RATIONAL GROUPTHINK

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

Abstract. We study how long-lived rational agents learn from repeatedly observing each others' actions. We find that in the long run, information aggregation fails, and the fraction of private information transmitted goes to zero as the number of agents gets large. With Normal signals, in the long-run, agents learn less from observing the actions of any number of other agents than they learn from seeing *three* other agents' signals. We identify *rational groupthink*—in which agents ignore their private signals and choose the same action for long periods of time—as the cause of this failure of information aggregation.

1. Introduction

Recently, there has been a renewed interest in understanding social learning, i.e. the ability of agents to learn by observing each others' actions. The key question in this literature is how well information is aggregated. As the analysis of the beliefs of long-lived Bayesian agents is challenging (e.g., Cripps, Ely, Mailath, and Samuelson, 2008), most of this literature focuses either on short-lived agents (e.g., Dasaratha et al., 2018; Mueller-Frank and Arieli, 2018) or on non-rational belief dynamics such as the DeGroot model (see Golub and Jackson, 2010; Jadbabaie et al., 2013) or quasi-Bayesian agents (see Molavi, Tahbaz-Salehi, and Jadbabaie, 2015).

By using new techniques from large deviation theory we are able to overcome the difficulty associated with the analysis of Bayesian beliefs and manage to analyze social learning with long-lived rational agents. Our main result is that social learning will fail for Bayesian agents in a large society: An arbitrarily large group of Bayesian agents observing each others' actions will only learn as fast as a small group of agent observing each others' signals directly. For example, when signals are normal, 4 agents sharing their signals learn faster

[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, Wade Hann-Caruthers, Benjamin Golub, Rainie Heck, Paul Heidhues, Shachar Kariv, Navin Kartik, Steven Morris, Luciano Pomatto, Larry Samuelson, Lones Smith, Juuso Toikka, 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.

than a group of n agents who observe each others' actions (but not signals). This failure of information aggregation is caused by the endogenous correlation in the agents' actions. As it is well known a large number of correlated signals might convey less information than a small number of independent signals.[1]

Whereas signals are independent, the agents' actions become endogenously correlated. This correlation is an immediate consequence of the incentive to learn from each others' actions. For example, if agent 1 takes an action that is optimal in some state of the world the other agents will infer that agent 1's private belief indicates that this state is relatively likely and will themselves take this action with greater probability. A greater number of agents increases this correlation as agents share more common information. The (perhaps surprising) insight of our analysis is that as the number of agents grows, the correlation increases to an extent that completely out-weighs the gain of the additional independent private signals. We show that asymptotically this failure of information aggregation holds for any signal structure, any utility function and any number of agents.

What inference an agent draws from the actions of another agent depends on her belief about the other agents' beliefs. Thus, agents' actions may depend on their higher order beliefs. This poses a significant challenge for the exact characterization of behavior. We circumvent this problem by focusing on long-term, asymptotic probabilities, and by analyzing a phenomenon that we call "rational groupthink". We loosely define rational groupthink to be the event that *all* agents take the wrong action for many periods, despite all having private signals that indicate otherwise. Importantly, this behavior arises in our model as a consequence of Bayesian updating, and is not driven by an assumed desire for conformity. Through a recursive argument we are able to estimate the asymptotic probability of rational groupthink (see Subsection 4.3) and find that due to rational groupthink agents in a large group learn almost as slowly as they do in autarky. Hence, in this sense, rational groupthink prevents almost all information aggregation.[2]

Rational groupthink occurs after a consensus on an action is formed in the initial periods, making it optimal for every agent to continue taking the consensus action, even when her private information indicates otherwise. Indeed, we show that typically, after a wrong consensus forms, all agents quickly observe private signals which provide strong evidence for choosing the correct action, and yet a long time may pass until any of them breaks the wrong consensus (Proposition 1). Thus a situation arises in which each agent's private information

[1]This point has been made for example by Clemen and Winkler (1985).

[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 places 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."

indicates the correct action, and yet, because of the group dynamics, all agents choose the wrong action. We thus find the name rational groupthink to be an apt description.

We study the effect of increasing the group size. On the one hand, 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 rational groupthink becomes more or less likely. We show that the inefficiency (measured as the share of information that is lost) associated with the rational groupthink effect becomes arbitrarily large as the size of the group increases. Our first main result shows that, even as the number of agents goes to infinity, the speed of learning from actions stays bounded by a constant (Theorem 1), whereas the speed of learning from the aggregated signals, which is proportional to the number of agents, goes to infinity (Fact 2). Thus, in a large group, almost no information is aggregated; the agents' belief when observing only actions has the same precision as would result from observing a vanishingly small fraction of the available private signals. Specifically, 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 number of agents! Hence, at most a fraction of $4/n$ of the private information is transmitted through actions (Corollary 1). We proceed beyond Normal signals to show that for any signal distribution at most a fraction of c/n of the private information is transmitted through actions, for some constant c that depends *only* on the distribution of the private signals (Proposition 2).

