THE SPEED OF SOCIAL LEARNING

MATAN HAREL[1], ELCHANAN MOSSEL[2], PHILIPP STRACK[3], AND OMER TAMUZ[4]

Abstract. We study how effectively a group of rational agents learns from repeatedly observing each others' actions. We find that, in the long-run, observing discrete actions of others is significantly less informative than observing their private information: only a fraction of the private information is transmitted. We study how this fraction depends on the distribution of private signals.

In a large society, where everyone's actions are public, this fraction tends to zero, i.e., only a vanishingly small share of the information is aggregated. We identify *groupthink* as the cause of this failure of information aggregation: As the number of agents grows, the actions of each individual depend more and more on the past actions of others, thus revealing less private information.

1. Introduction

In many economic situations, the costs and benefits of possible choices are initially unknown, but can be learned over time. Frequently, agents learn not only from their own experiences, but also from observing the choices made by others facing the same decision. For example, a monopolistic seller in a local market learns about the optimal price both by observing her own demand, as well as by observing the prices other sellers charge in similar markets. Likewise, observing who one's social network friends support might influence who one believes to be the better candidate in an election.

In many such situations information arrives over time, and all agents eventually learn. Two important questions arise: How quickly do they learn, and more interestingly, how does the nature of their social interaction affect their speed of learning?

We study a group of agents who interact repeatedly and try to learn a common state from private signals, as well as from the actions of the others. Every period, each agent

[1]Institut des Hautes Études Scientifiques, Paris
[2]Massachusetts Institute of Technology
[3]University of California, Berkeley
[4]California Institute of Technology
We thank seminar audiences in Berkeley, Berlin, Bonn, Caltech, Chicago, Düsseldorf, Duke, Microsoft Research New England, Medellín, Montreal, NYU, Penn State, Pittsburgh, Princeton, San Diego, UPenn, USC and Washington University, as well as Nageeb Ali, Ben Brooks, Dirk Bergemann, Kim Border, Federico Echenique, Benjamin Golub, Rainie Heck, Steven Morris, Luciano Pomatto, Lones Smith, Leeat Yariv and others for insightful comments and discussions. Elchanan Mossel is supported by ONR grant N00014-16-1-2227 and NSF grant CCF 1320105. Matan Harel was partially supported by the IDEX grant of Paris-Saclay.

observes a private signal, takes an action to maximize her expected utility, and observes others' actions. We study settings with purely informational externalities, i.e., each agent's utility is independent of the others' actions, and hence agents care about others' actions only because they provide information. As in the herding literature (Bikhchandani et al., 1992; Banerjee, 1992; Smith and Sørensen, 2000; Chamley, 2004), private signals are independent of actions, and consequently agents have no experimentation motive. Since each agent can learn the state from her private signals, the question is not whether or not she learns the state eventually, but rather how quickly she does so.

When agents' action spaces are sufficiently rich (e.g., when sellers choose prices from a continuum), actions reveal beliefs, and thereby reveal all the payoff relevant information that is contained in the private signals. In this case no information is lost, and the joint decision problem reduces to that of a single agent. However, when actions are discrete (e.g., when sellers must choose one of finitely many prices), we show that information is lost. We quantify the information loss in terms of the *speed of learning* – the exponential rate at which the probability of choosing a suboptimal action vanishes – and identify two forces that lead to inefficiently slow learning.

The first force, which we call *the coarseness effect*, is directly due to the fact that discrete actions are only a coarse signal about beliefs. We study this force in an environment with two agents, where agent 1 observes the actions of agent 2, while agent 2 observes only her own private signals. Since actions are observed unidirectionally, the actions an agent takes do not influence the information she will receive in future periods. Hence, it is optimal for each agent to behave myopically - that is, to choose the action that maximizes her expected utility in each period.

We find that in the long-run agent 1, who can observe 2's actions, learns as quickly as she would learn if she would observe some fixed fraction of agent 2's private signals (but not her actions). For Normal signals, this fraction equals $9/16 = 56.25\%$. That is, agent 1 learns equally quickly when she observes agent 2's actions (but not her signals) and when she sees 56.25% of agent 2's private signals (but not her actions). We calculate this fraction for arbitrary distributions of signals, and show that it ranges from zero to one, so that, depending on the distribution, it may be the case that almost all of the information is lost, or it may be that almost none is lost. Perhaps surprisingly, this fraction is independent of the agent's utility and set of actions. Especially, it does not depend on how finely one discretizes a continuous action space. Thus, the speed of learning in the continuous action model *cannot* be approximated using the speed of learning in discrete actions models, no matter how fine the discretization.

The second force, which we call *the groupthink effect*, emerges when agents observe each other bidirectionally. In this case an agent's actions depend on her higher order beliefs, and thus observing her actions may reveal less information about her private signals. We study this force in an environment where $n \geq 2$ agents all observe each others' actions. For tractability, we assume that agents are myopic.[1]

Groupthink occurs when a consensus on an action forms in the initial periods, making it optimal for an agent to continue taking the consensus action, even when her private information indicates otherwise. This is similar to the phenomenon of herding in models where the agents act sequentially. We show that typically, after a wrong consensus forms, all agents quickly observe private signals providing strong evidence for choosing the correct action, and yet a long time may pass until any of them breaks the wrong consensus. This leads to long periods of little information aggregation and a slow speed of learning.[2]

With more agents, each individual agent is less likely to break a wrong consensus. On the other hand, the number of potential dissenters is larger, and so a priori it is not obvious whether groupthink becomes more or less likely. We show that the inefficiency (measured as the share of information that is lost) associated with the groupthink effect *increases* with the number of agents. Quantitatively, even as the number of agents goes to infinity, the speed of learning from actions stays bounded by a constant, whereas the speed of learning from the aggregated signals, which is proportional to the number of agents, goes to infinity. Thus, almost all information is lost; the agents' belief has the same precision as would result from observing a vanishingly small fraction of the available private signals. For example, for normal signals, a group of n agents observing each others' actions learns asymptotically slower than a group of 4 agents who share their private signals; this holds for any n! Hence, at most a fraction of $4/n$ of the private information is transmitted through actions.

A natural alternative measure of the speed of learning is the expected discounted payoff loss, relative to a case where all private signals are public. This measure depends on the mistake probabilities in the early periods of the game, which, even for the single agent case, are in general intractable. As a consequence, previous work has studied asymptotic (long run) rates of learning (e.g., Vives, 1993; Chamley, 2004; Duffie and Manso, 2007; Duffie et al., 2009, 2010; Jadbabaie et al., 2013; Molavi et al., 2015), and we do the same. In particular, Jadbabaie et al. (2013) and Molavi et al. (2015) study the rate of learning in an almost identical setting, with boundedly rational agents. Asymptotic rates have been studied in other settings in which it is difficult to analyze the short-term dynamics (e.g., Hong and

[1]Myopic behavior is the common solution concept in much of the social learning literature (for example Sebenius and Geanakoplos, 1983; Parikh and Krasucki, 1990; Bala and Goyal, 1998; Duffie et al., 2010, 2009; Duffie and Manso, 2007; Gale and Kariv, 2003; Vives, 1993); we discuss this further below.

[2]Our prediction seems to be in line with the findings in the empirical literature: Da and Huang (2016, page 5) find in a study on forecasters "that private information may be discarded when a user place weights on the prior forecasts [of others]. In particular, errors in earlier forecasts are more likely to persist and appear in the final consensus forecast, making it less efficient."

Shum (2004) and Hörner and Takahashi (2016)), and more generally asymptotic results are common in the field of learning (e.g., Kalai and Lehrer (1993)).

Using asymptotic rates to quantify the speed of learning has the disadvantage that these rates, in general, convey no information about the initial time periods (and thus expected payoffs). The reason for this is that the rate of learning describes the probability of making a mistake only up to an error term which might be large in early periods, but vanishes in later periods. However, asymptotic rates are tractable, and — as we show — have the further advantage of being independent of many details of the model, providing a measure that is robust to changes in such model parameters as the agents' prior or the exact utility function.

Moreover, the approximation error made in early periods when considering asymptotic rates need not be large. In fact, in some natural examples, the asymptotic rates correspond closely to the short-term dynamics and provide the correct economic intuition. We highlight this by studying a canonical setting of a large group of agents with Normal private signals, where, as the size of the group is increased, the total precision of their signals is kept constant. In this case, our results on the learning rates show that the speed of learning tends to zero as the size of the group grows. By carefully analyzing the learning dynamics in this setting, we show that this result does not only hold in the long-run, but also in the initial periods, where the agents learn less and less, the larger the group gets. Specifically, after revealing some information by their action in the first period, the agents - with high probability for large groups - get locked into groupthink and ignore their subsequent private signals in many of the initial periods, choosing instead to follow the majority opinion of the first period.

Our paper is closely related to models of rational herding (Bikhchandani et al., 1992; Banerjee, 1992; Smith and Sørensen, 2000; Chamley, 2004), as we use the same conditional i.i.d. structure of signals, and utilities depend only on one's own actions and the state. The main difference is that in most herding models, each agent acts only once, whereas in our model, agents take actions repeatedly. An implication of this interaction is a feedback effect where an agent's action today influences other agents' future actions, which in turn change her own future actions. This entails an additional dimension relative to the herding literature: the complexity and importance of higher order beliefs. Agents' actions depend on beliefs of arbitrarily high order, since, unlike in the herding literature, is it not sufficient to reason about others' beliefs, but one must also reason about their beliefs regarding one's own beliefs and so on. A contribution of this paper is to provide an analysis of this interaction, circumventing the calculation of beliefs, which in such contexts is well known to be intractable (as discussed for example by Gale and Kariv (2003)[3]).

In our repeated action setting there may be a strategic incentive to change ones own action in order to gain more information from future actions of others. This effect does not exist for rational myopic agents, who do not value future information, and we assume myopic agents. The same choice is made in most of the learning literature (where signals are private and agents interact repeatedly) either explicitly (e.g., Sebenius and Geanakoplos, 1983; Parikh and Krasucki, 1990; Bala and Goyal, 1998; Keppo et al., 2008), or implicitly, by assuming that there is a continuum of agents (e.g., Vives, 1993; Gale and Kariv, 2003; Duffie and Manso, 2007; Duffie et al., 2009, 2010). A possible justification for this approach is that reasoning about the informational effect of one's actions in such setups requires a level of sophistication that seems unrealistic in many applications.[4]

In the herding literature agents either learn or do not learn the state, depending on whether private signals have bounded likelihood ratios (Smith and Sørensen, 2000). In our model, the distinction between unbounded and bounded private signals is not important, since the aggregate of each agent's private information suffices to learn the state. When agents fail to learn in the herding literature, it is because they disregard their private signals and follow the actions of their predecessors. Similar phenomena are described in other works (for example Vives, 1993; Bala and Goyal, 1998; Mossel et al., 2015), and the same basic mechanism underlies our groupthink effect.

Potential applications of our results appear in settings in which agents repeatedly learn from each other. These include the dissemination of information in developing countries (e.g., Conley and Udry (2010); Banerjee et al. (2013) among many studies), the adoption of opinions on social networks, and prediction markets where forecasters observe the forecasts of others (see Da and Huang (2016)). The main lesson we offer is that too much communication can slow down learning, through the groupthink effect. This lesson may be of interest to those who study these interactions, as well as to those who design them.

[3]Gale and Kariv (2003, p.20): "Speeds of convergence can be established analytically in simple cases. For more complex cases, we have been forced to use numerical methods. The computational difficulty of solving the model is massive even in the case of three persons. However, the results are sufficiently dramatic that they suggest the same might be true for more general cases. This is an important subject for future research."

[4]The strategic experimentation literature studies the incentive to change one's action in order to learn more from others' future actions. We exclude this strategic incentive by assuming that information arrives independently of actions and that agents are myopic. In most of this literature signals are public, and thus all agents share the same belief, whereas for us overcoming the difficulty imposed by complex higher order beliefs is the main challenge.

2. Setup

Time is discrete and indexed by $t \in \{1, 2, \ldots\}$. Each period, each agent $i \in \{1, 2, \ldots, n\}$ first observes a signal (or shock) $s_t^i \in \mathbb{R}$ and then takes an action $a_t^i \in A$. The set of possible actions is finite: $|A| < \infty$.

2.1. **States and Signals.**

There is an unknown state

$$\Theta \in \{l, h\}$$

randomly chosen by nature, with probability $p_0 = \mathbb{P}[\Theta = h] \in (0, 1)$. Signals $s_0^i, s_1^i, \ldots$ are i.i.d, across agents and over time, conditional on the state Θ, with distribution μ_Θ. The distributions μ_h and μ_l are mutually absolutely continuous[5] and hence no signal perfectly reveals the state. As a consequence the log-likelihood ratio of every signal

$$\ell_t^i = \log \frac{\mathrm{d}\mu_h}{\mathrm{d}\mu_l}(s_t^i)$$

is well defined (i.e., $|\ell_t^i| < \infty$) and we assume that it has finite expectation $|\mathbb{E}[\ell_t^i]| < \infty$. We also assume that priors are generic[6], so as to avoid the expository overhead of treating cases in which the agents are indifferent between actions; the results all hold even without this assumption.

Our signal structure allows for bounded as well as unbounded likelihoods. A commonly used special case of our general signal structure are *normal signals* $s_t^i \sim \mathcal{N}(m_\theta, \sigma^2)$ with mean m_θ depending on the state and variance σ^2. Another example is that of *binary signals* $s_t^i \in \{l, h\}$ which are equal to the state with constant probability $\mathbb{P}[s_t^i = \Theta \mid \Theta] = \phi > 1/2$.

2.2. **Actions and Payoffs.**

Agent i's payoff (or utility) in period t depends on her action a_t^i and next period's signal s_{t+1}^i, and is given by $u(s_{t+1}^i, a_t^i)$.[7] The signal can be interpreted as a shock (like demand or interest rate) which influences the payoffs of the different actions of the agent. Note that $u(\cdot, \cdot)$ does not depend on the agent's identity i or the time period t.

This model is equivalent to a model where the agent's utility $\bar{u}(\Theta, a_t^i)$ is unobserved and depends directly on the state. Formally, we can translate the model where the utility depends on the signal into the model where it depends on the state by setting it equal to the expected

[5]That is, every event with positive probability under one measure has positive probability under the other.

[6]That is, chosen from a Lebesgue measure one subset of $[0, 1]$.

[7]Note, that observing the utility $u(s_{t+1}^i, a_t^i)$ does not provide any information beyond the signal s_{t+1}^i and therefore past signals $(s_1^i, \ldots, s_{t+1}^i)$ are a sufficient statistic for the private information available to agent i when taking an action in period $t + 1$.

payoff conditional on the state θ[8]

$$\bar{u}(h, \alpha) := \mathbb{E}_h \left[u(s^i_{t+1}, \alpha)\right]$$
$$\bar{u}(l, \alpha) := \mathbb{E}_l \left[u(s^i_{t+1}, \alpha)\right] .$$

We denote by a^θ the action that maximizes the flow payoff in state θ, which we assume is unique

$$\alpha^\theta := \arg\max_{\alpha \in A} \bar{u}(\theta, \alpha) .$$

We call α^h, α^l the *certainty actions* and assume that they are distinct (i.e., $\alpha^h \neq \alpha^l$), as otherwise the problem is trivial.

