses, as compared to observing the last guess only.

2.4. Main result.

The Kullback-Leibler Divergence. For notational convenience, let $q := 1 - p$. Letting $\mu$ and $\nu$ be two measures with the same, finite support, we recall that the Kullback-Leibler divergence is defined as

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

where the quantity is infinite if $\nu$ is not absolutely continuous with respect to $\mu$. If $\nu$ is the Bernoulli distribution which assigns probability $p'$ to $+1$ and $1 - p'$ 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

$$D_{KL}(p' || p) = p' \log \frac{p'}{p} + (1 - p') \log \frac{1 - p'}{q}.$$

Fixing $p$, the function $D_{KL}(. || p)$ is nonnegative, has a unique zero at $p$, and is continuous and strictly convex.

Theorem 1 (The Asymptotic Rate of Learning).
(1) 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 player 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 \( p^* = \frac{1}{2} (1 + t^*) \).

(3) \( d_p \), the rate of learning when observing best guesses bidirectionally is bounded by
\[ d_p \leq D_{KL}(1 - \hat{p} \| p) < c_p, \]
where \( \hat{p} \) is the unique solution to \( 2D_{KL}(\hat{p} \| p) = D_{KL}(1 - \hat{p} \| p) \) satisfying \( \hat{p} < p \).

We note that (1) follows almost immediately from Sanov’s Theorem, (2) and (3) require more detailed analysis and constitute the main technical effort of this paper.

3. Discussion of the Main Result

Theorem 1 has various implications for the value of social learning in our setup. First, it allows us to describe the costs stemming from not observing the other player’s outcomes directly, but only observing her actions. If a player could observe all of the other player’s signals, her asymptotic rate of learning would be \( 2a_p \), i.e. twice the rate she could achieve on her own. In a situation where she only observes the other player’s actions (but not her outcomes) she learns at a rate of \( b_p < a_p \). Thus,
\[ \ell^b_p = \frac{2a_p - b_p}{a_p} \]
describes the relative loss from observing actions instead of outcomes. More precisely, a player is equally likely to be right in the long run when she either observes all actions by the other player or \( 1 - \ell^b_p \) of her signals. Figure 1 illustrates that the loss from only observing actions is between \( 7/16 \approx 44\% \) and \( 50\% \), and is higher for more informative signals (\( p \) close to 1). The red line in Figure 1 is the lower bound on the loss when the two players observe each other, obtained in Theorem 1. This shows that the additional loss caused by bidirectional observations is at least an additional \( 18 - 20\% \). Intuitively, if player 2 bases her decisions on the observations of player 1’s actions, then player 2’s actions reveal less information to player 1.

A further point illustrated by Figure 1 is that the losses due to social (as opposed to direct) learning are increasing in the informativeness of the signal \( p \).
Figure 1. The relative loss $\ell_p^b$ from observing actions instead of signals (in blue) and a lower bound on the loss when the two players observe each others’ actions (in red).

4. Proof of Theorem 1

4.1. Preliminaries.

4.1.1. Sanov’s Theorem. We first state a version of Sanov’s Theorem (see, e.g., Dembo and Zeitouni, 1998, Theorem 2.1.10), in the context in which we will use it here. For notational convenience, we say that any measure $\mu$ supported on $\{+1, -1\}$ is in $H \subseteq [0, 1]$ if $\mu(+1) \in H$.

4.1.2. Notation. In the proofs below we will use $P^+[\cdot]$ to denote $P[\cdot|\Theta = +1]$, and likewise $P^-[\cdot]$ to denote $P[\cdot|\Theta = -1]$. We also denote $S_n^i = \sum_{k \leq n} X_k^i$ the difference between the number of $+1$ and $-1$ signals player $i$ observed.