As a robustness test, we complement our results on the asymptotic rate by an analysis of the exact probability with which the wrong action is chosen in a given period for Normal signals. We study 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. Our second main result shows that in this setting, our asymptotic finding—that for large groups almost no information is aggregated through actions as a consequence of rational groupthink—holds starting already from the *second* period. We show that in *every period* the probability with which an agent chooses the correct action when she observes others converges (as the number of agents goes to infinity) to the probability with which she would choose the correct action if she could only observe the actions taken by others in the first period (Theorem 2). Thus, information fails to aggregate not only asymptotically, but already after the first period.

An important advantage of asymptotic rates is that they are tractable. Beyond this, we show that asymptotic rates have the advantage of being independent of many details of the model, providing a measure that is robust to changes in model parameters such as the agents' prior or the exact utility function. For similar reasons of tractability and robustness, many previous works have studied asymptotic (long run) rates of learning in various settings.[3]

Most of the preceding literature studies situations where each agent observes a *single* signal and agents try to infer the others' signals from repeatedly observing their actions. Geanakoplos and Polemarchakis (1982); Sebenius and Geanakoplos (1983); Parikh and Krasucki (1990); Mossel et al. (2015) give conditions under which agents actions agree in the long-run. Rosenberg, Solan, and Vieille (2009) also study agreement, in a more general social learning setting. The question of how well information is aggregated in such settings was considered in an important paper by Vives (1993), who studies rate at which information is aggregated through noisy prices.

In contrast to this literature we allow for agents to repeatedly observe signals about the state of the world. The only other articles which we are aware of that tackle this problem are Jadbabaie et al. (2013) and Molavi et al. (2015). Both study asymptotic rates of learning under (non-rational) linear belief updating rules in complex observational networks. The focus of both papers differs from ours: Jadbabaie et al. (2013) and Molavi et al. (2015) allow for complex network structures, but impose simple linear belief updating rules. In contrast, we study the complexities associated with Bayesian learning, but assume that all actions are commonly known. Interestingly, our results contrast the findings of Jadbabaie et al. (2013) and Molavi et al. (2015); while in their model information is efficiently aggregated, in our model it is not. This is a consequence of the difference in the rationality assumptions.

Gale and Kariv (2003) use numerical methods to characterize the asymptotic rates with which rational agents learn, and emphasize the importance of understanding the rates at which Bayesian agents learn from each other.[4]

Our work is also related to models of rational herding (Bikhchandani, Hirshleifer, and Welch, 1992; Banerjee, 1992), 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 crucial difference is that in herding models each agent acts only once, whereas in our model agents take actions repeatedly. We thus show that the failure of information aggregation is not particular to sequential models in which agents act only once, but more generally extends to situations of repeated interactions. Our main finding, the rational groupthink effect, has no analogue in sequential herding models, since, in these models, once a herd starts, it is not true that every agent's

[3] Examples of papers studying the rate of learning are Vives (1993); Chamley (2004); Duffie and Manso (2007); Duffie, Malamud, and Manso (2009); Duffie, Giroux, and Manso (2010). Asymptotic rates also have been studied in other settings in which it is difficult to analyze the short-term dynamics (e.g., Hong and Shum, 2004; Hörner and Takahashi, 2016). Jadbabaie et al. (2013) and Molavi et al. (2015) study the rate of learning in an almost identical setting, with boundedly rational agents.

[4] 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 [...] This is an important subject for future research."

private signal indicates the correct action. In our model we show that this happens, in the long run, with high probability.

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)).

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) $s_t^i \in \mathbb{R}$, takes an action $a_t^i \in A$, and finally observes the actions taken by others this period. 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_t^i are i.i.d, across agents i and over time t, 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.[7] Our main example is that of *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 $\mathbb{P}[s_t^i = \Theta \mid \Theta] = \phi > 1/2$.

[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]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.

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 $u(s_{t+1}^i, a_t^i)$.[8] 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 payoff conditional on the state θ[9]

$$\begin{aligned}\bar{u}(h,\alpha) &:= \mathbb{E}_h\left[u(s_{t+1}^i,\alpha)\right]\\ \bar{u}(l,\alpha) &:= \mathbb{E}_l\left[u(s_{t+1}^i,\alpha)\right].\end{aligned}$$

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.

It is an important feature of this model that externalities are purely informational, i.e., each agent's utility is independent of the others' actions, and hence agents care about others' actions only because they may provide information. Furthermore, private signals are independent of actions, and so agents have no experimentation motive; they learn the same information from their signals, irregardless of the actions that they take.

2.3. **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 that maximizes the expected payoff at that period. In this 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, and we make this assumption for tractability, as does most of the learning literature.[10] A possible justification

[8] 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 $t+1$.

[9] 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.

[10] Indeed, 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).