2.3. **Information.** Each agent observes only her own signals, and not the signals of others. To learn about the state, agents try to infer the signals of others from their actions. We consider three closely related information structures: i) a single agent acting in autarky, ii) two agents where one agent unidirectionally observes the other's actions, but not vice versa, and iii) the case where n agents observe each others' actions bidirectionally.

2.4. **Agents' Behavior.** We assume throughout that agents are Bayesian and myopic: they completely discount future payoffs, and thus at every time period choose the action the maximizes the payoff at that period. This assumption is without loss of generality, except in the case of bidirectional observation. We denote by p^i_t the posterior probability that agent i assigns to the event $\Theta = h$ at the beginning of period t. As an agent's posterior belief p^i_t is a sufficient statistic for her expected payoff, her action a^i_t depends only on p^i_t. Formally, there exists a function $a^\star \colon [0, 1] \to A$ such that with probability one[9]

$$a^i_t = a^\star(p^i_t) .$$

2.5. **Examples.**

2.5.1. *Matching the State.* A simple example which suffices to understand all the economic results of the paper is the case of two actions $A = \{l, h\}$ where the agent's expected utility equals one if she matches the state, i.e.

$$\bar{u}(\theta, \alpha) = \begin{cases} 1 & \text{if } \alpha = \theta \\ 0 & \text{if } \alpha \neq \theta \end{cases} .$$

[8]Throughout, we denote by $\mathbb{E}_\theta [\cdot] := \mathbb{E} [\cdot \mid \Theta = \theta]$ and $\mathbb{P}_\theta [\cdot] := \mathbb{P} [\cdot \mid \Theta = \theta]$ the expectation and probability conditional on the state.

[9]We here say "with probability one" only to rule out the zero probability event that the agent is indifferent.

In this case the agent simply takes the action to which her posterior belief assigns higher probability:

$$a_t^i = \begin{cases} h & \text{if } p_t^i > \frac{1}{2} \\ l & \text{otherwise} \end{cases} .$$

2.5.2. *Monopolistic Sellers.* As an application, consider local monopolistic sellers who want to learn about the demand for their product and the associated optimal price. Each seller acts in a different market, so that there are no payoff externalities. The distribution of demand, however, is the same, so that the realized demand in other markets is informative about future demands in a seller's home market.

For concreteness, assume that the sellers are shop owners who are selling a new product, and that in the high state the number of people entering the store to inquire about the product is Poisson with mean ρ_h, while in the low state it is Poisson with mean ρ_l, which is less than ρ_h. After learning the price each customer decides whether or not to buy, depending on her private valuation. Customers' private valuations for the product are independent of the state, and so, after having entered the store, customers reveal no new information about the state. Thus, the information a seller learns about the state from her own customers is independent of the price she sets.

When marginal profits are not constant in the volume of sales, a seller will want to set one price if the state is high, another price if the state is low, and potentially intermediate prices when she is unsure about the state. Consequently, each seller wants to learn the state and does so not only by observing the demand in her store, but also by observing the prices set by other sellers.

3. Results

In this section we describe our results on the speed of learning under the different informational assumptions (autarky, unidirectional and bidirectional observation). Section 4 derives the learning dynamics in detail and explains how they lead to the results of this section.

Recall that we consider three information structures: i) a single agent acting in autarky, ii) two agents where one agent can observe the other's actions, but not vice versa, and iii) the case where n agents observe each others' actions bidirectionally.

In all three cases we derive results on the speed at which agents learn the state. More precisely, we consider the probability with which an agent i takes a suboptimal action in period t:

$$a_t^i \neq \alpha^\theta .$$

As the action is suboptimal (given knowledge of the state) we refer to this event as agent i "making a mistake", even though she takes the action which is optimal given her information.

3.1. **Autarky.** In the single agent case, this probability is well known to decay exponentially, with a rate r_a that can be calculated explicitly in terms of the cumulant generating functions $\lambda_h = -\log \mathbb{E}_h\left[\mathrm{e}^{-z\,\ell}\right]$ and $\lambda_l(z) := -\log \mathbb{E}_l\left[\mathrm{e}^{z\,\ell}\right]$:[10]

Theorem 1. *The probability that a single agent in autarky chooses the wrong action in period t satisfies*[11]

$$\mathbb{P}\left[a_t \neq \alpha^{\Theta}\right] = \mathrm{e}^{-r_a\cdot t+o(t)}\,, \tag{1}$$

where

$$r_a := \sup_{z\geq 0} \lambda_h(z) = \sup_{z\geq 0} \lambda_l(z).$$

This type of autarky result is classical and can be found, for example, in studies of Bayesian hypothesis testing (see, e.g. Cover and Thomas (2012, pages 314-316)). For us it serves as a benchmark for the case when agents try to learn from the actions of others.

3.2. **Unidirectional Observation.** When agent 1 can observe agent 2's actions, but not vice versa, we show that her probability of making a mistake decays exponentially as well, and calculate the rate in terms of the Fenchel conjugates[12] (also known as the convex conjugate or Legendre Transform) of λ_θ, $\lambda_\theta^\star(\eta) = \sup_{z\geq 0} \lambda_\theta(z) - z\cdot\eta$.

Theorem 2. *The probability with which agent 1 makes a mistake when she observes all past actions of agent 2 unidirectionally satisfies*

$$\mathbb{P}\left[a_t^1 \neq \alpha^{\Theta}\right] = e^{-r_u\cdot t+o(t)}\,,$$

where

$$r_u := r_a + \min\left\{\lambda_h^\star(r_a), \lambda_l^\star(r_a)\right\}.$$

The rate r_u depends only on the signal distributions μ_l, μ_h, and not on the utility u, the agents' prior, or the set of actions A. Of course, changing the set of actions can have a large impact on the short-term dynamics: a very rich set will reveal much information in the early periods, while not changing the asymptotic speed. The reason that a larger set of actions will not change the speed of learning is that in the long-run only the certainty actions convey information to agent 1 (see Proposition 12).

We find that the rate r_u is always strictly larger than r_a, so that agent 1 learns faster than she would have in autarky. Also, it is always strictly lower than $2\,r_a$, which is the rate at which agent 1 would learn if she could observe all of agent 2's private signals.

[10]Here ℓ is a random variable with a distribution that is equal to that of any of the log-likelihood ratios ℓ_t^i. The definition of the cumulant generating function differs by a sign from the usual one.

[11]Here, and elsewhere, we write $o(t)$ to mean a lower order term. Formally a function $f\colon \mathbb{R}\to\mathbb{R}$ is in $o(t)$ if $\lim_{t\to\infty} f(t)/t = 0$.

[12]Our definition of the Fenchel conjugate differs by a sign from the usual one.

Likewise, if agent 2 were to take continuous actions that reveal her beliefs, then agent 1's speed of learning would be $2\,r_a$. Thus observing discrete actions rather than continuous ones causes a drop in the speed of learning. This drop is independent of how many finite actions are available to agent 2, or what her exact utility function is: for any such choice the learning rate of agent 1 is the same $r_u < 2\,r_a$. We thus call this the *coarseness effect*: a slowdown in learning that is the result of observing coarse actions rather than precise beliefs.

Agent 1 learns from agent 2's actions which state she believes to be more likely. However, agent 1 does not learn the exact strength of agent 2's belief. This information, about the certainty of agent 2, would have been useful to agent 1 as it would have allowed her to decide whether to follow her own private signal or agent 2's last period action in case the two disagree. This loss of information causes learning to be slower than it would have been if agent 1 were to observe 2's signals directly.

It is important to note that the rate r_u with which agent 1 learns has a clear economic meaning also for intermediate values $(r_a, 2\,r_a)$: If the rate equals r_u then the speed at which the agent learns equals the speed at which she would learn, if she were to observe the fixed fraction

$$\frac{r_u}{r_a} - 1$$

of agent 2's signals. To illustrate this, we consider the example of Normal signals:

Proposition 3. *Let μ_θ be the normal distribution with mean m_θ and variance $\sigma^2 > 0$. In this case $r_u = \frac{25}{16} r_a$.*

This implies that agent 1 learns as fast as she would learn if she observed $9/16 \approx 56\%$ of agent 2's private signals, instead of her actions. Equivalently, $7/16 \approx 44\%$ of agent 2's private information is lost en route to agent 1, due to the coarseness effect.

An interesting class of signal structures is that of symmetric distributions μ_l, μ_h, such that the induced log likelihood ratio ℓ conditioned on $\Theta = h$ is identical to the distribution of $-\ell$, conditioned on $\Theta = l$. For example, normal private signals are symmetric, as are binary private signals. In this case we find that agent 1 asymptotically learns at least as quickly as she would learn if she were able to observe half of agent 2's signals directly

Proposition 4. $r_u > {}^3\!/_2\, r_a$ *whenever the signal distributions are symmetric.*

We further discuss symmetric distributions in Section 4.5.

In Section 4.6 we discuss other private signals such as Poisson, binary and trinary signals, and show that the calculation of these rates is often tractable. Numerically, we find that for trinary signals, r_u can be arbitrarily close to $2\,r_a$ - the case in which almost no information is lost. For (asymmetric) Poisson signals, we find that r_u can be arbitrarily close to r_a, so that almost all information is lost. This shows that our results are sharp, in the sense that no further bounds on the speed of learning can be obtained without imposing additional restrictions on the distributions of signals.

3.3. **Bidirectional Observation.** Finally, we consider the case where n agents observe each others' actions. When $n = 2$, we compare the probability of mistake to the probability of mistake by agent 1 (the observer) in the unidirectional observation case. We find that this probability is exponentially higher:

Theorem 5. *If agent 1 and agent 2 observe each other bidirectionally, there is an $r_2 < r_u$ such that*

$$\mathbb{P}\left[a_t^1 \neq \alpha^\Theta\right] = \mathbb{P}\left[a_t^2 \neq \alpha^\Theta\right] \geq \mathrm{e}^{-r_2 \cdot t + o(t)}.$$

Thus, the fact that agent 2 can now also observe agent 1 hurts agent 1 in the long run, as compared to the setting in which agent 2 did not observe agent 1. It should be again emphasized that this statement does not imply anything about the probability of error in early periods; as a statement about exponential rates it claims that for large enough t, the probability of mistake by agent 1 in the bidirectional case will be exponentially larger than in the unidirectional case. Furthermore, agent 2 learns potentially at a faster rate in the bidirectional case as she might benefit from observing agent 1's actions.

Our second main result is that for any number of agents the speed of learning is bounded from above by a constant:

Theorem 6. *Suppose n agents all observe each others' past actions. Given the private signal distributions, there exists a constant $\bar{r}_b > 0$ such that for any number of agents*

$$\mathbb{P}\left[a_t^i \neq \alpha^\theta\right] \geq \mathrm{e}^{-\bar{r}_b \cdot t + o(t)}.$$

Thus adding more agents (and with them more private signals and more information) cannot boost the speed of learning past some bound, and as n tends to infinity more and more of the information is lost. In the case of normal signals $\bar{r}_b = 4\,r_a$, and thus, regardless of the number of agents, the probability of mistake is eventually higher than it would be if 4 agents shared their private signals. Thus for large groups most of the private signals are effectively lost.

To prove both of these theorems we calculate the asymptotic probability of the event that all agents choose the wrong certainty action in almost all time periods up to time t. We call this event groupthink and show that its probability is already high, which implies that the probability that one particular agent errs at time t is also high. Intuitively, when a wrong consensus forms by chance in the beginning, it is hard to break and can last for a long time, with surprisingly high probability. This is due to the fact that agents require their private signals to be relatively strong in order to choose a dissenting action.

In fact, conditioned on groupthink, it holds, with high probability, that the private signals of *each* agent, which initially indicated the wrong action, eventually indicate the *correct action*, but are still ignored due to the overwhelming information provided by the actions of others (Theorem 29). We thus find the term groupthink an apt description of the phenomenon.

Note again, that Theorem 6 is a statement about asymptotic rates. In fact, if one increases the number of agents while holding the private signal distributions fixed, the probability of the agents choosing correctly at any given period $t > 1$ approaches 1. Still, the rate is bounded, and so the probability of a mistake at later time periods is higher than what one might expect.

3.4. **An Analysis of Early Period Mistake Probabilities.** One could alternatively increase the number of agents while holding the total amount of information available to them constant. As an example, we consider n agents who each receive Normal private signals with fixed conditional means ± 1 and variance n. If such signals were publicly observable they would be informationally equivalent to a single Normal signal with variance 1 each period.

Thus, as the total amount of information available is bounded, the probability of choosing correctly in the early periods does not tend to one as the number of agents increases. Furthermore, as each agent learns at most as much from seeing the other agents' actions as she would from seeing 3 other agents' signals (see Section 4.6.1), and the informativeness of private signals goes to zero as $n \to \infty$, the speed of learning tends to zero.

Moreover, a detailed analysis in the case where agents want to match the state (Section 2.5.1) shows that already in the first periods, as the number of agents increases they learn less and less from each other's actions, and so the asymptotic result "kicks in" early on (in the second period):

Theorem 7. *Suppose n agents have normal private signals with conditional distributions $\mathcal{N}(\pm 1, n)$ and want to match the state, so that $\bar{u}(\theta, a) = \mathbf{1}_{\{a=\theta\}}$. Then, for every t, the probability that all agents in the periods $\{2, 3, \ldots, t\}$ choose the action that the majority of the agents chose in period 1 converges to one as n goes to infinity.*

Thus the private signals of periods $\{2, \ldots, t\}$ are with high probability not used in these periods. Consequently, the actions in these periods are correct only if the action taken by the majority in the first period is correct. This probability is bounded by $\Phi(1) \sim 0.84$ for any n. Of course, this probability can be arbitrarily close to $1/2$ if the private signal distributions has a larger variance. In this case, almost all information is lost even in early periods, if the number of agents is sufficiently high. In Figure 3.1 we show how the probability that all agents take the wrong action in the second period increases when the number of agents increases.

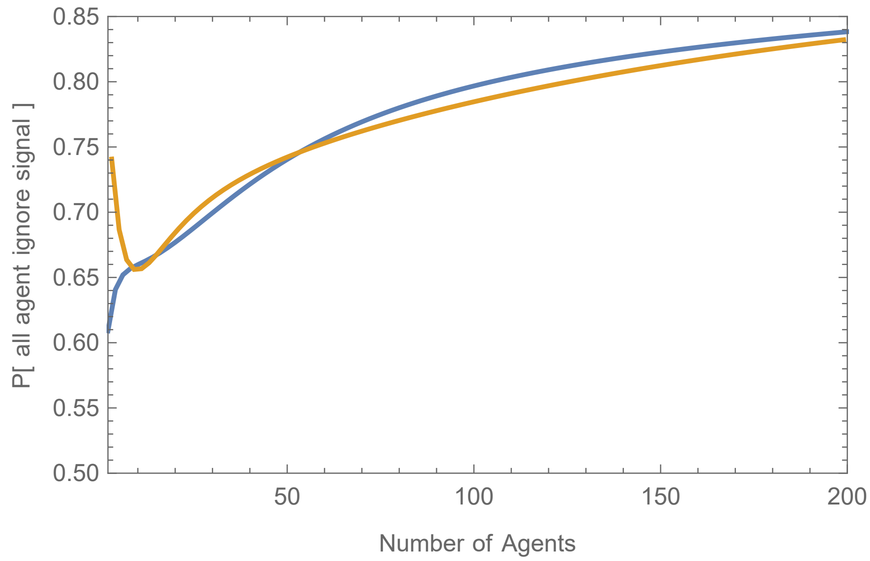

Figure 3.1. In the setting of Theorem 7, this figure charts the probability that all agents ignore their private signal in the second period, choosing instead to follow the choice of the majority in the first period, as a function of the number of agents. We plot even (blue line) and odd (orange line) numbers of agents in different colors to make the chart more readable.