Theorem 2 (Sanov’s Theorem). For any $a < b$ we have that

$$\lim_{n \to \infty} -\log \frac{1}{n} P^+[2a - 1 \leq S_n^i/n \leq 2b - 1] = \begin{cases} D_{KL}(a \| p) & \text{when } a > p \\ D_{KL}(b \| p) & \text{when } b < p. \end{cases}$$

4.2. No observations. Without loss of generality, we consider player 1. To find the value of $A_n^1$, we first calculate the log-likelihood ratio of $\Theta$, given $\{X_k^1\}_{k \leq n}$:

$$l_A := \log \frac{P[\Theta = +1 | \{X_k^1\}]}{P[\Theta = -1 | \{X_k^1\}]} = \log \frac{P^+[\{X_k^1\}]}{P^-[\{X_k^1\}]} ,$$

where the second equality is Bayes’ Rule. Note that $A_n^1 = +1$ if $l_A > 0$, $A_n^1 = -1$ if $l_A < 0$ and $A_n^1 = 0$ if $l_A = 0$. 7
Since the variables are independent, we can explicitly compute the right hand side to be

\[ l_A = S_n^1 \cdot \log(p/q). \]  

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 (1) and by conditioning on \( \Theta = +1, \) we can conclude that

\[ \mathbb{P}[A_n^1 \neq \Theta] = \mathbb{P}^+[S_n^1 \leq 0]. \]  

Since the \( X_k^{i} \)'s are independent conditional on \( \Theta, \) we can now apply Corollary 2. In terms of empirical measure, \( \{S_n^1 \leq 0\} \) is equivalent to \( L_n \in [0, \frac{1}{2}]. \) Therefore, since \( p > \frac{1}{2}, \)

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

where the last equality is Sanov’s Theorem, and the second equality is a consequence of (2). The formula in Theorem 1 is given by explicitly evaluating the Kullback-Leibler divergence.

4.3. Observing the final best guess. We now move on to calculate \( b_p. \) As before, 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

\[ l_B := \log \frac{\mathbb{P}^+[\{X_k^1\}, A_n^2]}{\mathbb{P}^-[\{X_k^1\}, A_n^2]} \]

As for \( l_A \) and \( A_n^1, \) the sign of \( l_B \) determines \( B_n^1. \)

Using the conditional independence of the players’ signals, we see that

\[ l_B = \log \frac{\mathbb{P}^+[\{X_k^1\}]}{\mathbb{P}^-[\{X_k^1\}]} + \log \frac{\mathbb{P}^+[A_n^2]}{\mathbb{P}^-[A_n^2]} \]

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

\[ \log \frac{\mathbb{P}^+[A_n^2]}{\mathbb{P}^-[A_n^2]} = A_n^2 \log \frac{\mathbb{P}[A_n^2 = \Theta]}{\mathbb{P}[A_n^2 \neq \Theta]}, \]

where we follow the usual convention that \( 0 \cdot \log 0 = 0. \)

For large \( n, \) the probability that \( A_n^2 \neq \Theta \) is approximately \( \exp(-a_p n), \) on a logarithmic scale; this follows from (3). Formally, for any \( \varepsilon > 0 \) sufficiently small and sufficiently large \( n, \) the following inequality holds:

\[ \exp[a_p(1 - \varepsilon)n] \leq \frac{\mathbb{P}[A_n^2 = \Theta]}{\mathbb{P}[A_n^2 \neq \Theta]} \leq \exp[a_p(1 + \varepsilon)n]. \]

Plugging all this back into (4), we conclude that

\[ S_n^1 \log(p/q) + A_n^2 \cdot [a_p(1 - \varepsilon)n] \leq l_B \leq S_n^1 \log(p/q) + A_n^2 \cdot [a_p(1 + \varepsilon)n]. \]
Recalling that
\[ t^* := \frac{a_p}{\log(p/q)} , \]
we can write the previous display as
\[ S^1_n + A^2_n \cdot t^*(1 - \varepsilon)n \leq \frac{t^*}{a_p} \cdot l_B \leq S^1_n + A^2_n \cdot t^*(1 + \varepsilon)n . \]

Since \(|A^2_n| \leq 1|\), it follows that if \(|S^1_n| > t^*(1 + \varepsilon)n\), then \(\text{sign}(l_B) = \text{sign}(S^1_n)\) regardless of \(A^2_n\). Furthermore, If \(|S^1_n| \leq t^*(1 - \varepsilon)n\), the sign of \(l_B\) is determined by the sign of \(A^2_n\), meaning \(B^1_n = A^2_n\). Note that the bounds we provided here do not give us a definite decision if \(|S^1_n|\) is close to \(t^*n\). It therefore follows immediately from the right inequality in (5) that
\[ \mathbb{P}\left[ B^1_n \neq \Theta \right] = \mathbb{P}^+\left[ B^1_n \neq +1 \right] \geq \mathbb{P}^+[S^1_n < -t^*(1 + \varepsilon)n] . \]