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.[11]

We denote by p_t^i 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_t^i is a sufficient statistic for her expected payoff, her action a_t^i depends only on p_t^i. Formally, there exists a function $a^\star\colon [0,1] \to A$ such that with probability one[12]

$$a_t^i = a^\star(p_t^i)\,.$$

As information arrives independently of actions, and because agents are myopic, our model is not one of strategic experimentation: there are incentives to change one's action in order to learn more from one's own future signals, or from others' future actions. With these potentially confounding effect removed, we are left with a distilled model that allows us to study how observing others' actions differs from observing their signals.

2.4. **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. More precisely, at the end of each period an agent observes the actions taken by all other agents in this period.

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.

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

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.

[11]We conjecture that all our results generalize to the case of non-myopic agents, but this extension requires substantial technical innovation, beyond the techniques developed in this paper.

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

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; section 4 derives the learning dynamics in detail and explains how they lead to the results of this section. We consider the probability with which an agent i takes a suboptimal action in period t:

$$a_t^i \neq \alpha^\Theta \, .$$

We refer to this event as agent i "making a mistake" by "choosing the wrong action", even though she takes the action which is optimal given her information. As a benchmark we first briefly discuss the classical single agent case.

3.1. **Autarky.** In the single agent case $n = 1$, the probability of a suboptimal action is known to decay exponentially, with a rate r_a that can be calculated explicitly in terms of the cumulant generating functions $\lambda_h(z) = -\log \mathbb{E}_h\left[\mathrm{e}^{-z\,\ell}\right]$ and $\lambda_l(z) := -\log \mathbb{E}_l\left[\mathrm{e}^{z\,\ell}\right]$:[13]

Fact 1 (Speed of learning in autarky)**.** *The probability that a single agent in autarky chooses the wrong action in period t satisfies*[14]

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

where

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

[13]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.

[14]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 $\lim_{t\to\infty} f(t)/t = 0$.

This type of autarky result is classical in the statistics literature and can be found, for example, in studies of Bayesian hypothesis testing; see, e.g. Cover and Thomas (2006, pages 314-316). For us it serves as a benchmark for the case when agents try to learn from the actions of others. We prove Fact 1 in the Appendix, for the convenience of the reader.

Note, that the long-run probability of a mistake is 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. However, it is an asymptotic measure and in general does not say anything formally about what happens in early periods. Of course, the same is true for many statistical results, like the Central Limit Theorem, which nevertheless provide helpful intuition about what happens in finite periods.

3.2. **Many agents.** We now turn to the case where there are $n \geq 2$ agents. We first consider the benchmark case where all signals are observed by all agents. Since there is no private information, all agents hold the same beliefs, and this case reduces to the single agent case, but where n signals are observed in every period. After t periods the agents will have observed $n \cdot t$ signals, and so, by Fact 1, their probability of taking the wrong action will be the probability of error after $n \cdot t$ periods in the autarky setting.

Fact 2 (Speed of learning with public signals)**.** *When signals are public, the probability that any agent i chooses the wrong action in period t satisfies*

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

Having considered this benchmark case, we turn to our main model, in which $n \geq 2$ agents observe each others' actions, but signals are private. Our main result is that for any number of agents the speed of learning is bounded from above by a constant:

Theorem 1. *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*

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

In particular, this holds for $\bar{r}_b = \min\left\{\mathbb{E}_h\left[\ell\right], -\mathbb{E}_l\left[\ell\right]\right\}$. When private signals are Normal then one can take $\bar{r}_b = 4r_a$.

An immediate corollary from Theorem 1 and Fact 2 is the following result.

Corollary 1. *There exists a fixed group size k such that for any arbitrarily large group size n, the probability that any agent chooses the wrong action is eventually lower with k agents* *and public signals than with n agents and private signals. When signals are Normal we can take $k = 4$.*

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 almost all of the private signals are effectively lost, i.e. not aggregated in the decisions of others.

3.2.1. *Rational groupthink.* To prove this theorem 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 "*rational 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 rational groupthink, it holds, with high probability, that the private signals of each agent, which initially indicated the wrong action, eventually strongly indicate the *correct action*, but are still ignored due to the overwhelming information provided by the actions of others. We thus find the term rational groupthink an apt description of the phenomenon. We formally express this in the following proposition.

Proposition 1. *In the long run, conditional on the state being high and all agents taking the **incorrect, low** certainty action in every period, the private signals of every agent indicate the **correct, high** certainty action. That is, for every agent i and $\varepsilon > 0$ it holds that*

$$\lim_{t\to\infty} \mathbb{P}_h\left[q_i^t > 1-\varepsilon \mid a_s^j = \alpha^l \text{ for all } s \le t \text{ and all } j\right] = 1,$$

where $q_i^t = \mathbb{P}\left[\Theta = h \mid s_1^i, \ldots, s_t^i\right]$ is the probability assigned to the high state given only agent i's signals.

The analogous statement holds in the low state.