4. Learning Dynamics

In this section we analyze the learning dynamics under the different informational assumptions, in detail. We explain how agents interpret each other's actions and how they choose their own. The analysis of these learning dynamics is related to questions in random walks and large deviations theory. Proving our results requires some mathematical innovation, the details of which we defer to the appendix.

4.1. **Preliminaries.** As an agent's objective is linear in her posterior belief p_t^i, the set of beliefs where she takes a given action is an interval. It will be convenient to define the agent's log-likelihood ratio (LLR) $L_t^i := \log p_t^i/(1-p_t^i)$. As the LLR is a monotone transformation of the agent's posterior belief, and as a myopic agent's action is determined by her posterior, the same holds true in terms of LLRs. This can be summarized in the following claim.

Lemma 8. *There exist disjoint intervals* $(\underline{L}(\alpha), \overline{L}(\alpha)) \subset \mathbb{R} \cup \{-\infty, +\infty\}$, *one for each action* $\alpha \in A$, *such that, with probability one,* $a_t^i = \alpha$ *if and only if* $L_t^i \in (\underline{L}(\alpha), \overline{L}(\alpha))$.

To characterize the agent's actions it thus suffices to characterize her LLR. Note, that for the certainty action α^l it holds that $\underline{L}(\alpha^l) = -\infty$, and that analogously $\overline{L}(\alpha^h) = +\infty$.

4.2. **Autarky.** As a benchmark, we first describe the autarky setting where a single agent acts by himself. In this section we omit the superscript signifying the agent.

Evolution of Beliefs. In autarky, the posterior probability the agent assigns to the high state before taking an action in period t is $P_t = \mathbb{P}[\theta = h \mid s_1, \ldots, s_t]$. Applying Bayes' rule yields that the LLR L_t follows a random walk with increments $\ell_t = \log \frac{\mathrm{d}\mu_h}{\mathrm{d}\mu_l}(s_t)$ equal to the LLR of the signals the agent observed:

$$L_t = L_0 + \sum_{\tau=1}^{t} \ell_\tau \,. \tag{2}$$

Probability of Mistakes. As a consequence of Lemma 8, the probability that the agent chooses the wrong action in period t when the state equals θ is given by

$$\mathbb{P}_\theta\left[a_t \neq \alpha^\theta\right] = \begin{cases} \mathbb{P}_h\left[L_t \leq \underline{L}(\alpha^h)\right] & \text{if } \theta = h \\ \mathbb{P}_l\left[L_t \geq \overline{L}(\alpha^l)\right] & \text{if } \theta = l \end{cases} . \tag{3}$$

Hence, to calculate the probability of a mistake one needs to calculate the probability that the LLR is in a given interval. By (2) the LLR is the sum of increments which are i.i.d. conditional on the state, and hence $(L_t)_t$ is a random walk.

The short-run probability that a random walk is within a given interval is hard to calculate and depends very finely on the distribution of its increments.[13] As this makes it impossible – even in the single agent case – to obtain any general results on the probability that the agent makes a mistake, we focus on the long-run probability of mistakes, which can be analyzed for general signal structures. The long-run behavior of random walks has been studied in *large deviations theory*, with one of the earliest result due to Cramér (1944), who studied these questions in the context of calculating premiums for insurers. We will use some of the ideas and tools from this theory in our analysis; a self-contained introduction is given in Appendix A.

Beliefs. We define the *private LLR* R_t as the LLR calculated only based on an agent's private signals:

$$R_t := L_0 + \sum_{\tau=1}^{t} \ell_\tau.$$

In the single agent case the private signals are all the available information, so $L_t = R_t$, but this will no longer be the case once we consider more agents. Regardless of the number of agents and the information available to them, the private LLR is a random walk with steps ℓ_t, if we condition on the state. We can therefore use large deviation theory to estimate the probability that the private LLR R_t deviates from its expectation, conditional on the state. Let ℓ have the same distribution as each ℓ_t, and define $\lambda_\theta : [0,1] \to \mathbb{R}$, the cumulant

[13]The only exception are a few cases where the distribution of the LLR L_t is known in closed form for every t, such as the Gaussian case. Even in the Gaussian case it seems to us intractable to calculate in closed form the mistake probability in early periods in the multi-agent case.

generating function of the increments of the LLR in state θ by

$$\lambda_h(z) := -\log \mathbb{E}_h\left[\mathrm{e}^{-z\,\ell}\right] \qquad \lambda_l(z) := -\log \mathbb{E}_l\left[\mathrm{e}^{z\,\ell}\right] \,,$$

and denote its Fenchel conjugate by

$$\lambda^\star_\theta(\eta) := \sup_{z\geq 0} \lambda_\theta(z) - \eta \cdot z.$$

The following lemma (proven in the appendix) derives properties of $\lambda_\theta, \lambda^\star_\theta$ which follow from the fact that ℓ_t is the LLR of the signal in period t.

Lemma 9. *$\lambda_\theta(z)$ and $\lambda^\star_\theta(\eta)$ are finite for all $z \in [0,1]$ and $\eta \in (\mathbb{E}_l[\ell], \mathbb{E}_h[\ell])$. Furthermore,*

$$\lambda_h(z) = \lambda_l(1-z) \text{ and } \lambda^*_h(\eta) = \lambda^*_l(-\eta) - \eta\,. \tag{4}$$

The private log likelihood ratio R_t concentrates around its conditional expectations, which are $\mathbb{E}_l[\ell] \cdot t$ in the low state, and $\mathbb{E}_h[\ell] \cdot t$ in the high state. The asymptotic probability that the private LLR lies anywhere in between these two extremes is given by the next lemma.

Lemma 10. *For any $\mathbb{E}_l[\ell] < \eta < \mathbb{E}_h[\ell]$ it holds that*[14]

$$\mathbb{P}_h\left[R_t \leq \eta \cdot t + o(t)\right] = \mathrm{e}^{-\lambda^\star_h(\eta)\cdot t + o(t)}$$
$$\mathbb{P}_l\left[R_t \geq \eta \cdot t + o(t)\right] = \mathrm{e}^{-\lambda^\star_l(-\eta)\cdot t + o(t)}.$$

The proof of Lemma 10 in the Appendix uses the properties of λ_θ and $\lambda^\star_\theta$ established in Lemma 9 to verify that the increments of the LLR process in both states are such that large deviation theory results are applicable. Lemma 10 allows us to calculate the probability of a mistake conditional on each state, immediately implying the next theorem, from which Theorem 1 follows immediately.

Theorem 11. *The probability that an agent in autarky chooses the wrong action in period t satisfies*

$$\mathbb{P}_\theta\left[a_t \neq \alpha^\theta\right] = \mathrm{e}^{-r_a \cdot t + o(t)}\,, \tag{5}$$

where

$$r_a = \lambda^\star_h(0) = \lambda^\star_l(0) = \sup_{z\geq 0} \lambda_h(z) = \sup_{z\geq 0} \lambda_l(z)\,.$$

This result is classical and can be found, for example, in studies of Bayesian hypothesis testing (see, e.g. Cover and Thomas (2012, pages 314-316)). We introduce it to familiarize the reader with the notation and tools that we will use in the sequel. We also note that

[14] Here each $o(t)$ denotes a different function, so that the first line can be alternatively written as follows: For every $f(\cdot)$ with $\lim_{t\to\infty} f(t)/t = 0$ there exists a $g(\cdot)$ with $\lim_{t\to\infty} g(t)/t = 0$ such that $\mathbb{P}_h\left[R_t \leq \eta \cdot t + f(t)\right] = \mathrm{e}^{-\lambda^\star_h(\eta)\cdot t + g(t)}$.

it is possible to strengthen the results by replacing the lower order $o(t)$ term by $O(\log(t))$ using the Bahadur-Rao exact asymptotics method (see Dembo and Zeitouni (1998, Pages 110-113) for a detailed derivation). However, such precision will provide little additional economic insight while significantly complicating the proofs, and thus we will not pursue it.

Note, that the long-run probability of a mistake does not depend on the cut-offs $\overline{L}(\alpha^l)$ and $\underline{L}(\alpha^h)$ and is hence independent of set of actions A and the utility function u. It is also independent of the prior. Thus quantifying the speed of learning using the exponential rate has both advantages and disadvantages: the rate is independent of many details of the model and depends only on the private signal distributions. It is also tractable and can be explicitly calculated for many distributions (see Section 4.6). However, it is an asymptotic measure and in general does not say anything formally about what happens in early periods. The same is true for many statistical results, like the Central Limit Theorem, which nevertheless provide helpful intuition about what happens in finite periods.

4.3. **Unidirectional Observation and the Coarseness Effect.** In the previous section we analyzed how a single agent learns in autarky. We now turn to a two agent setting where agent 1 learns from the actions of agent 2, who himself acts in autarky. The results of the autarky case will be crucial for understanding what inference agent 1 draws from the actions of agent 2.

In this setting, in addition to her own signals $s_1^1, \ldots, s_t^1$, agent 1 observes agent 2's past actions $a_1^2, \ldots a_{t-1}^2$. Agent 2 only observes her own signals $s_1^2, \ldots, s_t^2$. As agent 2 acts in autarky, she behaves as described in Section 4.2. For example, in the "matching the state" setting (Section 2.5.1), the agents' actions will be given by

$$a_t^2 = \begin{cases} h & \text{if } \mathbb{P}\left[\Theta = h \mid s_1^2, \ldots, s_t^2\right] > \frac{1}{2} \\ l & \text{otherwise} \end{cases}$$

(as in autarky) and

$$a_t^1 = \begin{cases} h & \text{if } \mathbb{P}\left[\Theta = h \mid s_1^1, \ldots, s_t^1, a_1^2, \ldots, a_{t-1}^2\right] > \frac{1}{2} \\ l & \text{otherwise} \end{cases}.$$

Only the Last Action Unidirectionally. To get an intuition let us first assume that agent 1 observes only agent 2's last action a_{t-1}^2. More precisely, we analyze the probability with which agent 1 will take a wrong action in period t after observing her first t private signals and the action of agent 2 at time $t-1$. We study this setup to explain the inference problem of agent 1 and later, in Section 4.3, extend the analysis to the case where agent 1 sees all of agent 2's previous actions.

Bayes rule yields that the LLR of agent 1 when agent 2 takes the action α is given by

$$L_t^1 = R_t^1 + I_t(a_{t-1}^2)\,, \tag{6}$$

where $I_t(a_{t-1}^2)$ is the amount by which agent 1's log-likelihood is shifted when she observes agent 2 take action a_{t-1}^2 in period $t-1$:

$$I_t(\alpha) := \log \frac{\mathbb{P}_h\left[a_{t-1}^2 = \alpha\right]}{\mathbb{P}_l\left[a_{t-1}^2 = \alpha\right]}\,.$$

As the signals of agent 1 and agent 2 are independent, R_t^1 is a random walk conditional on the state. The next proposition shows that there are three different types of inference $I_t(\alpha)$ agent 1 can draw from agent 2's behavior.

Proposition 12. *The function $I_t(a)$ satisfies*

$$I_t(a) = \begin{cases} -r_a \cdot t + o(t) & \text{if } \alpha = \alpha^l \\ +r_a \cdot t + o(t) & \text{if } \alpha = \alpha^h \\ o(1) & \text{if } \alpha \notin \{\alpha^l, \alpha^h\} \end{cases}\,.$$

When agent 2 takes a certainty action $\alpha \in \{\alpha^l, \alpha^h\}$ agent 1 believes that agent 2 has strong evidence for the state in which agent 2's action is optimal. If agent 2 does not take a certainty action $\alpha \notin \{\alpha^l, \alpha^h\}$ agent 1 believes that agent 2 must have gotten a sequence of very uninformative signals as she knows that agent 2's belief is bounded away from certainty. As a consequence the influence that agent 2's action has on agent 1's LLR $I_t(\alpha)$ vanishes for large t in this case.

The fact, that the amount by which a full certainty action of agent 2 shifts agent 1's belief is asymptotically linear in the period t, with slope equal to the rate r_a, follows as, by Theorem 11, the probability of a mistake in autarky vanishes at the rate r_a:

$$\begin{aligned} I_t(\alpha^l) &= \log \frac{\mathbb{P}_h\left[a_{t-1}^2 = \alpha^l\right]}{\mathbb{P}_l\left[a_{t-1}^2 = \alpha^l\right]} = \log \mathbb{P}_h\left[a_{t-1}^2 = \alpha^l\right] - \log \mathbb{P}_l\left[a_{t-1}^2 = \alpha^l\right] \\ &= \log\left(e^{-r_a \cdot t + o(t)}\right) - o(1) \\ &= -r_a \cdot t + o(t)\,. \end{aligned}$$

Intuitively, as agent 1 knows that agent 2, who acts in autarky, will take a suboptimal action approximately with probability $\mathrm{e}^{-r_a \cdot t}$, agent 1 shifts her LLR by approximately $-r_a \cdot t$ when she sees that agent 2 chose α^l, and shifts by $+r_a \cdot t$ when she sees agent 2 chose α^h. When agent 1 sees agent 2 take an action that is not optimal in either state she concludes that agent 2 is uninformed and ignores her action.

Probability of Mistakes. Let us first consider the case of the high state. Recall that the LLR of agent 1 is the sum of the LLRs of her private signals R^1_τ as well as the inference $I_t(a^2_{t-1})$ she draws from agent 2's action

$$L^1_t = R^1_t + I_t(a^2_{t-1}) = \begin{cases} R^1_t - r_a \cdot t + o(t) & \text{if } a^2_{t-1} = \alpha^l \\ R^1_t + r_a \cdot t + o(t) & \text{if } a^2_{t-1} = \alpha^h \\ R^1_t + o(1) & \text{if } a^2_{t-1} \notin \{\alpha^l, \alpha^h\}, \end{cases} \tag{7}$$

where the second equality follows from Proposition 12. As shown in Lemma 8 agent 1 makes a mistake in the high state (i.e., does not choose α^h) whenever her likelihood is below $\underline{L}(a^h)$. Thus, when $a^2_{t-1} = \alpha^l$, agent 1 does not choose α^h whenever $R^1_t \leq r_a \cdot t + o(t)$. We can estimate the probability of this event using Lemma 10: it is $\mathrm{e}^{-\lambda^\star_h(r_a)\cdot t + o(t)}$. A similar calculation for the other two cases yields

$$\begin{aligned} \mathbb{P}_h\left[a^1_t \neq \alpha^h \mid a^2_{t-1} = \alpha\right] &= \mathbb{P}_h\left[L^1_t \leq \underline{L}(\alpha^h) \mid a^2_{t-1} = \alpha\right] \\ &= \begin{cases} \mathrm{e}^{-\lambda^\star_h(+r_a)\cdot t + o(t)} & \text{if } \alpha = \alpha^l \\ \mathrm{e}^{-\lambda^\star_h(-r_a)\cdot t + o(t)} & \text{if } \alpha = \alpha^h \\ \mathrm{e}^{-\lambda^\star_h(0)\cdot t + o(t)} & \text{if } \alpha \notin \{\alpha^l, \alpha^h\} \end{cases} . \end{aligned} \tag{8}$$

To calculate the overall probability of a mistake in state h we calculate the probability with which the three above cases occur in Appendix C. We illustrate this type of calculation here by solving the simplest case where agent 1 chooses a wrong action and agent 2 chooses the correct action α^h. By Theorem 11 the probability that agent 2 chooses the correct action $a^2_{t-1} = \alpha^h$ satisfies

$$\mathbb{P}_h\left[a^2_{t-1} = \alpha^h\right] = 1 - \mathrm{e}^{-r_a + o(t)}$$

As a consequence the probability that agent 1 chooses a wrong action and agent 2 chooses the correct action equals

$$\begin{aligned} \mathbb{P}_h\left[a^1_t \neq \alpha^h \text{ and } a^2_{t-1} = a^h\right] &= \mathbb{P}_h\left[a^1_t \neq \alpha^h \mid a^2_{t-1} = a^h\right] \times \mathbb{P}_h\left[a^2_{t-1} = \alpha^h\right] \\ &= \mathrm{e}^{-\lambda^\star_h(-r_a)\cdot t + o(t)}\left(1 - \mathrm{e}^{-r_a \cdot t + o(t)}\right) \\ &= \mathrm{e}^{-\lambda^\star_h(-r_a)\cdot t + o(t)} . \end{aligned}$$

The analysis of the other two cases in Appendix C uses similar arguments and leads to the following result:

Proposition 13. *The probability that agent 1 makes a mistake if she observes agent 2's last action unidirectionally satisfies*

$$\mathbb{P}_\theta\left[a^1_t \neq \alpha^h\right] = e^{-r_u \cdot t + o(t)} ,$$

where $r_u := r_a + \min\{\lambda_h^\star(r_a), \lambda_l^\star(r_a)\} = \min\{\lambda_l^\star(-r_a), \lambda_h^\star(-r_a)\}$.

