UNIVERSITY OF CAMBRIDGE
Faculty of Economics

Institute for New Economic Thinking

Cambridge Working Papers in Economics

Cambridge-INET Working Paper Series No: 2020/39

Cambridge Working Papers in Economics: 2085

LEARNING IN A SMALL/BIG WORLD

Benson Tsz Kin Leung
(University of Cambridge)

This paper looks into how learning behavior changes with the complexity of the inference problem and the individual's cognitive ability, as I compare the optimal learning behavior with bounded memory in small and big worlds. A learning problem is a small world if the state space is much smaller than the size of the bounded memory and is a big world otherwise. I show that first, optimal learning behavior is almost Bayesian in small worlds but is significantly different from Bayesian in big worlds. Second, ignorant learning behaviors, e.g., availability heuristic, correlation neglect, persistent over-confident, are never optimal in small worlds but could be optimal in big worlds. Third, different individuals are bound to agree in small worlds but could disagree and even be bound to disagree in big worlds. These results suggest that the complexity of a learning problem, relative to the cognitive ability of individuals, could explain a wide range of abnormalities in learning behavior.

Keywords: Learning, Bounded Memory, Bayesian, Ignorance, Disagreement
JEL Codes: D83, D91

Learning in a Small/Big World*

Benson Tsz Kin Leung†

August 27, 2020

Abstract

This paper looks into how learning behavior changes with the complexity of the inference problem and the individual's cognitive ability, as I compare the optimal learning behavior with bounded memory in small and big worlds. A learning problem is a small world if the state space is much smaller than the size of the bounded memory and is a big world otherwise. I show that first, optimal learning behavior is almost Bayesian in small worlds but is significantly different from Bayesian in big worlds. Second, ignorant learning behaviors, e.g., availability heuristic, correlation neglect, persistent over-confident, are never optimal in small worlds but could be optimal in big worlds. Third, different individuals are bound to agree in small worlds but could disagree and even be bound to disagree in big worlds. These results suggest that the complexity of a learning problem, relative to the cognitive ability of individuals, could explain a wide range of abnormalities in learning behavior.

Keywords: Learning, Bounded Memory, Bayesian, Ignorance, Disagreement

JEL codes: D83, D91

*I am grateful to Jacques Crémer, Renato Gomes, Philippe Jehiel, Hamid Sabourian, Mikhael Safronov, Larry Samuelson, Tak-Yuen Wong and seminar audiences for insightful discussion and comments.

†University of Cambridge. Email: btkl3@cam.ac.uk.

1 Introduction

Many experimental and empirical studies have documented different behaviors of belief formation that systematically differ from the Bayesian model,[1] e.g., the use of heuristics (Kahneman et al. (1982)), correlation neglect (Enke and Zimmermann (2019)), persistent over-confidence (Hoffman and Burks (2017)), inattentive learning (Graeber (2019)), etc. Although different mechanisms have been proposed for different phenomena, our understanding of these learning behaviors is still inadequate. In particular, few attempts explain how the prominence of these "abnormalities" changes across different situations, and even fewer propose a more unified mechanism to explain different types of "abnormalities". Informally, these departures from the Bayesian model are often attributed to the complexity of employing the Bayes rule. However, to the best of my knowledge, no study formally analyzes how the complexity of an inference problem affects individuals' learning behaviors. Are "abnormalities" less prominent in less complicated problems? How do learning behaviors change with the complexity of the inference problems? This paper aims to answer these questions and explain different "abnormal" learning behaviors in light of complexity.

Every day we form beliefs over many uncertainties to guide our decision making, from predicting the weather and deciding whether to bring an umbrella outdoor, where to get chocolate in a particular supermarket, to estimating and preparing for the impact of Brexit. Given our limited cognitive ability, intuitively the complexity of the inference problem would affect the way we form beliefs. After several trips to the same supermarket, we would be fairly sure about where to look for chocolate; but even after collecting numerous data about the stock market, we make mistakes in our investment decisions. We are also more likely to disagree on complicated problems, e.g., the impact of Brexit or global warming, but agree on simpler problems, whether to bring an umbrella outdoor. Similarly, given an inference problem, different individuals with different levels of cognitive ability would perceive the level of complexity differently, learn differently, and could disagree with each other even after receiving the same and many pieces of information.