Note that Proposition 1 is *not* a consequence of the law of large numbers, as conditional on taking the wrong action the distribution of signals is not independent. Indeed, the result of Proposition 1 does not hold in the single agent case, where—in sharp contrast—conditional on choosing the wrong action the agent holds wrong beliefs. It shows that in a multi-agent learning problem agents will (with high probability) have received correct signals even conditioned on choosing the wrong action. This phenomenon, which does not have an analogue in sequential herding models, seems striking, as it does not involve irrationality, and yet results in a group taking an action which contradicts each and every member's private information.

3.2.2. *Early Period Mistake Probabilities.* Theorem 1 is a statement about asymptotic rates. In fact, if one were to increase 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. Thus, a more interesting setting is one in which, as we increase the number of agents, we decrease the informativeness of each agent's signal, while keeping fixed the amount of information available to all agents together.

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. In this setting, Theorem 1 implies that the speed of learning would be inversely proportional to the size of society, and in particular would tend to zero as n tends to infinity.

To test the robustness of this asymptotic speed of learning result, we perform a detailed analysis of the early periods, showing that, as the number of agents increases, they learn less and less from each other's actions. Thus, the asymptotic result of Theorem 1, which stated that the agents learn little from each other's actions in the long run, "kicks in" early on (in fact, already in the *second* period), in the sense that with high probability the agents learn nothing from each other's actions after the first period.

Theorem 2. *Suppose n agents have normal private signals with conditional distributions $\mathcal{N}(\pm 1, n)$ and want to match the state*[15]*, 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 strong enough to induce a deviation from the first period consensus. 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 have a larger variance. In this case, almost all information is lost even in early periods, if the number of agents is sufficiently high.

The intuition behind this result is the following: after observing the first round actions, the probability that a particular agent will have a strong enough signal to deviate from the majority opinion (action) is small. In fact, it is so small that the probability that *no*

[15] See Section 2.5.1.

agent deviates is almost one, and moreover it takes many periods until any agent has a strong enough signal to deviate. When agents observe that no one has deviated, it further strengthens (if not by much) their belief in the majority opinion, thus again delaying the breaking of the consensus. Of course, when the initial consensus is wrong, eventually it is broken.

4. Learning Dynamics

In this section we analyze the learning dynamics in detail and explain how we prove the results of Section 3. We discuss 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, which we view as a contribution of this paper.

4.1. **Preliminaries.**

As an agent's expected utility for a given action 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 lemma.

Lemma 1. *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 classical 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 1, 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.[16] 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 results 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 for the convenience of the reader.

Beliefs. We define the *private LLR* R_t as the LLR calculated based only 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. To this end, let ℓ have the same distribution as each ℓ_t, define $\lambda_\theta : \mathbb{R} \to \mathbb{R}$, the cumulant 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

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

[16]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 Normal case. Even in the Normal case it seems to us intractable to calculate in closed form the mistake probability in early periods in the multi-agent case.

Given these definition, we are ready to state the basic classical large deviations estimate that we use in this paper.

Lemma 2. *For any* $\mathbb{E}_l[\ell] < \eta < \mathbb{E}_h[\ell]$ *it holds that*[17]

$$\mathbb{P}_h[R_t \le \eta \cdot t + o(t)] = \mathrm{e}^{-\lambda_h^\star(\eta)\cdot t + o(t)}$$
$$\mathbb{P}_l[R_t \ge \eta \cdot t + o(t)] = \mathrm{e}^{-\lambda_l^\star(-\eta)\cdot t + o(t)}.$$

This Lemma states that the probability that the random walk R_t deviates from its (conditional) expectation is exponentially small, and decays with a rate that can be calculated exactly in terms of $\lambda_h^\star$ or $\lambda_l^\star$. The proof of Lemma 2 in the Appendix uses the properties of λ_θ and $\lambda_\theta^\star$ to verify that the increments of the LLR process in both states are such that large deviation theory results are applicable. Lemma 2 allows us to calculate the probability of a mistake conditional on each state, immediately implying Fact 1, which states that[18]

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

where $r_a = \lambda_h^\star(0) = \lambda_l^\star(0)$.

4.3. **Many Agents and the Groupthink Effect.** In this section we consider $n \ge 2$ agents. Each agent observes a sequence of private signals $s_1^i, \dots, s_t^i$, and the action taken by other agents in previous periods $(a_\tau^j)_{\tau<t, j\neq i}$. In this setting we prove Theorem 1. 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

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

The Probability that All Agents Make a Mistake in Every Period. To bound the probability of mistake, we consider the event G_t that all agents choose the action α^l in all time periods up to t:

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

[17] 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 $\mathbb{P}_h[R_t \le \eta \cdot t + f(t)] = \mathrm{e}^{-\lambda_h^\star(\eta)\cdot t + g(t)}$.

[18] We note that it is possible to strengthen this result 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.