Observing All Past Actions Unidirectionally. We have so far determined the rate at which agent 1 makes a mistake when she can observe only the last action of agent 2. As agent 1 can always ignore additional information, it follows that her speed of learning is (weakly) greater when she can observe all past actions of agent 2.

Thus the question is whether agent 1 can learn useful information from agent 2's earlier actions, and whether the rate at which agent 1 learns if she observes all of agent 2's actions is strictly higher than the rate at which she learns when she can only observe agent 2's last action.

Agent 2's last period action reveals which state she considers more likely. Hence, her previous actions will only help to assess how much more likely agent 2 considers that state. This information is helpful for agent 1 as it allows her to ignore agent 2's action whenever she believes her evidence to be relatively weak, which is the case when agent 2 switched her action recently. As a consequence, agent 1 is *strictly* less likely to make a mistake whenever all actions of agent 2 are public information.

Our next result shows, however, that in the long-run this effect is vanishingly small, and agent 1 does not significantly benefit from this information: the rate at which she learns exactly equals the rate when she can only observe the last action of agent 2.

Theorem 14. *The probability with which agent 1 makes a mistake when she observes all past actions of agent 2 unidirectionally satisfies*

$$\mathbb{P}\left[a_t^1 \neq \alpha^\Theta\right] = e^{-r_u \cdot t + o(t)}.$$

Here $r_u = r_a + \min\{\lambda_h^\star(r_a), \lambda_l^\star(r_a)\} = \min\{\lambda_l^\star(-r_a), \lambda_h^\star(-r_a)\}$ is the same rate that is defined in Proposition 13, and thus this theorem is a restatement of Theorem 2.

The proof of Theorem 14 argues that it is, in the long-run, approximately equally likely that agent 2 takes the wrong certainty action in the last period and that agent 2 takes the wrong certainty action in all periods. As we explain in the appendix, this result follows from the same logic as Bertrand's classical Ballot Theorem (Bertrand (1887)), which states that the probability that a random walk deviates significantly from its expectation at time t is comparable to the probability that it deviates at all periods prior to time t. Specifically, the ratio between these probabilities scales like $1/t$, and therefore, having a sub-exponential dependence on t, does not affect the exponential rate.

To understand how fast agent 1 learns in this setting, we first note that $\min\{\lambda_h^\star(r_a), \lambda_l^\star(r_a)\} > 0$, and so $r_u > r_a$ (see Claim 24). Thus the rate of learning when observing agent 1 is strictly higher than when learning in autarky. Second, $r_u < 2r_a$ (Claim 25), so that the speed of learning is strictly less than it would be if agent 1 had access to all of agent 2's private signals.

An illustrative example is that of Normal signals, where $s_t^i \sim \mathrm{N}(m_\theta, \sigma^2)$ for some $m_l \neq m_h$ and $\sigma > 0$. We calculate in Section (4.6.1) that in this case $\lambda_l^\star(\eta) = \lambda_h^\star(\eta) = \rho(\eta/\rho - 1)^2/4$, where $\rho = (m_h - m_l)^2/(2\sigma^2)$. Thus $r_a = \rho/4$, and, as one would expect, the rate of learning in autarky is higher the more separated the private signal distribution. More interestingly, it follows that $r_u = \frac{25}{16} r_a$, for any value of σ^2.

4.4. **Bidirectional Observation and the Groupthink Effect.** In this section we consider $n \geq 2$ agents. Each agent observes a sequence of private signals $s_1^i, \ldots, s_t^i$, and the action taken by other agents in previous periods $(a_\tau^j)_{\tau<t, j\neq i}$. In this setting we prove Theorems 5 and 6. As before, we consider myopic agents who completely discount future payoffs, and thus at each period choose the action that maximizes their expected payoffs at that period. For example, in the "matching the state" setting (Section 2.5.1), the agents' actions will be given by

$$a_t^i = \begin{cases} h & \text{if } \mathbb{P}\left[\Theta = h \mid (s_\tau^2)_{\tau \leq t}, (a_\tau^j)_{\tau<t,\, j\neq i}\right] > \frac{1}{2} \\ l & \text{otherwise} \end{cases}.$$

The Probability that All Agents Make a Mistake in Every Period. Let G_t be the event that all agents choose the action α^l in all time periods up to t:

$$G_t = \cap_{i=1}^n \cap_{\tau=1}^t \left\{a_\tau^i = \alpha^l\right\}.$$

To simplify the exposition we assume in the main text that G_t has strictly positive probability.[15]

Conditioned on $\Theta = h$, the event G_t is the event that all the agents are, and always have been, in unanimous agreement on the *wrong* action α^l. We thus call G_t the *groupthink* event. The probability of G_t provides a lower bound on the probability that an agent makes a mistake in period t, conditioned on $\Theta = h$.

This event can be written as $G_t^1 \cap \cdots \cap G_t^n$, where G_t^i is the event that agent i chooses the wrong action α^l in every period $\tau \leq t$. To calculate the probability of G_t, it would of course have been convenient if these n events were independent, conditioned on Θ. However, due to the fact that the agents' actions are strongly intertwined, these events are not independent; given that agent 1 played α^l in all time periods, agent 2 is more likely to do the same. This

[15] This is the case, for example, if the prior is not too extreme relative to the maximal possible private signal strength, or if the private signals are unbounded. Otherwise, it may be the case that agents never take the wrong certainty action in some initial periods, for example if the prior is extreme and the private signals are weak. In Appendix E we drop this assumption, slightly change the definition of G_t, and formally show that all our results also hold in general.

poses a difficulty in the analysis of this model that is a direct consequence of the fact that the agents' actions are intricately dependent on their higher order beliefs.

Decomposition in Independent Events. Perhaps surprisingly, it turns out that G_t can nevertheless be written as the intersection of conditionally independent events. We now describe how this can be done.

Lemma 15. *There exists a threshold $(q_\tau)_\tau$ such that the event G_t equals the event that no agent's private LLR R^i hits the threshold q before period t*

$$G_t = \bigcap_{i=1}^{n} \{R^i_\tau \leq q_\tau \textit{ for all } \tau \leq t\}\,.$$

The proof of Lemma 15 in Appendix E shows this result recursively. Intuitively, whenever G_{t-1} occurs, all agents take action α^l up to time $t-1$. Hence, conditional on G_{t-1}, whether agent i takes the action α^l at time t depends only on R^i_t. As α^l is the most extreme action it follows that the set of private LLRs where the agent takes the action α^l must be a half-infinite interval and is thus characterized by a threshold q. By symmetry, this is the same threshold for all agents.

Calculating the Thresholds. To calculate the q_t's we consider agent j's decision problem at time t, conditioned on G_{t-1}. The information available to her is her own private signals (and in particular her private log-likelihood ratio R^j_t, which is a sufficient statistic for Θ), and in addition the fact that all other agents have chosen α^l up to this point. But the latter observation is equivalent to knowing that all the other agent's private log-likelihood ratios have been under the thresholds q in all previous time periods! Formally, knowing G_{t-1} is equivalent to knowing that

$$W^i_{t-1} := \{R^i_\tau \leq q_\tau \text{ for all } \tau \leq t-1\}$$

has occurred for all $1 < i \leq n$.

This leads to the following recursive characterization of the threshold q.

Proposition 16. *The threshold q_t is characterized by the recursive relation*

$$q_t = \overline{L}(\alpha^l) - (n-1) \cdot \log \frac{\mathbb{P}_h\left[W^1_{t-1}\right]}{\mathbb{P}_l\left[W^1_{t-1}\right]}\,. \tag{9}$$

Having established this additional connection between the events W^i_t and the thresholds $(q_\tau)_\tau$, we next explain that q_t is asymptotically linear with some slope q, so that

$$q_t = q \cdot t + o(t),$$

and estimate the probability of the event W^i_t in terms of q.

We show in the appendix (Claim 27) that $\mathbb{P}_l\left[W^i_{t-1}\right]$ is essentially a constant. It follows that if we take the limits of (9) divided by t, the contribution of the denominator on the right hand side vanishes, and we arrive at

$$\lim_{t\to\infty}\frac{q_t}{t} = -(n-1)\cdot\lim_{t\to\infty}\frac{1}{t}\log\mathbb{P}_h\left[W^i_t\right]. \tag{10}$$

Note, that it remains to be explained why this limit exists, which indeed we do, in Appendix E. Thus, if we denote the left hand side of the above display by $q = \lim_{t\to\infty} q_t/t$, then

$$\mathbb{P}_h\left[W^i_t\right] = \mathrm{e}^{-\frac{q}{n-1}\cdot t+o(t)}. \tag{11}$$

Since $G_t = \bigcap_{i=1}^n W^i_t$, and since these W^i_t's are conditionally independent, we have that

$$\mathbb{P}_h\left[G_t\right] = \mathbb{P}_h\left[W^i_t\right]^n = \mathrm{e}^{-q\cdot\frac{n}{n-1}\cdot t+o(t)}. \tag{12}$$

We thus turn to calculating q. To this end we show in the appendix (Theorem 23) that the probability of the event W^i_t, which is the event that the private LLR R^i_τ is below the threshold q_τ for all $\tau \leq t$, is comparable to the probability of the subevent $\{R^i_t \leq q_t\}$ that the private LLR is below the threshold just at the last time period t. In particular, both have the same rate, conditioned on $\Theta = h$. The large deviations estimate (Lemma 10) implies that

$$\mathbb{P}\left[R^i_t \leq q_t\right] = \mathrm{e}^{-\lambda^\star_h(q)\cdot t+o(t)},$$

and so, since W^i_t has the same rate as $\{R^i_t \leq q_t\}$,

$$\mathbb{P}_h\left[W^i_t\right] = \mathrm{e}^{-\lambda^\star_h(q)\cdot t+o(t)}$$

as well. Thus it follows from (11) that q, the asymptotic slope of the thresholds (q_t), is the solution to the fixed point equation

$$q = (n-1)\lambda^\star_h(q). \tag{13}$$

Note that q depends only on the private signal distributions, through $\lambda^\star_h$. Since $\lambda^\star_h$ is non-negative and decreasing, this equation will always have a unique solution. For example, when $n = 2$ and private signals are normal, we get that $q = 4\frac{\sqrt{2}-1}{\sqrt{2}+1}r_a \approx 0.69 r_a$.

Having calculated q, we now also know from (12) the rate

$$r_g = \frac{n}{n-1}q$$

of the event G_t that all agents take the wrong action in all periods up to time t. This provides a bound on the speed of learning in this setting, conditioned on $\Theta = h$.

A simple convexity argument now yields the following claim.

Claim 17. For any number of agents n it holds that $r_g < \mathbb{E}_h\left[\ell\right]$.

Noting, that

$$\mathbb{P}_h\left[a_t^i \neq \alpha^h\right] \geq \mathbb{P}_h\left[G_t\right] = \mathrm{e}^{-r_g\cdot t+o(t)},$$

and repeating this calculation when conditioning on the low state, we have proved Theorem 6, for $\bar{r}_b = \min\{\mathbb{E}_h[\ell], -\mathbb{E}_l[\ell]\}$. In the case of Normal private signals, $\bar{r}_b = 4r_a$.

In the case of $n = 2$ agents, we get that q is given by the fixed point equation $q = \lambda_h^\star(q)$, and that

$$\mathbb{P}_h\left[a_t^i\right] \geq \mathbb{P}_h\left[G_t\right] = \mathrm{e}^{-2q\cdot t+o(t)}.$$

In the normal signal case this rate is about $1.37r_a$, and is in particular less than $r_u = \frac{25}{16}r_a \approx 1.56r_a$. The following claim states that this is true in general:

Claim 18. $2\cdot q < r_u$.

This completes the proof of Theorem 5 and the analysis of the case of bidirectional observations. The next section will consider the special case of symmetric private signal distributions, in which more can be said about the speeds of learning.

4.5. **Symmetric Private Signal Distributions.** Many natural examples of private signal distribution are, in a sense, invariant to renaming the states. Formally, we say that the private signal distributions are symmetric if the distribution of the induced log likelihood ratio ℓ conditioned on $\Theta = h$ is identical to the distribution of $-\ell$, conditioned on $\Theta = l$. This, for example, is the case for Normal private signals where $\mu_\theta = \mathcal{N}(m_\theta, \sigma^2)$. For symmetric signals we can make a few additional observations.

First, we note that symmetry implies that $\lambda_h = \lambda_l$ and likewise $\lambda_h^\star = \lambda_l^*$. We thus, in this section, omit the subscripts. Since in general $\lambda_h(z) = \lambda_l(1-z)$, we have that in this symmetric case $\lambda(z) = \lambda(1-z)$, or that λ is symmetric to reflection around $z = 1/2$.

Recall that r_a, the rate of learning in autarky, is equal to $\lambda^\star(0) = \max_{z\geq 0}\lambda(0)$. Since $\lambda(z) = \lambda(1-z)$, and since λ is concave, this maximum is attained at $z = 1/2$, and so we have that $r_a = \lambda(1/2)$. It thus follows from (16) that

$$r_a = -\log\int\sqrt{\frac{\mathrm{d}\mu_l}{\mathrm{d}\mu_h}(s)}\mathrm{d}\mu_h(s).$$

This is also known as the Bhattacharyya distance (Bhattacharyya (1943)) between the two distributions, and is a commonly used measure of the distance between distributions.

Recall that $r_u = \lambda^*(-r_a) = \max_{z\geq 0}\lambda(z) + r_a z$. Substituting $z = 1/2$ yields

$$r_u \geq \lambda(1/2) + r_a/2 = \frac{3}{2}r_a.$$

Since the maximum is not attained at $z = 1/2$ (since $\lambda'(1/2) = 0 \neq r_a$), we have strict inequality, and so $r_u > \frac{3}{2}r_a$. Thus, in the symmetric case there is a stronger lower bound on the speed of learning in the unidirectional case, or, equivalently, an upper bound on how much can be lost due to the coarseness effect. We state this formally in the following claim.