The probability that \(S^1_n < -t^*(1 + \varepsilon)n\) is the same as the empirical measure of lying in the interval \([0, q^* - \frac{1}{2} t^* \varepsilon)\), where
\[ p^* = \frac{1}{2}(1 + t^*) , \]
as earlier, and \(q^* = 1 - p^*\). Therefore, by Sanov’s Theorem 2,
\[ D_{KL}(q^* - \frac{1}{2} t^* \varepsilon || p) = \lim -\frac{1}{n} \log \mathbb{P}^+[S^1_n < -t^*(1 + \varepsilon)n] \]
\[ \geq \limsup -\frac{1}{n} \log \mathbb{P}\left[ B^1_n \neq \Theta \right] \]
where inequality is a consequence of (6). The above equation shows a lower bound on the probability that player 1 has such a strong wrong signal that she makes a mistake irrespective of what player 2 does. Since \(\varepsilon\) is arbitrary, and by the continuity of \(D_{KL}(. || p)\)
\[ D_{KL}(q^* || p) \geq \limsup -\frac{1}{n} \log \mathbb{P}\left[ B^1_n \neq \Theta \right] . \]

The next paragraph establishes the matching lower bound. If player 1 believes the state to be \(-1\) and and player 2 played \(A^2_n = +1\) it follows from inequality (5) that
\[ S^1_n < -t^*(1 - \varepsilon)n . \]

Similarly, if player 1 believes the state to be \(-1\) and and player 2 played \(A^2_n = -1\) then
\[ S^1_n < t^*(1 - \varepsilon)n . \]

Thus, the event \(\{B^1_n = -1\}\) is included in the union of \(\{S^1_n < -t^*(1 - \varepsilon)n\}\) and \(\{S^1_n < t^*(1 - \varepsilon)n, A^2_n = -1\}\).
\[ \mathbb{P}^+[B^1_n = -1] \leq \mathbb{P}^+[S^1_n < t^*(1 - \varepsilon)n, A^2_n = -1] + \mathbb{P}^+[S^1_n < -t^*(1 - \varepsilon)n, A^2_n = +1] \]
\[ \leq \mathbb{P}^+[S^1_n < t^*(1 - \varepsilon)n, A^2_n = -1] + \mathbb{P}^+[S^1_n < -t^*(1 - \varepsilon)n] . \]
We have here a sum of two probabilities, which we will analyze separately. First,
\[ \lim_{n \to \infty} -\frac{1}{n} \log \mathbb{P}^+ [S_n^1 < -t^*(1 - \varepsilon)n] = D_{KL}(q^* + \frac{1}{2}t^*\varepsilon \| p). \]

The second probability factors into a product of two independent events, one concerning \( S_n^1 \) and the other \( A_n^2 \), by the conditional independence of the signals:
\[ \mathbb{P}^+[S_n^1 < t^*(1 - \varepsilon)n, A_n^2 = -1] = \mathbb{P}^+[S_n^1 < t^*(1 - \varepsilon)n] \cdot \mathbb{P}^+[A_n^2 = -1]. \]

Now, the probability that \( S_n^1 < t^*(1 - \varepsilon)n \) is the same as the empirical measure lying in the interval \( [0, p^* - \frac{1}{2}t^*\varepsilon] \). Note that, \( p^* < p \); this follows from the definition of \( p^* \) and a straightforward (but tedious) computation. Hence
\[ \lim_{n \to \infty} -\frac{1}{n} \log \mathbb{P}^+[S_n^1 < t^*(1 - \varepsilon)n, A_n^2 = -1] = D_{KL}(p^* - \frac{1}{2}t^*\varepsilon \| p) + D_{KL}(\frac{1}{2} \| p). \]

Now, we can upper bound the sum by twice the maximum of the two probabilities. Therefore,
\[ \lim_{n \to \infty} -\frac{1}{n} \log \mathbb{P}^+[B_n^1 \neq +1] \geq \min \left\{ D_{KL}(q^* + \frac{1}{2}t^*\varepsilon \| p), D_{KL}(p^* - \frac{1}{2}t^*\varepsilon \| p) + D_{KL}(\frac{1}{2} \| p) \right\} \]

Note that, if \( \varepsilon = 0 \), the two quantities are exactly identical:
\[ D_{KL}(q^* \| p) = D_{KL}(p^* \| p) + D_{KL}((\frac{1}{2} \| p). \]

The equality follows by substituting the definition of \( t^* \) into the explicit formula for the Kullback-Leibler divergence. Heuristically, this can be thought of as an alternate definition of the variational principle that defines \( t^* \). There are two reasons to incorrectly guess \( \Theta \): one is due to a very strong signal (i.e. \( S_n^1 < -t^*n \)) in the wrong direction, and the other is a weak signal \( (S_n^1 < t^*n) \) in the correct direction coupled with an incorrect guess by the other player. By the definition of \( t^* \), both of these errors turn out to have the same asymptotic error rate.