In this paper, I analyze a simple model to compare the optimal learning behavior of an individual in small and big world problems.[2] The former refers to situations where the complexity of the inference problem is low relative to the cognitive ability

[1] There are plenty of examples, see for example the seminal work of Kahneman et al. (1982), Kahneman (2011) and the section 3 of the review article Rabin (1998).

[2] The terms "small worlds" and "big worlds" are inspired by the seminal work of Savage (1972). Differently, in Savage (1972) (and also the related study of Mailath and Samuelson (2020)), big worlds refer to complicated inference problems where it is difficult for individuals to form a prior belief on states and signal structures, or even to construct the state space. In contrast, this paper assumes individuals have clear (but perhaps subjective) ideas on the state space and signal structures but defines big worlds based on the size of the state space relative to the cognitive ability of individuals.

	Small Worlds (small $\frac{N}{M}$)	Big Worlds (big $\frac{N}{M}$)
Is learning close to Bayesian?	Yes	No
Is ignorance in learning "optimal"?	No	Could be
Is disagreement persistent?	No	Could be

Table 1: Differences in optimal learning in small/big worlds

of the individual, and the latter refers to situations where the complexity of the inference problem is high relative to the individuals' cognitive ability.[3] In other words, I analyze how the relative complexity affects learning behavior.

Consider an individual who tries to learn the true state of the world from a finite state space $\Omega = \{1, 2, \cdots, N\}$. In each period $t = 1, \cdots, \infty$, before receiving a signal $s_t \in S$, he takes an action $a_t \in \mathscr{A} = \Omega$ which he aims to match with the true state. To model limited cognitive ability, I assume his belief is confined to a M sized automaton that captures bounded memory, as in the seminal work of Hellman and Cover (1970).[4] A belief updating mechanism with a M sized automaton comprises a (potentially stochastic) transition rule[5] $\mathscr{T} : M \times S \to \triangle M$ and a (potentially stochastic) decision rule $d : M \to \triangle\mathscr{A}$. That is, at each period t, he starts in a memory state $m_{t-1} \in \{1, \cdots, M\}$, receives a signal s_t, then transits to memory state a m_t according to $\mathscr{T}(m_t, s_t)$ and takes an action a_t according to $d(m_t)$.

In contrast to Bayesian, the individual with bounded memory has a much coarser idea of the likelihood of different states of the world, and the coarseness decreases in M. Thus, N measures the absolute complexity/size of the world, and M measures the cognitive ability of the individual. I define small worlds as cases where $N/M \to 0$, or in general when N/M is very small.[6] In contrast, big worlds refer to cases when N/M is bounded below from some strictly positive real number. As previously mentioned, whether a problem is a small or big world problem depends on the relative size of the world with respect to the individual's cognitive ability.

I compare the characteristics of the optimal updating mechanisms $(\mathscr{T}^*, d^*)$ that maximizes the asymptotic probability that the individual matches his actions with the true state, in small and big worlds. I show three differences in learning behavior, which are summarized in table 1. The first comparison is about whether asymptotic learning is close to Bayesian in small and big worlds. I show that in small worlds,

[3]Intuitively, if individuals have super cognitive/memory ability, any complicated problem would look like a small problem.

[4]See Compte and Postlewaite (2012), Wilson (2014), Monte and Said (2014) and Basu and Chatterjee (2015) in the Economic literature that model bounded memory with finite automaton.

[5]The Bayesian analogue of this bounded memory setting would be M equals the space of probability simplex and the transition rule equals to Bayesian formula.