Proposition 19. *When private signals are symmetric then* $r_u > \frac{3}{2} r_a$.

The example of asymmetric Poisson private signals that we consider in the next section shows that this it not true in general, and that r_u can be arbitrarily close to r_a for asymmetric signals.

4.6. **The Speed of Learning for Different Signal Distributions.** In this section we calculate the speeds of learning for some examples of private signal distributions. Our goal is to show that these calculations are often tractable, and to discuss the range of possible values of r_a and r_u, in particular showing that without further restrictions on the signal distributions one cannot obtain results restricting r_u/r_a beyond $(1, 2)$ for general distributions and $(3/2, 2)$ for symmetric distributions.

4.6.1. *Normal Private Signals.* The simplest example is that of Normal private signals, where the distribution of private signals is either $\mathcal{N}(m_l, \sigma^2)$ or $\mathcal{N}(m_h, \sigma^2)$ for some $m_l \neq m_h$. In this case, if we denote $\rho = (m_h - m_l)^2/(2\sigma^2)$, then an easy calculation shows that

$$\lambda_h(z) = -\rho \cdot z \cdot (z - 1)$$

and

$$\lambda_h^\star(\eta) = \frac{\rho}{4}(\eta/\rho - 1)^2$$

for $\eta \leq \rho$ and $\lambda_h^\star(\eta) = 0$ for $\eta \geq \rho$. Substituting the results of Theorems 11 and 14, we have that $r_a = \rho/4$ and $r_u = \frac{25}{16} r_a$. Thus, in the Normal case, learning with unidirectional observations is always faster by a factor of 25/16 than learning on ones own, regardless of the means and variance of the signals.

When there are n agents observing bidirectionally, it is easy to calculate, using (13), that the rate of the groupthink event G_t in which all agents take a wrong action in all time periods up to time t is

$$r_g = 4\frac{(n - \sqrt{n})^2}{(n-1)^2} r_a.$$

In particular, when there are only two agents, $r_g = 8\frac{\sqrt{2}-1}{\sqrt{2}+1} r_a \approx 1.37\, r_a$, which is less than r_u, as guaranteed by Claim (18).

For n agents, the upper bound on the speed of learning, $\bar{r}_b$, is, in this case, equal to ρ, which is equal to $4 \cdot r_a$, so that learning from actions in a group of any size is slower than learning directly from the private signals of 4 agents.

4.6.2. *Poisson Private Signals.* When private signals are Poisson with mean either ρ_l or ρ_h then the expressions for λ_h, $\lambda_h^\star$, r_a and r_u are rather complex. However, they can still be calculated analytically, for example by using Mathematica. Mathematica's symbolic engine can also be used to formally show that for fixed ρ_l, as one takes larger and larger ρ_h, the ratio r_u/r_a tends to one. Thus, in the case of very different means of the two distributions, learning in the unidirectional case can be negligibly faster than learning on ones own. If we fix ρ_l but this time let ρ_h tend to ρ_l, then r_u/r_a tends to 25/16, as in the Normal case.

4.6.3. *Binary Private Signals.* A commonly used example of private signals is that of symmetric binary signals $s_t^i \in \{l, h\}$ which are equal to the state with constant probability

$$\mathbb{P}\left[s_t^i = \Theta \mid \Theta\right] = \phi > {}^1/_2.$$

In this case

$$\lambda_h(z) = -\log\left[\phi\left(\frac{\phi}{1-\phi}\right)^{-z} + (1-\phi)\left(\frac{\phi}{1-\phi}\right)^{z}\right].$$

If we denote by

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

the Kullback-Leibler divergence between the Bernoulli distributions $B(p)$ and $B(q)$, then we can calculate that

$$r_a = D_{\mathrm{KL}}({}^1/_2||\phi)$$

and

$$r_u = r_a + D_{\mathrm{KL}}\left(\hat{\phi}||\phi\right),$$

where

$$\hat{\phi} = \frac{1}{2}\left(1 + \frac{r_a}{\log(\phi/(1-\phi))}\right).$$

Figure 4.1 shows that r_u/r_a varies between ${}^{25}/_{16}$ (as ϕ approaches ${}^1/_2$ and the signals become uninformative) and ${}^3/_2$ (as ϕ tends to 1). Thus, as in the case of Poisson signals, the ratio r_u/r_a of less and less informative signals approaches that of Normal signals.

4.6.4. *Trinary Private Signals.* Another interesting class of private signal distributions are those that are supported on three values, which we take to be $\{-1, 0, 1\}$. We assume 0 to be uninformative, and thus occurring with the same probability ξ in either state. Intuitively, $1-\xi$ corresponds to the probability of getting an informative signal. Conditional on getting an informative signal, we assume the signal to equal $+1$ with probability p in the high state, and -1 with probability p in the low state. This signal distribution is symmetric, and thus by Proposition 19 we have that $r_u/r_a > {}^3/_2$. Numerical calculations suggest that for any $0 < \xi < 1$, as p tends to 1, the ratio r_u/r_a tends to 2. Thus, in this case, the speed of

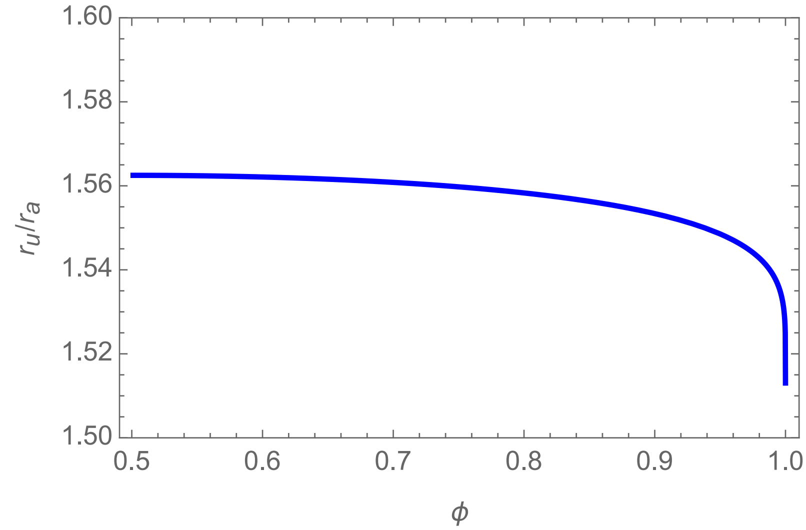

FIGURE 4.1. The ratio r_u/r_a as a function of the precision ϕ in the symmetric binary signal case.

learning from actions is close to the speed of learning from signals, and hence almost no information is lost.

5. Conclusion

We introduce asymptotic rates as a measure of the speed of learning in models of repeated interaction. While asymptotic rates have no formal implications for short run dynamics, they do provide intuition for the forces at work.[16]

This article leaves many open questions which could potentially be analyzed using our approach. What happens when the state changes over time? What happens with payoff externalities, for example when agents have a motive to coordinate? Of particular interest is the study of a more complex societal structure of the agents: how fast do they learn for a given network of observation, which is not the complete network?

References

Venkatesh Bala and Sanjeev Goyal. Learning from neighbours. *The Review of Economic Studies*, 65(3):595–621, 1998.

Abhijit Banerjee, Arun G Chandrasekhar, Esther Duflo, and Matthew O Jackson. The diffusion of microfinance. *Science*, 341(6144):1236498, 2013.

Abhijit V Banerjee. A simple model of herd behavior. *The Quarterly Journal of Economics*, pages 797–817, 1992.

[16]The same can be said for many statistical results. A classical example is the Central Limit Theorem, which provides intuition for the behavior of finite sums of i.i.d. random variables, even though, as an asymptotic statement, it has no formal implication except at the limit.

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

A Bhattacharyya. On a measure of divergence between two statistical population defined by their population distributions. *Bulletin Calcutta Mathematical Society*, 35:99–109, 1943.

Sushil Bikhchandani, David Hirshleifer, and Ivo Welch. A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of Political Economy*, pages 992–1026, 1992.

Christophe Chamley. *Rational herds: Economic models of social learning.* Cambridge University Press, 2004.

Timothy G Conley and Christopher R Udry. Learning about a new technology: Pineapple in ghana. *The American Economic Review*, 100(1):35–69, 2010.

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

Harald Cramér. On a new limit theorem of the theory of probability. *Uspekhi Mat. Nauk*, 10:166–178, 1944.

Zhi Da and Xing Huang. Harnessing the wisdom of crowds. 2016.

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

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

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

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

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

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

Han Hong and Matthew Shum. Rates of information aggregation in common value auctions. *Journal of Economic Theory*, 116(1):1–40, 2004.

Johannes Hörner and Satoru Takahashi. How fast do equilibrium payoff sets converge in repeated games? *Journal of Economic Theory*, 165:332–359, 2016.

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

Ehud Kalai and Ehud Lehrer. Rational learning leads to nash equilibrium. *Econometrica*, pages 1019–1045, 1993.

Jussi Keppo, Lones Smith, and Dmitry Davydov. Optimal electoral timing: Exercise wisely and you may live longer. *The Review of Economic Studies*, 75(2):597–628, 2008.

Pooya Molavi, Alireza Tahbaz-Salehi, and Ali Jadbabaie. Foundations of non-bayesian social learning. *Columbia Business School Research Paper*, 2015.

Elchanan Mossel, Allan Sly, and Omer Tamuz. Strategic learning and the topology of social networks. 2015.

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

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

Lones Smith and Peter Sørensen. Pathological outcomes of observational learning. *Econometrica*, 68(2):371–398, 2000.

Daniel W Stroock. *Mathematics of probability*, volume 149. American Mathematical Society, 2013.

Xavier Vives. How fast do rational agents learn? *The Review of Economic Studies*, 60(2): 329–347, 1993.

Appendix A. The cumulant Generating Functions, their Fenchel Conjugates, and Large Deviations Estimates

Large Deviations of Random Walks. The long-run behavior of random walks has been studied in large deviations theory. We now introduce some tools from this literature, which will be crucial to understanding the long-run behavior of agents.

Let $X_1, X_2, \dots$ be i.i.d random variables with $\mathbb{E}[X_t] = \mu$ and $Y_t = \sum_{\tau=1}^{t} X_t$ the associated random walk with steps X_t. By the law of large numbers we know that Y_t should approximately equal $\mu \cdot t$. Large deviation theory characterizes the probability that Y_t is much lower, and in particular smaller than $\eta \cdot t$, for some $\eta < \mu$. Under some technical conditions, this probability is exponentially small, with a rate $\lambda^\star(\eta)$:

$$\mathbb{P}[Y_t < \eta \cdot t + o(t)] = \mathrm{e}^{-\lambda^\star(\eta)\cdot t + o(t)},$$

or equivalently stated

$$\lim_{t\to\infty} -\frac{1}{t} \log \mathbb{P}[Y_t < \eta \cdot t + o(t)] = \lambda^\star(\eta).$$

The rate $\lambda^\star$ can be calculated explicitly and is the *Fenchel Conjugate* of the *cumulant generating function* of the increments

$$\lambda^\star(\eta) := \sup_{z\geq 0} \left(-\log \mathbb{E}\left[\mathrm{e}^{-z\,X_1}\right] - \eta \cdot z\right).$$

The first proof of a "large deviation" result of this flavor is due to Cramér (1944), who studied these questions in the context of calculating premiums for insurers. A standard textbook on large deviations theory is Dembo and Zeitouni (1998).

In this section we provide an independent proof of this classical large deviations result, and prove a more specialized one suited to our needs. We consider a very general setting: we make no assumptions on the distribution of each step X_t, and in particular do not need to assume that it has an expectation.

Denoting $X = X_1$, The cumulant generating function λ is (up to sign, as compared to the usual definition) given by

$$\lambda(z) = -\log \mathbb{E}\left[\mathrm{e}^{-z\,X}\right].$$

Note that when the right hand side is not finite it can only equal $-\infty$ (and never $+\infty$).

Proposition 20. *λ is finite on an interval I, on which it is concave and on whose interior it is smooth (that is, having continuous derivatives of all orders).*

Proof. Note that I contains 0, since $\lambda(0) = 0$ by definition. Assume $\lambda(a)$ and $\lambda(b)$ are both finite. Then for any $r \in (0,1)$

$$\lambda(r\cdot a + (1-r)\cdot b) = -\log \mathbb{E}\left[\mathrm{e}^{-(r\cdot a+(1-r)\cdot b)\cdot X}\right] = -\log \mathbb{E}\left[\left(\mathrm{e}^{-a\cdot X}\right)^r \cdot \left(\mathrm{e}^{-b\cdot X}\right)^{1-r}\right],$$

which by Hölder's inequality is at least $r\cdot\lambda(a) + (1-r)\cdot\lambda(b)$. Hence λ is finite and concave on a convex subset of $\mathbb{R}$, or an interval. We omit here the technical proof of smoothness; it can be found, for example, in Stroock (2013, Theorem 1.4.16). $\square$

It also follows that unless the distribution of X is a point mass (which is a trivial case), λ is strictly concave on I. We assume this henceforth. Note that it could be that I is simply the singleton $[0,0]$. This is not an interesting case, and we will show later that in our setting I is larger than that.

The Fenchel conjugate of λ is given by

$$\lambda^\star(\eta) = \sup_{z\geq 0} \lambda(z) - \eta\cdot z.$$

We note a few properties of $\lambda^\star$. First, since $\lambda(0) = 0$ and $\lambda(z) < \infty$, $\lambda^\star$ is well defined and non-negative (but perhaps equal to infinity for some η). Second, since λ is equal to $-\infty$ whenever it is not finite, the supremum is attained on I, unless it is infinity. Third, since λ is strictly concave on I, $\lambda(z) - \eta\cdot z$ is also concave there, and so the supremum is a maximum and is attained at a single point $z \in I$ whenever it is finite. Additionally, since λ is smooth on I, this single point z satisfies $\lambda'(z) = \eta$ if $z > 0$ (equivalently, if $\lambda^\star(\eta) > 0$).

I.e., if $\lambda'(z) = \eta$ for some z in the interior of I then

$$\lambda^*(\eta) = \lambda(z) - \eta \cdot z. \tag{14}$$

Finally, it is immediate from the definition that λ^* is weakly decreasing, and it is likewise easy to see that it is continuous. This, together with (14) and the fact that λ' is decreasing, yields that $\lambda^\star(\eta) = \lambda(0) = 0$ whenever $\eta \geq \sup_{z \geq 0} \lambda'(z)$. We summarize this in the following claim.

Proposition 21. *Let I be the interval on which λ is finite, and let $I^\star = \{\eta : \exists z \in \mathrm{int} I \text{ s.t. } \lambda'(z) = \eta\}$. Then*

(1) λ^* is continuous, non-negative and weakly decreasing. It is positive and strictly decreasing on I^*.
(2) $\lambda^\star(\eta) = 0$ whenever $\eta \geq \sup_{z \geq 0} \lambda'(z)$.
(3) If $\eta \in I^\star$ and $\lambda'(z) = \eta$ then $\lambda^*(\eta) = \lambda(z) - \eta \cdot z$.

Given all this, we are ready to state and prove our first large deviations theorem.