It hence follows, by the continuity of \( D_{KL}(\cdot \| p) \), that
\[ \lim_{n \to \infty} -\frac{1}{n} \log \mathbb{P} \left[ B_n^1 \neq \Theta \right] \geq D_{KL}(q^* \| p). \]

Hence, combining (7) and (8) we have shown that
\[ b_p = D_{KL}(p^* \| p) + D_{KL}(\frac{1}{2} \| p) = D_{KL}(p^* \| p) + a_p, \]

as required.

4.4. Observing all best guesses unidirectionally. As stated in the theorem, we wish to show that \( c_p = b_p \). Since \( B_n^i \) is measurable with respect to the \( \sigma \)-algebra generated by \( \{ X_k \} \).
and \( \{A_k^2\} \), we know that \( C_n^i \) is more likely to correctly estimate \( \Theta \) 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 \( \{C_n^i \neq +1\} \), and conditioned on \( \Theta = +1 \) had the asymptotic rate function \( b_p \) - i.e.

\[
\lim_{n \to \infty} \frac{1}{n} \log \mathbb{P}[F_n] \leq b_p + \varepsilon,
\]

for any \( \varepsilon > 0 \) sufficiently small. We will show that \( F_n = \{S_1^n < t^*(1 - \varepsilon)n, \{S_k^2 < 0\}_{k \leq n}\} \) satisfies these requirements.

First, we calculate the probability of \( F_n \), conditioned on \( \Theta = +1 \):

\[
\mathbb{P}^+[F_n] = \mathbb{P}^+[S_1^n < t^*(1 - \varepsilon)n] \cdot \mathbb{P}^+\{S_k^2 < 0\}_{k \leq n}.
\]

The first probability is easy to calculate via Corollary 2. For the second expression, we state a corollary of the reflection principle, sometimes referred to as the “Ballot Theorem” (see, e.g., [Durrett, 1996, pg. 198]):

**Theorem 3 (The Ballot Theorem).** For any negative integer \( x < 0 \) and any \( \theta \in \{+1, -1\} \)

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

As a consequence of the Ballot Theorem,

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

Substituting this into (9), the expression for the probability of \( F_n \), we find that

\[
-\frac{1}{n} \log \mathbb{P}^+[F_n] \leq \log \frac{n}{n} - \frac{1}{n} \left( \log \mathbb{P}^+[S_1^n < t^*(1 - \varepsilon)n] + \log \mathbb{P}^+[S_k^2 < 0] \right) .
\]

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

\[
D_{KL}(p^* - \frac{1}{2}t^* \varepsilon \| p) + D_{KL}\left(\frac{1}{2} \| p\right) .
\]

Furthermore, we have shown before that

\[
D_{KL}(p^* \| p) + D_{KL}\left(\frac{1}{2} \| p\right) = b_p .
\]

It therefore remains to be shown that \( F_n \) implies \( \{C_n^i \neq +1\} \). We define

\[
l_C := \log \frac{\mathbb{P}^+\{X_k^1, \{A_k^2\}\}}{\mathbb{P}^+\{X_k^1, \{A_k^2\}\}} .
\]
Applying Bayes’ rule and conditional independence, it follows that
\[ l_C = S_n^1 \log \left( \frac{p}{q} \right) + \log \left( \frac{\mathbb{P}^+[\{A_i^2\}]}{\mathbb{P}^-[\{A_i^2\}]} \right). \]
This is a function of \( S_n^1 \) and \( \{A_i^2\} \). If \( F_n \) occurs, \( S_n < (p^* - \varepsilon)n \), and \( A_k^2 = -1 \) for every \( 1 \leq k \leq n \). It is sufficient to show that, for these values, \( l_C \) is negative.