[6]Note that if the individual tracks his belief not with a finite automaton but with a real number statistics. The cardinality of the belief statistics is much bigger than N and the model collapses to a Bayesian model.

asymptotic learning behavior is very close to Bayesian, i.e., individuals almost perfectly learn the true state; while in big worlds, asymptotic learning is significantly different from Bayesian.

The second comparison concerns ignorance learning behaviors that mentioned at the beginning of the introduction. Consider for example the phenomenon of persistent over-confidence (Hoffman and Burks (2017), Heidhues et al. (2018)). Suppose the state of the world comprises ability and luck $\omega = (\text{ability}, \text{luck}) \in \{H, L\} \times \{H, L\}$ and the individual observes performance as signals. In the setting of bounded memory, persistent over-confidence happens when the individual never chooses the actions that correspond to the states $\{L\} \times \{H, L\}$ even after observing a long sequence of bad performance.[7] He only takes actions that correspond to either (H, H) or (H, L), thus acts as if he always believes he has high ability and only updates his belief on how much luck he has. Similarly, other heuristic or inattentive learning behaviors could also be modeled as the individual ignores some states of the world.[8] I show that such ignorance learning behaviors are never optimal in small worlds, but could be optimal in big worlds as the individual may focus learning on a subset of states given his scarce cognitive resources.

Lastly, I look into the phenomenon of asymptotic disagreement in small and big worlds. Consider two individuals who start with different prior beliefs over Ω but observe a sequence of public signals and have the same objective information structures; or two individuals with the same prior belief over Ω but observe a different sequence of private signals generated by different signal structures. I show that in small worlds, the two individuals almost always agree with each other in the sense that they almost always take the same action asymptotically. In contrast, in big worlds, the two individuals could be bound to disagree with each other even after receiving infinitely many signals. Moreover, in big worlds, asymptotic disagreement could be solely driven by differences in cognitive ability: even when two individuals have the same prior beliefs and observe a sequence of public signals with an objective signal structure, they could disagree with probability 1 as $t \to \infty$. This is in contrast with the existing literature that explains long-term disagreement based on differences in prior beliefs (Rabin and Schrag (1999)) or uncertainties in signal structures (Acemoglu et al. (2016)).

The paper is organized as follows. In the next section, I present the model setup. Then I analyze the optimal learning behavior in small world problems and big world problems in sections 3 and 4 respectively. In section 5, I conclude by presenting a detailed discussion of the results and the connection with the existing literature. Proofs and omitted results are presented in the appendix.

[7] As I assume one-to-one mapping from actions to states, if an individual never chooses an action, it is as if he never pay attention to the corresponding state.

[8] I provide another example with the availability heuristic in section 2.1.

2 Model Setting

Consider a world with N possible true states of the world, i.e., $\omega \in \Omega = \{1, 2, \cdots, N\}$, and a decision-maker (DM) who wants to learn the true state. I analyze two cases that correspond to where N is fixed and finite, and at the limit that N goes to infinite, because of their differences in both technical and economic implications.[9],[10] In each period $t = 1 \cdots, \infty$, the DM takes an action/guess $a_t \in \mathscr{A} = \Omega$ and he gets utility $u_N^\omega > 0$ if he infers correctly the true state as ω, i.e., $u_N(a, \omega) = u_N^\omega$ if $a = \omega$; otherwise, his utility equals 0. Thus, u_N^ω measures how important it is to identify the state ω. I assume no state is infinitely more important than any other state:

Assumption 1. *For all ω, $u_N^\omega \in [\underline{u}, \overline{u}]$ where $\underline{u} > 0$ and $\overline{u}$ is finite.*

A special case that satisfies the assumption is that u_N^ω is constant across ω, that is, the DM does not intrinsically discriminate any states. With some abuse of notations, sometimes I denote action $a = \omega$ by action ω. The (potentially subjective) prior belief of the DM is denoted as $(p_N^\omega)_{\omega=1}^N$ where $\sum_{\omega=1}^N p_N^\omega = 1$ and I make the following full support assumptions:

Assumption 2 (Full support prior with finite N)**.** *In the case where N is finite, $p_N^\omega > 0$ for all ω.*

Assumption 2' (Full support prior with infinite N)**.** *In the case where $N \to \infty$, for any sequence of set $\Omega_N \subseteq \Omega$ with cardinality $|\Omega_N|$ and a well-defined limit $\lim_{N\to\infty} \frac{|\Omega_N|}{N}$,*

$$\lim_{N\to\infty} \sum_{\omega\in\Omega_N} p_N^\omega > 0 \text{ if } \lim_{N\to\infty} \frac{|\Omega_N|}{N} > 0. \tag{1}$$

Equation (1) ensures any sequence of sets of states that has non-negligible measure in fraction at the limit, i.e., $\lim_{N\to\infty} \frac{|\Omega_N|}{N} > 0$, also has a non-negligible probability mass at the limit, i.e., $\lim_{N\to\infty} \sum_{\omega\in\Omega_N} p_N^\omega > 0$.

In each period after taking an action,[11] the DM receives a signal $s_t \in S$ which are independently drawn across different periods from a continuous distribution with p.d.f. f_N^ω in state ω.[12] I assume no signals perfectly rule out any states of the world:

Assumption 3 (No perfect signals with finite/infinite N)**.** *There exists $\varsigma > 0$ such that $\inf_{s\in S} f_N^\omega(s)/f_N^{\omega'}(s) > \varsigma$ for all ω, $\omega' \in \Omega$.*

[9]Roughly speaking, analyzing the case where $N \to \infty$ allows me to show that the behavioral implications depend on $\frac{N}{M}$ instead of the absolute value of N or M.

[10]As discussed in section 1.3.4 of Mailath and Samuelson (2020), arguably, the DM could include in the state space infinitely many potential variables that affect the signal distributions such that N always goes to infinity. I offer a brief discussion in section 5 on what constructs a state space for the DM based on existing experimental studies and the resulting implications.

[11]The order, i.e., whether the DM receives a signal before or after taking an action in each period, does not affect the result. The crucial assumption is that the action chosen by the DM at each period depends only on the memory state he is in (his "belief"), but does not (directly) depend on the signals he received.

[12]For ease of exposition, I assume signals follow a continuous distribution but the results hold with more general probability measures.

Assumption 3 says that the information structure in every state has the same support. In other words, given any signal realization s_t, the DM will not learn perfectly the true state. I also assume the following identifiability assumption such that no pairs of signal structures are the same.

Assumption 4 (Identifiability with finite N)**.** *In the case where N is finite, there are no ω and $\omega' \neq \omega$ such that $f_N^{\omega}(s) = f_N^{\omega'}(s)$ for (almost) all $s \in S$.*

In the case of infinite N, I make the following assumption about identifiability.

Assumption 4' (Identifiability with infinite N)**.** *In the case where $N \to \infty$, for all $\epsilon > 0$, there exists some $\xi > 0$ and some sequence of subsets of states $N_\xi \subseteq \Omega$ such that $\lim_{N\to\infty}\left(\sum_{\omega \in N_\xi} p_N^{\omega}\right) > 1-\epsilon$ and*

$$\lim_{N\to\infty} \inf_{\omega,\omega' \in N_\xi;\omega' \neq \omega} \left\{ -\log \frac{\int f_N^{\omega}(s) f_N^{\omega'}(s) ds}{\sqrt{\int (f_N^{\omega}(s))^2 ds}\sqrt{\int (f_N^{\omega'}(s))^2 ds}} \right\} > \xi,$$

such that signal structures with negligible Cauchy-Schwarz distances must have negligible probability mass at the limit where $N \to \infty$.[13] In standard Bayesian setting, the DM learns almost perfectly the true state when $t \to \infty$.[14] In contrast, in this paper, I focus on a bounded memory setting that will be outlined below.