Theorem 22. *For every η such that $\eta > \inf_{z \in I} \lambda'(z)$ it holds that*

$$\mathbb{P}\left[Y_t \leq \eta \cdot t + o(t)\right] = \mathrm{e}^{-\lambda^\star(\eta) \cdot t + o(t)}.$$

Proof. For the upper bound, we use a Chernoff bound strategy: for any $z \geq 0$

$$\mathbb{P}\left[Y_t \leq \eta \cdot t + o(t)\right] = \mathbb{P}\left[\mathrm{e}^{-z\,Y_t} \geq \mathrm{e}^{-z \cdot (\eta \cdot t + o(t))}\right],$$

and so by Markov's inequality

$$\mathbb{P}\left[Y_t \leq \eta \cdot t + o(t)\right] \leq \frac{\mathbb{E}\left[\mathrm{e}^{-z\,Y_t}\right]}{\mathrm{e}^{-z \cdot (\eta \cdot t + o(t))}}.$$

Now, note that $\mathbb{E}\left[\mathrm{e}^{-z\,Y_t}\right] = \mathrm{e}^{-\lambda(z) \cdot t}$, and so

$$\mathbb{P}\left[Y_t \leq \eta \cdot t + o(t)\right] \leq \mathrm{e}^{-(\lambda(z) - z \cdot \eta) \cdot t + z \cdot o(t)}.$$

Choosing $z \geq 0$ to maximize the coefficient of t yields

$$\mathbb{P}\left[Y_t \leq \eta \cdot t + o(t)\right] \leq \mathrm{e}^{-\lambda^\star(\eta) \cdot t + o(t)},$$

which is the desired lower bound.

We now turn to proving the upper bound. Denote by ν the law of X, and for some fixed z in the interior of I (to be determined later) define the probability measure $\tilde{\nu}$ by

$$\frac{\mathrm{d}\tilde{\nu}}{\mathrm{d}\nu}(x) = \frac{\mathrm{e}^{-zx}}{\mathbb{E}\left[\mathrm{e}^{-zX}\right]} = \mathrm{e}^{\lambda(z) - zx},$$

and let $\tilde{X}_t$ be i.i.d. random variables with law $\tilde{\nu}$. Note that

$$\mathbb{E}\left[\tilde{X}\right] = \frac{\mathbb{E}\left[X\mathrm{e}^{-zX}\right]}{\mathbb{E}\left[\mathrm{e}^{-zX}\right]} = \lambda'(z).$$

Now, fix any η_1, η_2 such that $\eta_1 < \eta_2 < \eta$ and $\lambda'(z) = \eta_2$ for some z in the interior of I; this is possible since $\eta > \inf_{z \in I} \lambda'(z)$. This is the z we choose to take in the definition of $\tilde{\nu}$. If we think of η_2 as being close to η then the expectation of $\tilde{X}$, which is equal to η_2, is close to η. We have thus "tilted" the random variable X, which had expectation μ, to a new random variable with expectation close to η.

We can bound

$$\mathbb{P}\left[Y_t \leq \eta \cdot t + o(t)\right] \geq \mathbb{P}\left[\eta_1 \cdot t \leq Y_t \leq \eta \cdot t + o(t)\right] = \int_{\eta_1 t}^{\eta t + o(t)} 1 \, \mathrm{d}\nu^{(t)},$$

where $\nu^{(t)}$ is the t-fold convolution of ν with itself, and hence the law of Y_t. It is easy to verify[17] that $\mathrm{d}\nu^{(\mathrm{t})}(\mathrm{y}) = \mathrm{e}^{zy - \lambda(z)\cdot t} \, \mathrm{d}\tilde{\nu}^{(t)}(y)$, and so

$$= \mathrm{e}^{-\lambda(\mathrm{z})\cdot \mathrm{t}} \int_{\eta_1 t}^{\eta t + o(t)} \mathrm{e}^{zy} \, \mathrm{d}\tilde{\nu}^{(t)}(y),$$

which we can bound by taking the integrand out of the integral and replacing y with the lower integration limit:

$$\geq \mathrm{e}^{(\eta_1 p - \lambda(z))\cdot t} \int_{\eta_1 t}^{\eta t + o(t)} 1 \, \mathrm{d}\tilde{\nu}^{(t)}.$$

Since the law of $\tilde{Y}_t = \sum_{\tau=1}^{t} \tilde{X}_t$ is $\tilde{\nu}^{(t)}$, this is equal to

$$= \mathrm{e}^{(\eta_1 z - \lambda(p))\cdot t} \mathbb{P}\left[\eta_1 \cdot t \leq \tilde{Y}_t \leq \eta \cdot t + o(t)\right].$$

Since $\eta_1 < \mathbb{E}\left[\tilde{\mathbb{X}}\right] < \eta$ we have that $\lim_t \mathbb{P}\left[\eta_1 \cdot t \leq \tilde{Y}_t \leq \eta \cdot t + o(t)\right] = 1$, by the law of large numbers. Hence

$$\liminf_{t\to\infty} \frac{1}{t} \log \mathbb{P}\left[Y_t \leq \eta \cdot t + o(t)\right] \geq \eta_1 z - \lambda(z),$$

which, by (14), and recalling that $z = (\lambda')^{-1}(\eta_2)$, can be written as

$$\liminf_{t\to\infty} \frac{1}{t} \log \mathbb{P}\left[Y_t \leq \eta \cdot t + o(t)\right] \geq -\lambda^*(\eta_2) - (\eta_2 - \eta_1) \cdot (\lambda')^{-1}(\eta_2).$$

Taking the limit as η_1 approaches η_2 yields

$$\liminf_{t\to\infty} \frac{1}{t} \log \mathbb{P}\left[Y_t \leq \eta \cdot t + o(t)\right] \geq -\lambda^*(\eta_2). \tag{15}$$

We now consider two cases. First, assume that $\eta \leq \sup_{z\geq 0} \lambda'(z)$. In this case we can choose η_2 arbitrarily close to η, and by the continuity of λ^* we get that

[17] See, e.g., Durrett (1996, Page 74) or note that the Radon-Nikodym derivative between the law of X and $\tilde{X}$ is $\mathrm{e}^{zx-\lambda(z)}$, and so the derivative between the laws of $(X_1, \ldots, X_t)$ and $\left(\tilde{X}_1, \ldots, \tilde{X}_t\right)$ is $\mathrm{e}^{z(x_1+\cdots+x_t)-\lambda(z)\cdot t}$.

$$\liminf_{t\to\infty} \frac{1}{t}\log \mathbb{P}\left[Y_t \leq \eta\cdot t + o(t)\right] \geq -\lambda^\star(\eta),$$

or equivalently

$$\mathbb{P}\left[Y_t \leq \eta\cdot t + o(t)\right] \geq \mathrm{e}^{-\lambda^\star(\eta)\cdot t + o(t)}.$$

The second case is that $\eta > \sup_{z\geq 0}\lambda'(z)$. In this case $\lambda^\star(\eta) = 0$ (Proposition 21). Also, (15) holds for any $\eta_2 < \sup_z \lambda'(z)$ and thus it holds for $\eta_2 = \sup_{z\geq 0}\lambda'(z)$. But then $\lambda^\star(\eta_2) = 0 = \lambda^\star(\eta)$, and so we again arrive at the same conclusion. □

The next Theorem is similar in spirit, and in some sense is stronger than the previous, as it shows that the same rate applies to the event that the sum is below the threshold at all time periods prior to t, rather than just at period t. It furthermore does not require the threshold to be linear, but only asymptotically and from one direction; both of these generalizations are important. This theorem will be useful in analyzing both the unidirectional setting and the bidirectional setting. Theorem 23 is not an established result, but rather a (modest) contribution of this paper.

Theorem 23. *For every η such that $\eta > \inf_{z\in I}\lambda'(z)$, and every sequence $\{y_t\}_{t\in\mathbb{N}}$ with $\liminf_t y_t/t = \eta$ and $\mathbb{P}\left[Y_t \leq y_t\right] > 0$ it holds that*

$$\mathbb{P}\left[\cap_{\tau=1}^{t}\{Y_\tau \leq y_\tau\}\right] = \mathrm{e}^{-\lambda^\star(\eta)\cdot t + o(t)}.$$

Proof. Let E_t be the event $\cap_{\tau=1}^{t}\{Y_\tau \leq y_\tau\}$. Let $\{t_k\}$ be a sequence such that $\lim_k y_{t_k}/t_k = \eta$. For every t let t' be the largest t_k with $t_k \leq t$. Then by inclusion we have that

$$\frac{1}{t}\log \mathbb{P}\left[E_t\right] \leq \frac{1}{t'}\log \mathbb{P}\left[Y_{t'} \leq y_{t'}\right].$$

Using the same Chernoff bound strategy of the proof of Theorem 22, we get that

$$\frac{1}{t}\log \mathbb{P}\left[E_t\right] \leq -\lambda^\star\left(y_{t'}/t'\right).$$

The continuity of λ implies that taking the limit superior of both sides yields

$$\limsup_t \frac{1}{t}\log \mathbb{P}\left[E_t\right] \leq -\lambda^\star\left(\eta\right),$$

or

$$\mathbb{P}\left[E_t\right] \leq \mathrm{e}^{-\lambda^\star(\eta)\cdot t + o(t)}.$$

To show the other direction, define (as in the proof of Theorem 22) $\tilde{X}_t$ to be be i.i.d. random variables with law $\tilde{\nu}$ given by

$$\frac{\mathrm{d}\tilde{\nu}}{\mathrm{d}\nu}(x) = \mathrm{e}^{\lambda(z) - zx},$$

where ν is the law of X, and $z \in I$ is chosen so that $\lambda'(z) = \eta_2$ for some $\eta_1 < \eta_2 < \eta$. Denoting $\epsilon = \eta - \eta_1$, it follows from inclusion that

$$\mathbb{P}[E_t] \geq \mathbb{P}[E_t \cap \{Y_t \geq y_t - \epsilon \cdot t\}] .$$

Now, the Radon-Nikodym derivative between the laws of $(X_1, \ldots, X_t)$ and $\left(\tilde{X}_1, \ldots, \tilde{X}_t\right)$ is $\mathrm{e}^{z(x_1+\cdots+x_t)-\lambda(z)\cdot t}$. Hence

$$\mathbb{P}[E_t] \geq \mathbb{E}[1_{E_t} \cdot 1_{Y_t \geq y_t - \epsilon\cdot t}] = \mathbb{E}\left[1_{\tilde{E}_t} \cdot 1_{\tilde{Y}_t \geq y_t-\epsilon\cdot t} \cdot \mathrm{e}^{z\tilde{Y}_t - \lambda(z)\cdot t}\right],$$

where $\tilde{E}_t$ is the event $\cap_{\tau=1}^{t}\left\{\tilde{Y}_\tau \leq y_\tau\right\}$. We can bound this expression by taking $\mathrm{e}^{z\tilde{Y}_t-\lambda(z)\cdot t}$ out of the integral and replacing it with the lower bound $y_t - \epsilon \cdot t$. This yields

$$\mathbb{P}[E_t] \geq \mathrm{e}^{z(y_t-\epsilon\cdot t)-\lambda(z)\cdot t} \cdot \mathbb{P}\left[\tilde{E}_t \cap \left\{\tilde{Y}_t \geq y_t - \epsilon \cdot t\right\}\right].$$

Now, since the expectation of $\tilde{Y}_t/t$ is strictly between $\eta = \liminf_t y_t/t$ and $\eta - \epsilon$, we have that $\lim_t \mathbb{P}\left[\tilde{Y}_t \geq y_t - \epsilon \cdot t\right] = 1$ by the weak law of large numbers. By the strong law of large numbers and the Markov Property of $\left\{\tilde{Y}_t\right\}$ we have that $\lim_t \mathbb{P}\left[\tilde{E}_t\right] > 0$; $\left\{\tilde{Y}_t\right\}$ is indeed Markov since $\left\{\tilde{X}_t\right\}$ are i.i.d. Thus $\lim_t \mathbb{P}\left[\tilde{E}_t \cap \left\{\tilde{Y}_t \geq y_t - \epsilon \cdot t\right\}\right] > 0$ and

$$\liminf_t -\frac{1}{t}\log \mathbb{P}[E_t] \leq z\cdot\eta_1 - \lambda(z).$$

Proceeding as in the proof of Theorem 22 yields that

$$\mathbb{P}[E_t] \geq \mathrm{e}^{-\lambda^\star(\eta)\cdot t + o(t)}. \qquad \square$$

Appendix B. Application of Large Deviation Estimates

In this section we prove a number of claims regarding the functions λ_θ and λ^*_θ. Recall that for $\theta \in \{h, l\}$

$$\lambda_h(z) := -\log\, \mathbb{E}_h\left[\mathrm{e}^{-z\,\ell}\right] \qquad \lambda_l(z) := -\log\, \mathbb{E}_l\left[\mathrm{e}^{z\,\ell}\right],$$

where ℓ is a random variable with the same law as any ℓ^i_t, and

$$\lambda^\star_\theta(\eta) = \max_z \lambda_\theta(z) - \eta\cdot z.$$

We first note that by the definition of λ_θ we have that

$$\lambda_h(z) = -\log \int \exp\left(-z\cdot\log\frac{\mathrm{d}\mu_{\mathrm{h}}}{\mathrm{d}\mu_l}(s)\right)\mathrm{d}\mu_h(s) = -\log\int\left(\frac{\mathrm{d}\mu_{\mathrm{l}}}{\mathrm{d}\mu_h}(s)\right)^z \mathrm{d}\mu_h(s). \tag{16}$$

It follows immediately that there is a simple connection between λ_h and λ_l

$$\lambda_l(z) = \lambda_h(1-z).$$

It also follows from (16) that the interval I on which λ_h is finite contains $[0,1]$. Since from the definitions we have that $\lambda_h'(0) = \mathbb{E}_h[\ell]$, and since $\lambda_h'(1) = \mathbb{E}_l[\ell]$ by the relation between λ_h and λ_l, two immediate corollaries of Theorem 22 are Lemmas 9 and 10. Furthermore, as for every η between $\mathbb{E}_h[\ell]$ and $\mathbb{E}_l[\ell]$ the maximum in the definition of $\lambda_h^\star$ is achieved for some $z \in (0,1)$, it follows that there is also a simple connection between $\lambda_h^\star$ and $\lambda_l^\star$:

$$\lambda_l^\star(\eta) = \lambda_h^\star(-\eta) - \eta. \tag{17}$$

We will accordingly state some results in terms of λ_h and $\lambda_h^\star$ only.

The following simple observation will be useful on several occasions:

Claim 24. Let $r_a = \lambda_h^\star(0)$. Then $r_a = \max_{z\in(0,1)} \lambda_h(z) = \max_{z\in(0,1)} \lambda_l(z) = \lambda_l^*(0)$, $r_a < \min\{\mathbb{E}_h[\ell], -\mathbb{E}_l[\ell]\}$, and $\min\{\lambda_h^\star(r_a), \lambda_l^\star(r_a)\} > 0$.