To simplify the exposition we assume in the main text that G_t has strictly positive probability.[19] 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 *rational groupthink* event. The event G_t implies that all agents made a mistake in period t, conditioned on $\Theta = h$. Thus calculating the probability of G_t will provide a lower bound on the probability that a particular agent makes a mistake.

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 the action α^l—which is optimal in the low state—in all previous time periods, agent 2 assigns a higher probability to the low state and is more likely to also play the same action. This poses a difficulty for the analysis of this model, which 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 3. *There exists a sequence of thresholds $(q_\tau)_\tau$ such that the event G_t equals the event that no agent's private LLR R^i hits the threshold q before period t*

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

The proof of Lemma 3 in Appendix C shows this result recursively. Intuitively, whenever G_{t-1} occurs, all agents took the action α^l up to time $t-1$. By the induction hypothesis this implies that the private LLR of all other agents was below the threshold q_τ in all previous periods. As conditional on the states the private LLR's of different agents are independent, whether agent i takes the action α^l at time t conditional on G_{t-1}, depends only on her private LLR R_t^i. 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_τ. By symmetry, this is the same threshold for all agents.

[19]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 C we drop this assumption, slightly change the definition of G_t, and formally show that all our results also hold in general.

Calculating the Thresholds. We now provide a sketch of the argument (omitting many technical details) which we use in the appendix to characterize the threshold q_t. The threshold q_t admits a simple interpretation: it determines how high a private LLR R^i_t an agent must have in order to break from the consensus, and not take action α^l at time t, after having seen everyone take it so far. To calculate the q_t's we consider agent j's decision problem at time $t+1$, conditioned on G_t. The information available to her is her own private signals (summarized in her private log-likelihood ratio R^j_{t+1}), 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_τ in all previous time periods. Formally, knowing G_t is equivalent to knowing that

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

has occurred for all agents $i \neq j$.

How does knowing that agent i's private LLR has been below q_τ in all previous periods (i.e. W^i_t occurred) influence agent j's posterior? To answer this question we consider the log-likelihood ratio induced by this event:

$$\log \frac{\mathbb{P}_h\left[W^i_t\right]}{\mathbb{P}_l\left[W^i_t\right]}. \tag{4}$$

We show in Proposition 5 in the appendix that the logarithm of the probability of the event W^i_t conditioned on $\Theta = h$ is asymptotically the same as that of the event $R^i_t \leq q_t$, i.e., the event that agent i's private LLR is below the threshold q_t at just the last period:

$$\log \mathbb{P}_h\left[W^i_t\right] \approx \log \mathbb{P}_h\left[R^i_t \leq q_t\right].$$

Proposition 5 is similar in spirit to the Ballot Theorem of Bertrand (1887), which implies that the probability that a random walk is below a constant threshold in all prior periods approximately equals (up to sub-exponential terms) the probability that the random walk is below this threshold in the last period. We generalize this result in Proposition 5 by showing that the probability that a random walk is below a non-constant threshold $(q_t)_t$ in all prior periods asymptotically equals the probability that the random walk is below the linear threshold $q \cdot t$ with slope $q = \liminf_t q_t/t$ equal to the infimum of the slopes of the original threshold. This proposition is not an established large deviations result, but rather a contribution of this paper.

In Proposition 7 in the appendix we show that q_t is in fact asymptotically linear, i.e. the limit $q = \lim_{t\to\infty} {}^{q_t}/_t$ exists. This implies that $\log \mathbb{P}_h\left[W^i_t\right] = \log \mathbb{P}_h\left[R^i_t \leq q_t\right] + o(t)$. Thus, the large deviations estimate given in Lemma 2 implies that

$$\log \mathbb{P}_h \left[W_t^i\right] = \log \mathbb{P}_h \left[R_t^i \leq q_t\right] + o(t) = \log \mathbb{P}_h \left[R_t^i \leq q \cdot t\right] + o(t) = -\lambda_h^\star(q) \cdot t + o(t)\,. \tag{5}$$

In Lemma 7 in the Appendix we show that conditional on $\Theta = l$, the probability of the event W_t^i that agent i takes the correct action α^l in every period is strictly positive, i.e. there exists a constant $C > 0$ such that $\mathbb{P}_l\left[W_t^i\right] \in [C, 1]$ for all t. Thus, the LLR induced by the event W_t^i is

$$\log \frac{\mathbb{P}_h\left[W_t^i\right]}{\mathbb{P}_l\left[W_t^i\right]} = \log \mathbb{P}_h \left[W_t^i\right] - \log \mathbb{P}_l \left[W_t^i\right] = -\lambda_h^\star(q) \cdot t + o(t).$$