By the formula for \( l_A \) above, we know that the event \( \{S_k^2 < 0\}_{k \leq n} \) implies \( \{A_k^2 = -1\}_{k \leq n} \). Therefore,
\[ \mathbb{P}^-[\{A_k^2 = -1\}_{k \leq n}] \geq \mathbb{P}^-[\{S_k^2 < 0\}_{k \leq n}] \geq \frac{1}{n} \mathbb{P}^-[S_n^2 < 0], \]
where the second inequality follows from Theorem 3. Given \( \Theta = -1 \), the mean of \( S_n^2 \) is \((q - p)n\), and its variance is \( K_p n \) for some \( K_p \) independent of \( n \). Thus, by Chebyshev’s Inequality,
\[ \mathbb{P}^-[S_n^2 < 0] \geq 1 - \varepsilon \]
for any \( \varepsilon > 0 \) and \( n \) high enough. Meanwhile, the event \( \{A_k^2 = -1\}_{k \leq n} \) implies \( \{S_k^2 \leq 0\}_{k \leq n} \), which, in turn, implies \( \{S_n^2 \leq 0\} \). Therefore,
\[ \mathbb{P}^+[\{A_k^2 = -1\}_{k \leq n}] \leq \mathbb{P}^+[S_n^2 \leq 0] \leq \exp[-(a_p - K_p' \varepsilon)n], \]
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 \( F_n \),
\[ l_C < (p^* - \varepsilon)n \log(p/q) + \log[n/(1 - \varepsilon)] - (a_p - K_p' \varepsilon)n. \]
Noting that \( p^* \cdot \log(p/q) = a_p \) by definition, we see that this quantity is bounded above by
\[ l_C \leq \log[n/(1 - \varepsilon)] - (\log(p/q) - K_p') \varepsilon n. \]
If we choose \( K_p' < \frac{1}{2} \log(p/q) \), the upper bound goes to negative infinity as \( n \) grows, and, in particular, \( l_C \) is negative when \( F_n \) occurs. This implies that \( C_n^1 = -1 \) under these conditions; since we assumed \( \Theta = +1 \), we conclude that \( F_n \) implies \( \{C_n^1 \neq +1\} \), and the proof is complete.

4.5. Observing all best guesses bidirectionally. Let
\[ F_n = \bigcap_{k \leq n, i \in \{1, 2\}} \{D_k^i \neq +1\} \]
be the event that both players guess wrongly at all time periods. Since \( F_n \) implies \( \{D_n^1 \neq +1\} \), it follows that
\[ d_p \leq \lim \inf \frac{1}{n} \log \mathbb{P}^+[F_n]. \]
We will therefore prove that $d_p < c_p$ by showing that the right hand side of the above display is strictly less than $c_p$.

To calculate the probability of $F_n$, we consider the perspective of an outside observer, who observes $\{D_i^k\}$ for $i = 1, 2$ and $l < k$, but has no access to the signals of either players. This outside observer can calculate, given $F_k$, whether there exists a trajectory which would cause player 1 (for example) to guess that $\Theta = +1$ in the $k$th turn, 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 $F_k$. By symmetry, $S^2_k$ must also be under $t_k$.

We now formalize this construction. We will define $t_k$, $W^1_k$ 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

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

For $k > 1$ let

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

Note that this is a positive number.

**Claim 4.**

$$W^1_n \cap W^2_n = F_n.$$  

**Proof.** 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 player one’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. \qed

We would now like to bound the asymptotic probability of

$$\frac{\mathbb{P}^+[W^2_k]}{\mathbb{P}^-[W^2_k]}.$$  

Since $t_k$ is positive, then the denominator can be bounded from below by the probability that $S^2_l$ is negative for all $l \leq k$, conditioned on $\Theta = -1$. This, in turn, is well known to be bounded from below by a constant $C$ independent of $k$ (but not of $p$), as a consequence of the negative drift of such a walk.

To bound the numerator from above, define

$$\hat{t} = \liminf_k t_k/k.$$  

and let $\hat{p} = \frac{1}{2}(1 + \hat{t})$ and $\hat{q} = 1 - \hat{p}$.