Proof. That $r_a = \max_{z\in(0,1)} \lambda_h(z) = \max_{z\in(0,1)} \lambda_l(z) = \lambda_l^*(0)$ follows immediately from the definitions. Now, note that $E_h[\ell_1] = \lambda_h'(0)$. Thus $r_a < \mathbb{E}_h[\ell]$ is a simple consequence of the fact that $r_a = \lambda_h^\star(0) = \max_{z\geq 0}\lambda(z)$, that this maximum is obtained in $(0,1)$, and that λ_h is strictly concave. It follows from the same considerations that $r_a < -\mathbb{E}_l[\ell]$. Finally, by Proposition 21, $\lambda_h^\star(r_a) > 0$ as $\lambda_h'(0) < r_a < \lambda_h'(1)$. The same arguments show that $r_a < -\mathbb{E}_l[\ell_1]$ and $\lambda_l^\star(r_a) > 0$. □

Proof of Theorem 11. Consider the case $\Theta = h$. As shown in Lemma 8 the probability that the agent makes a mistake is equal to the probability that the LLR is below $\underline{L}(\alpha^h)$. Thus, Lemma 10 allows us to characterize this probability explicitly:

$$\mathbb{P}_h\left[a_t^i \neq \alpha^\theta\right] = \mathbb{P}_h\left[R_t^i \leq \underline{L}(\alpha^h)\right] = \mathbb{P}_h\left[R_t^i \leq o(t)\right] = \mathrm{e}^{-\lambda_h^\star(0)\cdot t + o(t)}.$$

An analogous argument yields that $\mathbb{P}_l\left[a_t^i \neq \alpha^\theta\right] = \mathrm{e}^{-\lambda_l^\star(0)\cdot t+o(t)}$. By (4) $\lambda_h^\star(0) = \lambda_l^\star(0)$. □

Appendix C. Observing the Last Action Unidirectionally

In this section we prove Proposition 13. Assuming that agent 1 only observes the last action of agent 2, we would like to calculate $\mathbb{P}_h\left[a_1^t \neq \alpha^h\right]$. We can write this as

$$\mathbb{P}_h\left[a_t^1 \neq \alpha^h\right] = \mathbb{P}_h\left[a_t^1 \neq \alpha^h, a_{t-1}^2 = \alpha^h\right] + \mathbb{P}_h\left[a_t^1 \neq \alpha^h, a_{t-1}^2 = \alpha^l\right] + \mathbb{P}_h\left[a_t^1 \neq \alpha^h, a_{t-1}^2 \notin \left\{\alpha^h, \alpha^l\right\}\right].$$

We already calculated the first term: it is equal to $\mathrm{e}^{-\lambda_h^\star(-r_a)\cdot t+o(t)}$. To calculate the second term we write

$$\mathbb{P}_h\left[a_t^1 \neq \alpha^h \text{ and } a_{t-1}^2 = \alpha^l\right] = \mathbb{P}_h\left[a_t^1 \neq \alpha^h \mid a_{t-1}^2 = \alpha^l\right] \times \mathbb{P}_h\left[a_{t-1}^2 = \alpha^l\right] = \mathrm{e}^{-\lambda_h^\star(+r_a)\cdot t+o(t)} \times \mathbb{P}_h\left[a_{t-1}^2 = \alpha^l\right],$$

where the second equality is an application of (8). To estimate $\mathbb{P}_h\left[a^2_{t-1} = \alpha^l\right]$ we note that the event $a^2_{t-1} = \alpha^l$ is equal to the event $P^2_{t-1} \leq \underline{L}(\alpha^l)$. Therefore, by Lemma 10, $\mathbb{P}_h\left[a^2_{t-1} = \alpha^l\right] = \mathrm{e}^{-\lambda^\star(0)\cdot t+o(t)} = \mathrm{e}^{-r_a\cdot t+o(t)}$. Hence

$$\mathbb{P}_h\left[a^1_t \neq \alpha^h \text{ and } a^2_{t-1} = \alpha^l\right] = \mathrm{e}^{-(\lambda^\star_h(r_a)+r_a)\cdot t+o(t)}.$$

We are thus left with the estimation of the last addend, $\mathbb{P}_h\left[a^1_t \neq \alpha^h \text{ and } a^2_{t-1} \notin \left\{\alpha^h, \alpha^l\right\}\right]$. To this end we note that

$$\mathbb{P}_h\left[a^2_{t-1} \notin \left\{\alpha^h, \alpha^l\right\}\right] \leq \mathbb{P}\left[R^2_t \leq \underline{L}(\alpha^h)\right] = \mathrm{e}^{-r_a\cdot t+o(t)},$$

where the last equality is another consequence of Lemma 10. Therefore, by (8),

$$\mathbb{P}_h\left[a^1_t \neq \alpha^h \text{ and } a^2_{t-1} \notin \left\{\alpha^h, \alpha^l\right\}\right] = \mathrm{e}^{-2r_a t+o(t)}.$$

We thus have that

$$\mathbb{P}_h\left[a^1_t \neq \alpha^h\right] = \mathrm{e}^{-\lambda^\star_h(-r_a)\cdot t+o(t)} + \mathrm{e}^{-(\lambda^\star_h(r_a)+r_a)\cdot t+o(t)} + \mathrm{e}^{-2r_a t+o(t)}.$$

Recall that $\lambda^\star_l(\eta) = \lambda^\star_h(-\eta) - \eta$ (by (17)) and so $\lambda^\star_h(r_a) + r_a = \lambda^\star_l(-r_a)$. Hence

$$\mathbb{P}_h\left[a^1_t \neq \alpha^h\right] = \mathrm{e}^{-\lambda^\star_h(-r_a)\cdot t+o(t)} + \mathrm{e}^{-\lambda^\star_l(-r_a)\cdot t+o(t)} + \mathrm{e}^{-2r_a t+o(t)}.$$

We show in Claim 25 below that $\lambda^\star_h(-r_a) < 2r_a$, and likewise $\lambda^\star_l(-r_a) < 2r_a$. Given this, the last addend can be absorbed into the $o(t)$ term, and we have that

$$\mathbb{P}_h\left[a^1_t \neq \alpha^h\right] = \mathrm{e}^{-r_u\cdot t+o(t)},$$

where

$$r_u = \min\left\{\lambda^\star_h(-r_a), \lambda^*_l(-r_a)\right\} = r_a + \min\left\{\lambda^\star_l(r_a), \lambda^\star_h(r_a)\right\}.$$

By symmetry the same holds conditioned on $\Theta = l$, and so we have shown that

$$\mathbb{P}\left[a^1_t \neq \alpha^\theta\right] = \mathrm{e}^{-r_u\cdot t+o(t)}.$$

Claim 25. $\lambda^\star_h(-r_a) < 2r_a$ and $\lambda^\star_l(-r_a) < 2r_a$.

Proof. We show the former; the proof of the latter is identical. To this end, we first note that $-r_a > \lambda'_h(1)$ (Claim 24). It thus follows that the maximum in

$$\lambda^\star_h(-r_a) = \max_{z\geq 0} \lambda_h(z) + r_a z$$

is also obtained in $(0, 1)$, since the z in which it is obtained is the solution to $\lambda'(z) = -r_a$. Thus

$$\lambda^\star_h(-r_a) = \max_{z\in(0,1)} \lambda_h(z) + r_a z < \max_{z\in(0,1)} \lambda_h(z) + r_a = 2r_a.$$ ☐

This also concludes the proof of Proposition 13.

Appendix D. Observing all Actions Unidirectionally

In this section we show that when agent 1 observes all of agent 2's actions, then the speed of learning is identical to the speed in the case that she observes only the last action:

$$\mathbb{P}\left[a_t^1 \neq \alpha^\theta\right] = \mathrm{e}^{-r_u \cdot t + o(t)}.$$

One direction is immediate: observing all actions can only reduce the probability of error relative to observing the last action, and so we know that

$$\mathbb{P}\left[a_t^1 \neq \alpha^\theta\right] \leq \mathrm{e}^{-r_u \cdot t + o(t)}.$$

It thus remains to be shown that

$$\mathbb{P}\left[a_t^1 \neq \alpha^\theta\right] \geq \mathrm{e}^{-r_u \cdot t + o(t)}.$$

To show this we show that the probability of a smaller event already satisfies this inequality. Specifically, we condition (without loss of generality) on $\Theta = h$ and would like to consider the case that agent 2 chooses the wrong action α^l at all time periods up to time t. A technical difficultly arises from the fact that for some priors and signal distributions it may be that agent 2 never chooses α^l up to some period s, since the prior is too high and each signal is too weak. But of course there is some s such that $\mathbb{P}\left[a_t^2 = \alpha^l\right] > 0$ for all $t \geq s$. Accordingly, we define for each t the action $\alpha_\tau^{\min}$ to lowest (i.e., having the lowest $\overline{L}$) that is taken by agent 2 with positive probability at time t. By the above, $\alpha_\tau^{\min}$ is equal to α^l for all t large enough. We then prove the claim by showing that

$$\mathbb{P}_h\left[a_t^1 \neq \alpha^h, \cap_{1 \leq \tau \leq t}\{a_\tau^2 = \alpha_\tau^{\min}\}\right] = \mathrm{e}^{-r_u \cdot t + o(t)}. \tag{18}$$

That is, we show that even when agent 1 observes agent 2 take the wrong action at every period in which this is possible - even then agent 1 gets it wrong with probability that is comparable to the probability of mistake when observing only the last action. Denote by E_t the event

$$E_t = \cap_{1 \leq \tau \leq t}\left\{a_\tau^2 = \alpha_\tau^{\min}\right\}.$$

We first claim that

$$\mathbb{P}_h\left[E_t\right] = \mathrm{e}^{-r_a \cdot t + o(t)} \tag{19}$$

and that

$$\mathbb{P}_l\left[E_t\right] = e^{-o(t)}, \tag{20}$$

so that asymptotically this event has the same rate as the event $a_t^2 = \alpha^l$, for both possible values of Θ. Given this, the analysis is identical to the one carried out in Sec. 4.3, and likewise yields (18). It thus remains to calculate the conditional rates of E_t, and in particular to show that they are the same at the rates of the event $a_t^2 = \alpha^l$. The key insight from which this follows is the classical Ballot Theorem (Bertrand (1887)). It states that if $(X_1, X_2, \ldots)$ are i.i.d. random variables, and if $Y_t = \sum_{\tau=1}^t X_\tau$ then

$$\frac{1}{t}\mathbb{P}\left[Y_t \leq 0\right] \leq \mathbb{P}\left[\cap_{\tau=1}^t \left\{Y_\tau \leq\right\} 0\right] \leq \mathbb{P}\left[Y_t \leq 0\right],$$

and so in particular the event that $Y_t \leq 0$ has the same rate as the event that $Y_\tau \leq 0$ for all $\tau \leq t$. Instead of using the Ballot Theorem, we use our Theorem 23.

Indeed, noting that the event E_t can be written as

$$E_t = \cap_{1 \leq \tau < t} \left\{R_\tau^2 \leq \overline{L}(\alpha_\tau^{\min})\right\}.$$

Thus, if we define $X_t = \ell_t$ and $y_t = \overline{L}(\alpha_\tau^{\min}) - L_0$ then $\lim_t y_t/t = 0$ and Theorem 23 yields the desired rates. This completes the proof that when agent 1 observes all the actions of agent 2 then

$$\mathbb{P}\left[a_t^2 \neq \alpha^\theta\right] = \mathrm{e}^{-r_u \cdot t + o(t)}.$$

Appendix E. Bidirectional Observations

As in Section (4.3), we define for each t the action $\alpha_t^{\min}$ to be the lowest action (i.e., having the lowest $\overline{L}(\alpha)$) that is taken by any agent with positive probability at time t, and observe that $\alpha_t^{\min}$ is equal to α^l for all t large enough. We define

$$G_t = \cap_{i=1}^n \cap_{\tau=1}^t \left\{a_\tau^i = \alpha_\tau^{\min}\right\}.$$

Proof of Lemma 15. Note first, that each agent chooses action $\alpha_1^{\min}$ in the first period if the likelihood ratio she infers from her first private signal is at most $\overline{L}(\alpha_1^{\min})$. Hence

$$G_1 = \bigcap_{1 \leq i \leq n} \{a_1^i = \alpha_1^{\min}\} = \bigcap_{1 \leq i \leq n} \{R_1^i \leq \overline{L}(\alpha_1^{\min})\}.$$

Thus G_1 is an intersection of conditionally independent events. Assume now that all agents choose the action $\alpha_\tau^{\min}$ up to period $t-1$; that is, that G_{t-1} has occurred, which is a necessary condition for G_t. What would cause any one of them to again choose $\alpha_t^{\min}$ at period t? It is easy to see that there will be some threshold q_t^i such that, given G_{t-1}, agent i will choose $\alpha_t^{\min}$ if and only if her private likelihood ratio P_t^i is lower than q_t^i. By the symmetry of the equilibrium, q_t^i is independent of i, and so we will simply write it as q_t. It follows that

$$G_t = G_{t-1} \cap \bigcap_{1 \leq i \leq n} \{R_t^i \leq q_t\}.$$

Therefore, by induction, and if we denote $q_1 = \overline{L}(\alpha_\tau^{\min})$, we have that

$$G_t = \bigcap_{\substack{\tau \le t \\ 1 \le i \le n}} \{R_\tau^i < q_\tau\}.$$

Now, note that the event that agent i chooses $\alpha_\tau^{\min}$ in all periods is not independent of the event that some other agent j does the same. Still, by rearranging the above equation we can write G_t as an intersection of conditionally independent events:

$$G_t = \bigcap_{1 \le i \le n} \left(\bigcap_{1 \le \tau \le t} \{R_\tau^i \le q_\tau\} \right),$$

and if we denote

$$W_t^i = \bigcap_{1 \le \tau \le t} \{R_\tau^i \le q_\tau\},$$

then the W_t^i's are conditionally independent, and

$$G_t = \bigcap_{1 \le i \le n} W_t^i.$$

□

Proof of Proposition 16. Agent 1's log-likelihood ratio conditional on $\cap_{i=1}^n W_{t-1}^i$ at time t equals

$$L_t^1 = P_t^1 + \log \frac{\mathbb{P}_h\left[\cap_{i=1}^n W_{t-1}^i\right]}{\mathbb{P}_l\left[\cap_{i=1}^n W_{t-1}^i\right]}.$$

Since the W_{t-1}^i's are conditionally independent, we have that

$$L_t^1 = P_t^1 + \sum_{i=1}^n \log \frac{\mathbb{P}_h\left[W_{t-1}^i\right]}{\mathbb{P}_l\left[W_{t-1}^i\right]}.$$

Finally, by symmetry, all the numbers in the sum are equal, and

$$L_t^1 = P_t^1 + (n-1) \cdot \log \frac{\mathbb{P}_h\left[W_{t-1}^1\right]}{\mathbb{P}_l\left[W_{t-1}^1\right]}.$$

Now, the last addend is just a number. Therefore, if we denote

$$q_t = \overline{L}(\alpha^l) - (n-1) \cdot \log \frac{\mathbb{P}_h\left[W_{t-1}^1\right]}{\mathbb{P}_l\left[W_{t-1}^1\right]}, \tag{21}$$

then

$$L_n^1 = R_t^1 - q_t + \overline{L}(\alpha^l),$$

and $L_t^1 \le \overline{L}(\alpha^l)$ (and thus $a_t^1 = \alpha^l$) whenever $P_t^1 \le q_t$.

□

We recall the recursive relation between q_t and W_t^i:

$$q_t = \overline{L}(\alpha_t^{\min}) - (n-1) \cdot \log \frac{\mathbb{P}_h\left[W_{t-1}^1\right]}{\mathbb{P}_l\left[W_{t-1}^1\right]} \quad \text{and} \quad W_t^i = \bigcap_{1 \le \tau \le t} \{R_\tau^i \le q_t\}.$$

Claim 26. $q_t \ge \overline{L}(\alpha_t^{\min})$ for all t.