Since the event G_t that the private LLR of every agent is below q_τ in every period prior to t is the intersection of the individual events $G_t = \bigcap_{i=1}^n W_t^i$, and since these events $(W_t^i)_i$ are conditionally independent, we get that the log-likelihood ratio of G_t is simply a multiple of the LLR of W_t^1:

$$\log \frac{\mathbb{P}_h\left[G_t\right]}{\mathbb{P}_l\left[G_t\right]} = -(n-1) \cdot \lambda_h^\star(q) \cdot t + o(t).$$

The factor here is $n-1$ rather than n, since each agent observes only $n-1$ others. Thus, after observing G_t, agent j's posterior log-likelihood ratio will be the sum of her private LLR R_t^j and the LLR induced by observing G_t

$$L_t^j = R_t^j - (n-1) \cdot \lambda_h^\star(q) \cdot t + o(t)\,.$$

By Lemma 1, agent j will therefore take the action α^l in period $t+1$ if her signal is below $(n-1) \cdot \lambda_h^\star(q) \cdot t + o(t)$, which determines the new threshold q_{t+1}.

Thus, the threshold for the rational groupthink event at time $t+1$ will be

$$q_{t+1} = (n-1) \cdot \lambda_h^\star(q) \cdot t + o(t).$$

Dividing by t and taking the limit as t tends to infinity yields that (Proposition 7)

$$q = (n-1) \cdot \lambda_h^\star(q). \tag{6}$$

Note that q depends only on the private signal distributions, through $\lambda_h^\star$. Since $\lambda_h^\star$ is non-negative and decreasing, this equation will always have a unique solution. We have thus calculated q: it is the solution of the fixed point equation (6).

Intuitively, if the threshold is too high then it is likely that the others' private LLRs are below it, and so it is likely that they do not break the consensus. Thus an agent gains little information from observing them agreeing with the consensus, and her threshold for breaking the consensus will be low. This contradicts the initial assumption that the threshold is high. Likewise, if the threshold is too low, then an agent learns a lot by observing the consensus endure, and thus sets a high threshold for breaking it. The fixed point of (6) is the value in which these effects are equal.

Equation 6 determines the value of q, the slope of the threshold above which the agents break the consensus. We can use (5) to determine the probability of the event W_t^i that agent i does not break the consensus. Using the facts that the rational groupthink event G_t satisfies $G_t = \bigcap_{i=1}^n W_t^i$ and that the W_t^i's are conditionally independent, we thus have that

$$\mathbb{P}_h\left[G_t\right] = \mathbb{P}_h\left[W_t^i\right]^n = \mathrm{e}^{-q\cdot\frac{n}{n-1}\cdot t+o(t)}. \tag{7}$$

Consequently, the rate r_g of the event G_t that all agents take the wrong action in all periods up to time t is

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

We note that this rate can often be calculated explicitly. For example, for Normal private signals a straightforward calculation shows that

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

Finally, a convexity argument yields that this rate is bounded for any number of agents. We provide the proof in the appendix.

Proposition 2. *For any number of agents n it holds that $r_g < \mathbb{E}_h\left[\ell\right]$.*

As the rational groupthink event implies that all agents make a mistake, this provides a bound on the speed of learning, conditioned on $\Theta = h$:

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

Performing the corresponding calculation when conditioning on the low state, we have proved Theorem 1, for $\bar{r}_b = \min\left\{\mathbb{E}_h\left[\ell\right], -\mathbb{E}_l\left[\ell\right]\right\}$. In the case of Normal private signals, a tedious but straightforward calculation shows that $\bar{r}_b = 4r_a$.

5. Conclusion

We show that rational groupthink, a form of herding, occurs in a complex environment of agents who observe each other and take actions repeatedly. As a result, almost all information is lost when the group of agents is large. We use asymptotic rates as a measure of the speed of learning. As a robustness test, we show that the same effect holds also in the early periods, for the case of Normal signals.

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 incentive 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.

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

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.

Robert T Clemen and Robert L Winkler. Limits for the precision and value of information from dependent sources. *Operations Research*, 33(2):427–442, 1985.

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, 2006.

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

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

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

Krishna Dasaratha, Benjamin Golub, and Nir Hak. Social learning in a dynamic environment. 2018.

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.

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

Benjamin Golub and Matthew O Jackson. Naive learning in social networks and the wisdom of crowds. *American Economic Journal: Microeconomics*, 2(1):112–49, 2010.

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.

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. *Econometrica*, 83(5):1755–1794, 2015.

Manuel Mueller-Frank and Itai Arieli. Multi-dimensional social learning. *The Review of Economic Studies*, forthcoming, 2018.

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

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

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 $X_1, X_2, \ldots$ 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 3. *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 of Proposition 3. 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). □

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

$$\lambda^*(\eta) = \lambda(z) - \eta \cdot z. \tag{8}$$

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 (8) 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 proposition.