The DM is subject to a memory constraint such that he can only update his belief using a M memory states automaton. That is, in each period, his belief is represented by not a $N-1$ dimensional probability simplex but a memory state $m_t \in \{1, 2, \cdots, M\}$. Upon receiving a signal s_t in period t, the DM updates his belief from memory state m_t to $m_{t+1} \in \{1, 2, \cdots, M\}$. An updating mechanism specifies a (potentially stochastic) transition function between the M memory states given a signal $s \in S$, denoted as $\mathscr{T} : M \times S \to \triangle M$, and a (potentially stochastic) decision rule $d : M \to \triangle A$. Note that the updating mechanism $(\mathscr{T}, d)$ is restricted to be stationary across all $t = 1 \cdots, \infty$ to capture the idea of bounded memory.[15,16] Given an updating mechanism $(\mathscr{T}, d)$, the time-line of a given period t is summarized in figure 1.

An example of updating mechanism. Consider a simple example where there are two possible states, i.e., $N = 2$, and four memory states, i.e., $M = 4$. The DM has a uniform prior belief $p_2^1 = p_2^2 = 0.5$. For simplicity, assume there does not exists any uninformative signals, i.e., $\nexists$ s such that $f_2^1(s) = f_2^2(s)$. Figure 2 shows

[13] I provide an example that satisfies assumption 3 and 4' in appendix B.

[14] See Blackwell and Dubins (1962).

[15] As discussed in Hellman and Cover (1970), a non-stationary updating mechanism implicitly assumes the ability to memorize time, thus implicitly assumes a larger memory capacity.

[16] Note that switching between multiple M memory state automatons requires more than M memory states, as illustrated in appendix A. Appendix A also shows an example to illustrate that the current setting allows switching between smaller sized automatons, e.g., three automatons with $\frac{M+2}{3}$ memory states.

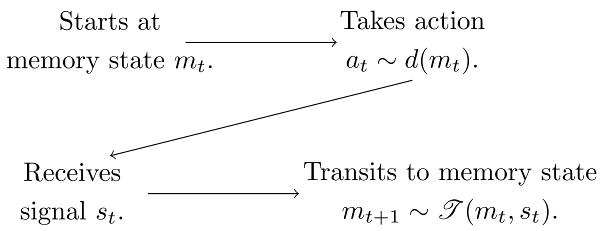

Figure 1: Time-line at period t given an updating mechanism $(\mathscr{T}, d)$.

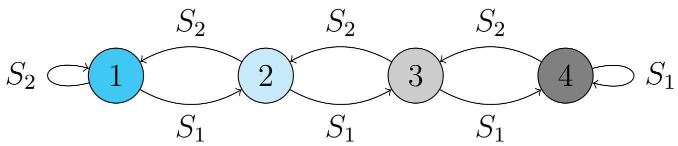

Figure 2: A simple updating mechanism with $N = 2$ and $M = 4$. S_1 denotes the set of signals where $f_2^1(s) > f_2^2(s)$ and $S_2 = S \setminus S_1$. The DM believes that state 1 is more likely as he moves towards the higher memory states. The DM's actions follow: $d(1) = d(2) = 2$ and $d(3) = d(4) = 1$.

a simple updating mechanism. The DM moves one memory state higher whenever he receives a signal supporting state 1 where $f_2^1(s) > f_2^2(s)$ and moves one memory state lower otherwise. Moreover, his decision rule is such that he chooses action 1 if and only if he is in memory state 3 or 4. With bounded memory, instead of tracking his belief in the segment $[0, 1]$, he only holds a rough idea on how likely state 1 or state 2 are: he believes that state 1 (resp. state 2) is "very likely" if he is in memory state 4 (resp. memory state 1), and is "likely" if he is in memory state 3 (resp. memory state 2). ♦