Proof. Let F_h and F_l be the cumulative distribution functions of a private log-likelihood ratio ℓ, conditioned on $\Theta = h$ and $\Theta = l$, respectively. Then it is easy to see that F_h stochastically dominates F_l, in the sense that $F_l(x) \ge F_h(x)$ for all $x \in \mathbb{R}$. It follows that the joint distribution of $\{R_\tau^i\}_{\tau \le t}$ conditioned on $\Theta = h$ dominates the same distribution conditioned on $\Theta = l$, and so $\mathbb{P}_h[W_t^1] \le \mathbb{P}_l[W_t^1]$. Hence $q_t \ge \overline{L}(\alpha_t^{\min})$. □

Claim 27. There is a constant $C > 0$ such that $\mathbb{P}_l[W_t^1] \ge C$ for all t.

Proof. Since the events W_t^1 are decreasing, we will prove the claim by showing that

$$\lim_{t\to\infty} \mathbb{P}_l\left[W_t^1\right] > 0,$$

which by definition is equivalent to

$$\lim_{t\to\infty} \mathbb{P}_l\left[\cap_{\tau \le t}\left\{R_\tau^i \le q_\tau\right\}\right] > 0.$$

Since $q_t \ge \overline{L}(\alpha_t^{\min})$, it suffices to prove that

$$\lim_{t\to\infty} \mathbb{P}_l\left[\cap_{\tau \le t}\left\{R_\tau^i \le \overline{L}(\alpha_\tau^{\min})\right\}\right] > 0.$$

To prove the above, note that agents eventually learn Θ, since the private signals are informative. Therefore, conditioned on $\Theta = l$, the limit of R_t^i as t tends to infinity must be $-\infty$. Thus, with probability 1, for all t large enough it does hold that $R_t^1 \le \overline{L}(\alpha_\tau^{\min})$. Since each of the events W_t^1 has positive probability, and by the Markov property of the random walk R_t^1, it follows that the event $\cap_\tau \left\{R_\tau^i \le \overline{L}(\alpha_\tau^{\min})\right\}$ has positive probability. Finally, by monotonicity

$$\lim_{t\to\infty} \mathbb{P}_l\left[W_t^1\right] > \mathbb{P}_l\left[\cap_\tau\left\{R_\tau^i \le \overline{L}(\alpha_\tau^{\min})\right\}\right] > 0.$$

□

It follows immediately from this claim and the definition of q_t that

$$\lim_{t\to\infty} \frac{q_t}{t} = -(n-1)\lim_t \frac{1}{t}\log \mathbb{P}_h\left[W_{t-1}^1\right]_t, \tag{22}$$

provided that the limit exists.

Let $\underline{q} = \liminf_{t\to\infty} q_t/t$. Since $W_t^i = \cap_{\tau=1}^t\left\{R_\tau^i \le q_\tau\right\}$, it follows from Theorem 23 that

$$-\lim{}_{t\to\infty}\frac{1}{t}\log \mathbb{P}_h\left[W_t^i\right] = \lambda_h^\star(\underline{q}),$$

provided that $\underline{q} > \inf_z \lambda_h'(z)$. But $\underline{q} \geq 0$ (Claim 26), and so this indeed holds. Thus, by (22), we have proved the following claim:

Proposition 28. *The limit* $q = \lim_{t\to\infty} \frac{q_t}{t}$ *exists, and*

$$q = (N-1)\lambda_h^\star(q)\,.$$

Proof of Claim 17. Recall that $\lambda_h^\star$ is strictly convex, and that $\lambda_h^\star(D) = 0$, where we denote $D = \mathbb{E}_h[\ell]$. Hence

$$\begin{aligned}\lambda_h^\star(q) &< \frac{q}{D}\lambda_h^\star(D) + \frac{D-q}{D}\lambda_h^\star(0) \\ &= \frac{D-q}{D}\lambda_h^\star(0).\end{aligned}$$

Substituting $(n-1)\lambda_h^\star(q)$ for q and simplifying yields

$$\lambda_h^\star(q) < \frac{D}{D/\lambda_h^\star(0) + n - 1}.$$

Since $\lambda_h^\star(0) < D$ (Claim 24) we have shown that

$$n\lambda_h^\star(q) < D,$$

and so

$$\frac{n}{n-1}q = n\lambda_h^\star(q) < D\,.$$ ☐

Proof of Claim 18. Since $r_u = \min\{\lambda_l^\star(-r_a), \lambda_h^\star(-r_a)\}$, in order to prove the claim we need to show that $2\cdot q < \lambda_h^\star(-r_a)$; the corresponding condition for the low state will follow by the same argument.

We consider two cases. If $q \geq r_a$ then $\lambda_h^\star(q) \geq \lambda_h^\star(0)$, since $q = \lambda_h^\star(q)$ and $r_a = \lambda_h^\star(0)$. By the monotonicity of $\lambda_h^\star$ (Proposition 21) it then follows that $q \leq 0$. But this is false (e.g., since it implies that $q = \lambda_h^\star(q) \geq \lambda_h^\star(0) > 0$), and so we have reached a contradiction.

Hence $q < r_a$, in which case $\lambda_h^\star(-q) < \lambda_h^\star(-r_a)$, since $\lambda_h^\star$ is strictly decreasing (Proposition 21). Now, since $q = \lambda_h^\star(q)$, we have that $2\cdot q = q + \lambda_h^\star(q) = \lambda_l^*(-q)$, where the last equality follows from the general fact (see Appendix B) that $\lambda_l^\star(\eta) = \lambda_h^\star(-\eta) - \eta$. Thus $2\cdot q < \lambda_h^\star(-r_a)$. ☐

We now turn to showing that conditioned on groupthink - that is, conditioned on the event G_t - all agents have, with high probability, a private LLR R_t^i that indicates the correct action. In fact, the LLR is arbitrarily close to $t\cdot q$, the asymptotic threshold for R_t^i above which groupthink ends.

Theorem 29. *For every* $\epsilon > 0$ *it holds that*

$$\lim_{t\to\infty} \mathbb{P}_h\left[R_t^i > t\cdot(q-\epsilon) \text{ for all } i \mid G_t\right] = 1,$$

where, as above, q is the solution to $q = (n-1)\lambda_h^\star(q)$.

Proof. By Theorem 22 we know that

$$\lim_{t\to\infty} -\frac{1}{t}\log \mathbb{P}_h\left[R_t^i \le t\cdot(q-\epsilon)\right] = \lambda_h^\star(q-\epsilon).$$

Since $\lambda_h^\star(q-\epsilon) > \lambda_h^\star(q)$ it follows that

$$\lim_{t\to\infty} -\frac{1}{t}\log \mathbb{P}_h\left[A_t\right] = n\cdot\lambda_h^\star(q-\epsilon) > n\cdot\lambda_h^\star(q),$$

where A_t is the event $\{R_t^i \le t\cdot(q-\epsilon) \text{ for all } i\}$. Since for t high enough the event A_t is included in G_t, and since

$$\lim_{t\to\infty} -\frac{1}{t}\log \mathbb{P}_h\left[G_t\right] = n\cdot\lambda_h^\star(q),$$

it follows that $\mathbb{P}_h\left[A_t \mid G_t\right]$ decays exponentially with t. Hence $\mathbb{P}_h\left[A_t^c \mid G_t\right] \to_t 1$, which is the claim we set to prove. □

We now turn to the proof of Theorem 7. We assume that each agent i observes a Normal signal $s_t^i \sim \mathcal{N}(m_\theta, n)$ with mean

$$m_\Theta = \begin{cases} +1 & \text{if } \Theta = h \\ -1 & \text{if } \Theta = l \end{cases}$$

and variance n.[18] Note, that for any number of agents the precision of the joined signal equals 1, and thus the total information the group receives every period is fixed, independent of n.

We assume that the prior belief assigns probability one-half to each state $p_0 = 1/2$ and that there are two actions $A = \{l, h\}$ and each agent just wants to match the state, as in the "matching the state" example (Section 2.5.1). As in the first period each agent bases her decision only on her own private signal, she takes the action h whenever her signal s_1^i is greater than 0 and the action l otherwise:

$$a_1^i = \begin{cases} h & s_1^i > 0 \\ l & s_1^i \le 0 \end{cases}.$$

[18] All results generalize to non-symmetric means, since only the difference $|m_h - m_l|$ enters the Bayesian calculations.

The private likelihood of each agent after observing the first t signals is given by

$$R_t^i = \log \frac{\prod_{\tau=1}^t \exp\left(-\frac{\left(s_\tau^i - 1\right)^2}{2\,n}\right)}{\prod_{\tau=1}^t \exp\left(-\frac{\left(s_\tau^i + 1\right)^2}{2\,n}\right)}$$
$$= \frac{2}{n} \sum_{\tau=1}^t s_\tau^i \,.$$

The probability that an agent takes the correct action Θ in period 1 (conditional only on her own first period signal) is thus given by

$$\mathbb{P}_h\left[\Theta = a_1^i\right] = \mathbb{P}_h\left[s_1^i \geq 0\right]$$
$$= 1 - \Phi\left(\frac{-m_h}{\sqrt{n}}\right)$$
$$= \Phi\left(\frac{1}{\sqrt{n}}\right) \,.$$

By symmetry, $\mathbb{P}_l[a_1^i = \Theta] = \Phi(1/\sqrt{n})$ as well. Denote $\pi_n = \Phi\left(\frac{1}{\sqrt{n}}\right)$ and by $w_1 = |\{i \in n \colon a_1^i = h\}|$ the number of agents taking the action $a_1^i = h$. Let $\kappa_n = \log(\pi_n/(1-\pi_n))$, and note that $2/\sqrt{n} \geq \kappa_n \geq 1/\sqrt{n}$.

As the action of each agent is independent, the LLR of agent i at the beginning of period 2 is given by

$$L_2^i = \frac{2}{n} \sum_{\tau=1}^2 s_\tau^i - (2\,w_1 - n)\,\kappa_n - \operatorname{sgn}(s_1^i)\,\kappa_n \,.$$

We define the private part of the LLR at the beginning of period 2 as

$$\hat{R}_2^i = \frac{2}{n} \sum_{\tau=1}^2 s_\tau^i - \operatorname{sgn}(s_1^i)\,\kappa_n$$

and the public part of the LLR as

$$L_2^p = (2\,w_1 - n)\,\kappa_n \,.$$

Let α_m be the action that the majority of the agents chose in the first period (with $\alpha_m = l$ in case of a tie). Note that $\alpha_m = h$ iff $L_2^p > 0$. Let E_t be the event that all agents take the first period majority action α_m in all subsequent periods up to time t, i.e., $a_s^i = \alpha_m$ for all $1 < s \leq t$.

Proposition 30. *The probability of E_t goes to one as the number of agents goes to infinity, i.e.,*

$$\lim_{n\to\infty} \mathbb{P}\left[E_t\right] = 1 \,.$$

This is a rephrasing of Theorem 7. We in fact provide a finitary statement and prove that $\mathbb{P}[E_t] \geq 1 - 20 \cdot t \cdot \sqrt{\frac{\log n}{n}}$.

We first show that the the probability of the event E_2 that all agents take the same action in period 2 goes to one. The LLR of agent i at the beginning of period 2 is given by

$$\begin{aligned} L_2^i &= \frac{2}{N} \sum_{\tau=1}^{2} s_\tau^i + (2\, w_1 - n)\, \kappa_n - \mathrm{sgn}(s_1^i)\, \kappa_n\, . \\ &= \hat{R}_2^i + L_2^p\, . \end{aligned}$$

To show that E_2 has high probability we show that with high probability it holds that L_2^p, the public belief induced by the first period actions, is large (in absolute value) and that the private beliefs are all small. Intuitively, this holds since both are (approximately) zero mean Gaussian, with L_2^p having constant variance and $\hat{R}_2^i$ having variance of order $1/\sqrt{n}$. It will then follow that with high probability the signs of L_2^p and L_2^i are equal for all i, which is a rephrasing of the definition of E_2.

Let A be the event that all of the private signals in the first t periods have absolute values at most $M = 4\sqrt{n \log n}$. Using the union bound (over the agents and time periods), this happens except with probability at most

$$\mathbb{P}[A^c] \leq t \cdot n \cdot \mathbb{P}\left[|s_t^i| > M\right] \leq t \cdot n \cdot 2 \cdot \Phi\left(-\frac{1}{2} M/\sqrt{n}\right);$$

the 1/2 factor in the argument of Φ is taken to account for the fact that the private signals do not have zero mean. Since $\Phi(-x) < e^{-\frac{x^2}{2}}$ for all $x < -1$, we have that

$$\mathbb{P}[A^c] \leq \frac{2 \cdot t}{n}.$$

Let

$$\hat{R}_\tau^i = \frac{2}{n} \sum_{\tau'=1}^{\tau} s_{\tau'}^i - \mathrm{sgn}(s_1^i)\, \kappa_n.$$

Thus the event A implies that

$$|\hat{R}_\tau^i| \leq \frac{2}{n} \cdot t \cdot M + \kappa_n \leq 8 \cdot t \cdot \sqrt{\frac{\log n}{n}} + \frac{2}{\sqrt{n}} \leq 9 \cdot t \cdot \sqrt{\frac{\log n}{n}}.$$

Let B be the event that the absolute value of the public LLR L_2^p is at least $9 \cdot t \cdot \sqrt{\frac{\log n}{n}}$; this is chosen so that the intersection of A and B implies E_2. Conditioned on $\Theta = h$, the random variable w_1 has the unimodal binomial distribution $B(n, \pi_n)$, which has mode $\lfloor (n+1) \cdot \pi_n \rfloor$. The probability at this mode is easily shown to be at most $1/\sqrt{n}$. The same applies conditioned on $\Theta = l$. It follows that the probability of B^c, which by definition is equal to the probability that $|w_1 - n/2| \leq \frac{1}{\kappa_n} 9 \cdot t \cdot \sqrt{\frac{\log n}{n}}$, is at most $\frac{2}{\kappa_n} 9 \cdot t \cdot \sqrt{\frac{\log n}{n}}$ times the probability of the mode, or

$$\mathbb{P}\left[B^c\right] \leq \frac{2}{\kappa_n} 9 \cdot t \cdot \sqrt{\frac{\log n}{n}} \cdot \frac{1}{\sqrt{n}} \leq 18 \cdot t \cdot \sqrt{\frac{\log n}{n}}.$$

Together with the bound on the probability of A, we have that

$$\mathbb{P}\left[\text{A and } B\right] \geq 1 - 20 \cdot t \cdot \sqrt{\frac{\log n}{n}},$$

and in particular

$$\mathbb{P}\left[E_2\right] \geq 1 - 20 \cdot \sqrt{\frac{\log n}{n}}.$$

We now claim that $A \cap B$ implies E_t. To see this, note that as $A \cap B$ implies E_2, the agents all observe at period 2 that no other agent has a strong enough signal to dissent with the first period majority. This only strengthens their belief in the first period majority, requiring them an even higher (in absolute value) threshold than L_2^p to choose another action; the formal proof of this statement is identical to the proof of Claim 26. But since, under the event $A \cap B$, each of their private LLRs $\hat{R}_\tau^i$ is weaker than L_2^p for all $\tau \leq$t, they will not do so at period 3, or, by induction, in any of the periods prior to period t. This completes the proof of 30, and thus of Theorem 7.