Proposition 4. *Let I be the interval on which λ is finite, and let $I^\star = \{\eta : \exists z \in \mathrm{int} I \;\, s.t. \;\, \lambda'(z) = \eta\}$. 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 3. *For every η such that $\eta > \inf_{z \in I} \lambda'(z)$ it holds that*

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

Proof of Theorem 3. For the upper bound, we use a Chernoff bound strategy: for any $z \ge 0$

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

and so by Markov's inequality

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

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

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

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

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

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

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

where $\nu^{(t)}$ is the t-fold convolution of ν with itself, and hence the law of Y_t. It is easy to verify[20] 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 $\eta_1 < \mathbb{E}\left[\tilde{X}\right] < \eta$ 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 (8), 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

$$\liminf_{t\to\infty} \frac{1}{t} \log \mathbb{P}\left[Y_t \leq \eta \cdot t + o(t)\right] \geq -\lambda^*(\eta_2). \tag{9}$$

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

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

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 4). Also, (9) 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 proposition 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

[20] 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 $\mathrm{e}^{z(x_1+\cdots+x_t)-\lambda(z)\cdot t}$.

the threshold to be linear, but only asymptotically and from one direction; both of these generalizations are important.

Proposition 5. *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 $\mathbb{P}[Y_t \le y_t] > 0$ it holds that*

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

Proof of Proposition 5. Let E_t be the event $\cap_{\tau=1}^{t} \{Y_\tau \le y_\tau\}$. 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 $t_k \le t$. Then by inclusion we have that

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

Using the same Chernoff bound strategy of the proof of Theorem 3, we get that

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

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

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

or

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

To show the other direction, define (as in the proof of Theorem 3) $\tilde{X}_t$ 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

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

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

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

where $\tilde{E}_t$ is the event $\cap_{\tau=1}^{t}\left\{\tilde{Y}_\tau \le y_\tau\right\}$. 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

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

Since the expectation of $\tilde{Y}_t/t$ is strictly between $\eta = \liminf_t y_t/t$ and $\eta - \epsilon$, we have that $\lim_t \mathbb{P}\left[\tilde{Y}_t \ge y_t - \epsilon\cdot t\right] = 1$ 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

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

Proceeding as in the proof of Theorem 3 yields that

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

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

$$\lambda_h(z) = -\log\int \exp\left(-z\cdot\log\frac{\mathrm{d}\mu_\mathrm{h}}{\mathrm{d}\mu_l}(s)\right)\mathrm{d}\mu_h(s) = -\log\int\left(\frac{\mathrm{d}\mu_\mathrm{l}}{\mathrm{d}\mu_h}(s)\right)^z\mathrm{d}\mu_h(s). \tag{10}$$

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

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

Furthermore, as for every η between $\mathbb{E}_h\left[\ell\right]$ and $\mathbb{E}_l\left[\ell\right]$ the maximum in the definition of $\lambda^\star_h$ is achieved for some $z \in (0,1)$, it follows that there is also a simple connection between $\lambda^\star_h$ and $\lambda^\star_l$:

$$\lambda^\star_l(\eta) = \lambda^\star_h(-\eta) - \eta. \tag{11}$$

We will accordingly state some results in terms of λ_h and $\lambda^\star_h$ only. It also follows from (10) that the interval I on which λ_h is finite contains $[0,1]$. Since from the definitions we have that $\lambda'_h(0) = \mathbb{E}_h\left[\ell\right]$, and since $\lambda'_h(1) = \mathbb{E}_l\left[\ell\right]$ by the relation between λ_h and λ_l, we have shown the following lemma.

Lemma 4. *$\lambda_\theta(z)$ and $\lambda^\star_\theta(\eta)$ are finite for all $z \in [0,1]$ and $\eta \in (\mathbb{E}_l\left[\ell\right], \mathbb{E}_h\left[\ell\right])$. Furthermore,*

$$\lambda_h(z) = \lambda_l(1-z) \text{ and } \lambda^*_h(\eta) = \lambda^*_l(-\eta) - \eta\,. \tag{12}$$

Proof of Lemma 2. Given Lemma 4, Lemma 2 is an immediate corollary of Theorem 3. ☐

The following simple observation will be useful on several occasions:

Lemma 5. *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 of Lemma 5. 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 4, $\lambda_h^\star(r_a) > 0$ 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 Fact 1. Consider the case $\Theta = h$. As shown in Lemma 1 the probability that the agent makes a mistake is equal to the probability that the LLR is below $\underline{L}(\alpha^h)$. Thus, Lemma 2 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 (12) $\lambda_h^\star(0) = \lambda_l^\star(0)$. ☐

Appendix C. Many Agents

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 3. 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

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

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

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

Proposition 6. *The threshold q_t is characterized by the recursive relation*

$$q_t = \overline{L}(\alpha^l) - (n-1) \cdot \log \frac{\mathbb{P}_h\left[W_{t-1}^1\right]}{\mathbb{P}_l\left[W_{t-1}^1\right]} \quad \text{and} \quad W_t^i = \bigcap_{1 \le \tau \le t} \{R_\tau^i \le q_t\}. \tag{13}$$