First, assume that $\hat{p} < p$, which will allow us to analyze all events using large deviation tools. By inclusion, $\mathbb{P}^+[W^2_k] \leq \mathbb{P}^+[S^2_k \leq t_k]$, and so, by Sanov’s Theorem 2

$$\mathbb{P}^+[W^2_k] \leq e^{-k[D_{KL}(1/2(1+t_k/k)||p) - \varepsilon]},$$

for every positive $\varepsilon$ and $k$ sufficiently large, assuming that $t_k/k$ is smaller than $2p - 1$. By the assumption that $\hat{p} < p$, this happens infinitely often.

Taking the limit inferior, and noting that $\varepsilon$ is arbitrary, we conclude

$$(11) \quad \liminf_k \frac{1}{k} \cdot \log \mathbb{P}^+[W^2_k] \geq D_{KL}(\hat{p}||p).$$

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

$$(12) \quad \mathbb{P}^+[W^2_k] \geq \mathbb{P}^+[W^2_m \cap \{S^2_l < (\hat{t} - \varepsilon)l\}_{m < l \leq k}].$$

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

$$\mathbb{P}^+[W^2_k] \geq q^m \cdot \mathbb{P}^+[\{S^2_l < (\hat{t} - \varepsilon)l\}_{m < l \leq k} \mid S^2_m = -m].$$

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

$$\mathbb{P}^+[W^2_k] \geq \frac{q^m}{k} e^{-(k-m)[D_{KL}(\hat{p} - \varepsilon)||p) + \varepsilon]}.$$
after possibly increasing $k$. Taking the limit inferiors and noting that $\varepsilon$ is arbitrary, we deduce

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

(13)

Hence, by (11), and the constant upper and lower bounds on $\mathbb{P}^{-}[W_k^2]$ from before, we have that

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

Applying this to the definition of $t_k$ (10), dividing both sides by $k$ and taking limit inferiors yields the following equation on $\hat{t}$:

\[ \hat{t} = \frac{D_{KL}(\hat{p}||p)}{\log(p/q)} = \frac{D_{KL}(\frac{1}{2}(1 + \hat{t})||p)}{\log(p/q)}. \]

(14)

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

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

recalling that $t^* = a_p/\log(p/q)$. Hence $q^* < \hat{q}$.

Returning to $F_k$,

\[ \mathbb{P}^+[F_k] = \mathbb{P}^+[W_k^1, W_k^2] = \mathbb{P}^+[W_k^1]^2. \]

Hence, by (13),

\[ \liminf_k \frac{1}{k} \cdot \log \mathbb{P}^+[F_k] \leq 2D_{KL}(\hat{p}||p). \]

It follows from (14) and elementary algebraic manipulations that $2D_{KL}(\hat{p}||p) = D_{KL}(\hat{q}||p)$. Therefore

\[ \liminf_k \frac{1}{k} \cdot \log \mathbb{P}^+[F_k] \leq D_{KL}(\hat{q}||p) < D_{KL}(q^*||p). \]

Since $c_p = D_{KL}(q^*||p)$, it follows that $d_p < c_p$, and the claim is proved, under the assumption $\hat{p} < p$.

To complete the analysis, we show that $\hat{p} < p$. Assume the contrary. The bounds proven for $\mathbb{P}^{-}[W_k^2]$ hold without change. Meanwhile, we can show that a far stronger lower bound holds for $\mathbb{P}^+[W_k^2]$. Recalling (12), for any $\varepsilon > 0$, there is an $m$ such that

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

The assumption $\hat{p} \geq p$ implies $\hat{t} > p - q$, and therefore the event $\{S_t^2 < (p - q - \varepsilon)l\}_{m < l \leq k}$ implies $\{S_t^2 < (\hat{t} - \varepsilon)l\}_{m < l \leq k}$. Applying the Ballot Theorem and Sanov’s Theorem again, we...
deduce that
\[ P^+[W_k^2] \geq \frac{q^m}{k} e^{-(k-m)[D_{KL}(p-\epsilon||p)+\epsilon]}, \]
after possibly increasing the value of $k$. Substituting this into (10), dividing through by $k$ and taking limits, we see that
\[ \hat{t} < D_{KL}(p - \epsilon||p) + \epsilon. \]
Since $\epsilon$ is arbitrary and $D_{KL}(\cdot||p)$ is continuous, we conclude that $\hat{t} = 0$, contradicting the lower bound on $\hat{p}$ and completing the proof.

References


(M. Harel) **University of Geneva**

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

(P. Strack) **U.C. Berkeley**

(O. Tamuz) **Massachusetts Institute of Technology**