Proof of Proposition 6. Agent 1's log-likelihood ratio conditional on $\cap_{i=1}^n W_{t-1}^i$ at time t equals b

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

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

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

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

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

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

$$q_t = \overline{L}(\alpha^l) - (n-1) \cdot \log \frac{\mathbb{P}_h\left[W^1_{t-1}\right]}{\mathbb{P}_l\left[W^1_{t-1}\right]}, \tag{14}$$

then

$$L^1_n = R^1_t - q_t + \overline{L}(\alpha^l),$$

and $L^1_t \leq \overline{L}(\alpha^l)$ (and thus $a^1_t = \alpha^l$) whenever $P^1_t \leq q_t$. □

Lemma 6. *$q_t \geq \overline{L}(\alpha^{\min}_t)$ for all t.*

Proof of Lemma 6. 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 $F_l(x) \geq F_h(x)$ for all $x \in \mathbb{R}$. It follows that the joint distribution of $\{R^i_\tau\}_{\tau \leq t}$ conditioned on $\Theta = h$ dominates the same distribution conditioned on $\Theta = l$, and so $\mathbb{P}_h\left[W^1_t\right] \leq \mathbb{P}_l\left[W^1_t\right]$. Hence $q_t \geq \overline{L}(\alpha^{\min}_t)$. □

Lemma 7. *There is a constant $C > 0$ such that $\mathbb{P}_l\left[W^1_t\right] \geq C$ for all t.*

Proof of Lemma 7. Since the events W^1_t are decreasing, we will prove the lemma by showing that

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

which by definition is equivalent to

$$\lim_{t\to\infty} \mathbb{P}_l\left[\cap_{\tau \leq t}\left\{R^i_\tau \leq q_\tau\right\}\right] > 0.$$

Since $q_t \geq \overline{L}(\alpha^{\min}_t)$, it suffices to prove that

$$\lim_{t\to\infty} \mathbb{P}_l\left[\cap_{\tau \leq t}\left\{R^i_\tau \leq \overline{L}(\alpha^{\min}_\tau)\right\}\right] > 0.$$

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^i_t as t tends to infinity must be $-\infty$. Thus, with probability 1, for all t large enough it does hold that $R^1_t \leq \overline{L}(\alpha^{\min}_\tau)$. Since each of the events W^1_t has positive probability, and by the Markov property of the random walk R^1_t, it follows that the event $\cap_\tau\left\{R^i_\tau \leq \overline{L}(\alpha^{\min}_\tau)\right\}$ has positive probability. Finally, by monotonicity

$$\lim_{t\to\infty} \mathbb{P}_l\left[W^1_t\right] > \mathbb{P}_l\left[\cap_\tau\left\{R^i_\tau \leq \overline{L}(\alpha^{\min}_\tau)\right\}\right] > 0.$$

□

It follows immediately from this Lemma 7 and Proposition 6 that

$$\lim_{t\to\infty} \frac{q_t}{t} = -(n-1)\lim_t \frac{1}{t} \log \mathbb{P}_h\left[W^1_{t-1}\right], \tag{15}$$

provided that the limit exists.

Let $\underline{q} = \liminf_{t\to\infty} q_t/t$. Since $W_t^i = \cap_{\tau=1}^t \{R_\tau^i \le q_\tau\}$, it follows from Proposition 5 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 $\underline{q} \ge 0$ (Lemma 6), and so this indeed holds. Thus, by (15), we have proved the following proposition:

Proposition 7. *The limit $q = \lim_{t\to\infty} \frac{q_t}{t}$ exists, and*

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

Proof of Proposition 2. Recall that $\lambda_h^\star$ is strictly convex, and that $\lambda_h^\star(D) = 0$, where we denote $D = \mathbb{E}_h\left[\ell\right]$. 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 $\lambda_h^\star(0) < D$ (Lemma 5) we have shown that

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

and so

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

We now turn to proving Proposition 1, which states that conditioned on rational groupthink—that is, conditioned on the event G_t—all agents have, with high probability, a private LLR R_t^i that strongly indicates the correct action. In fact, we prove a stronger statement, which implies Proposition 1: the private LLR is arbitrarily close to $q \cdot t$, the asymptotic threshold for R_t^i above which rational groupthink ends.

Proposition 8. *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 3 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 $\{R_t^i \leq t\cdot(q-\epsilon) \text{ for all } i\}$. 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. □

Appendix D. Early Period Mistake Probabilities

We now prove Theorem 2. 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.[21] 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:

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

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

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

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

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

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

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 9. *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 2. We in fact provide a finitary statement and prove that $\mathbb{P}\left[E_t\right] \geq 1 - 20 \cdot t \cdot \sqrt{\frac{\log n}{n}}$.

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 Normal, 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

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

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

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

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 Lemma 6. But since, under the event $A \cap B$, each of their private LLRs $\hat{R}^i_\tau$ 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 9, and thus of Theorem 2.