In this setup N represents how big (or complicated) the world is, while M represents the cognitive resources of the DM. When M increases, the DM has a finer "belief space" and more instruments to track beliefs. This gives a natural definition of small and big world problems based on relative or perceived complexity: an inference problem is a small world problem where N is much smaller than M and is a big world problem otherwise. The formal definition is as follows and summarized in table 2:

Definition 1 (Small and Big World with finite N)**.** *In the case where N is finite, the inference problem is a small world problem at the limit where $M \to \infty$ and is a big world problem if M is finite.*

In the case where $N \to \infty$, I analyze the interesting cases where $M \to \infty$ with slower and faster rate than N.[17]

Definition 1' (Small and Big World with infinite N)**.** *In the case where $N \to \infty$, the inference problem is a small world problem if $M \to \infty$ and $N = O(M^h)$ where*

[17] When M is finite and $N \to \infty$, most of the results are trivial. For example, the DM will not learn perfectly the true state.

	Small Worlds	Big Worlds
Finite N	At the limit of infinite M	Finite M
At the limit of infinite N	$N = O(M^h)$ where $h < 1$	$N = O(M^h)$ where $h \geq 1$

Table 2: Definition of small and big worlds.

$h < 1$ and is a big world problem if $M \to \infty$ and $N = O(M^h)$ where $h \geq 1$.

In short, small worlds are inference problems with $\frac{N}{M}$ close to 0 and big worlds are those with $\frac{N}{M}$ bounded away from 0. It is important to define small/big worlds based on the relative size of the state space with respect to the cognitive ability of the DM, instead of the absolute size of the state space. In particular, if the DM can track his belief using the real number space, i.e., M is uncountably infinite, then the model collapses to a standard Bayesian model and any inference problem with countable state space is a small world problem.[18]

This paper analyzes the asymptotic learning of the DM, i.e., the DM aims to choose an updating mechanism that maximizes his expected long run per-period utility:

$$E\left[\lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}u_N(a_t,\omega)\right].$$

Given state $\omega \in \Omega$, the sequence m_t, together with some specified initial memory state, forms a Markov chain over the signal space S. Note that fixing M, the transition function $\mathscr{T}$ could be represented by a transition matrix between the memory states given a signal realization $s \in S$:

$$\mathbf{Q}(s) = [\Pr\{\mathscr{T}(i,s) = j\}] = [q_{ij}(s)] \tag{2}$$

for $i, j = 1, 2, \cdots, M$ where $\sum_{j=1}^{M} q_{ij}(s) = 1$ and $q_{ij}(s) \geq 0$ for all i, j, s. Taking the expectation over s, the transition probability matrices under state ω follows:

$$\mathbf{Q}_N^\omega = \int \mathbf{Q}(s) f_N^\omega(s)\, ds. \tag{3}$$

The stationary probability distributions over the memory states, denoted by $\boldsymbol{\mu}_N^\omega = (\mu_{N1}^\omega, \mu_{N2}^\omega, \cdots, \mu_{NM}^\omega)^T$, is the solution to the following system of equations:

$$\boldsymbol{\mu}_N^\omega = (\boldsymbol{\mu}_N^\omega)^T \mathbf{Q}^\omega. \tag{4}$$

Given the stationary probability distribution over the memory states, the conditional probability of state ω given the DM is in memory state m equals $\frac{p_N^\omega \mu_{Nm}^\omega}{\sum_{\omega' \in \Omega} p_N^{\omega'} \mu_{Nm}^{\omega'}}$

[18]The result is obvious when $N = 2$. For $N > 2$, the result is implied by the fact that there exists an one-to-one mapping between $\mathscr{R}$ to $\mathscr{R}^{N-1}$ for all N.

so it would be optimal to choose $d^*(m, \mathscr{T}) = \arg\max_{\omega'} u^{\omega'} p_N^{\omega'} \mu_{Nm}^{\omega'}$. Thus, I often refer the transition function $\mathscr{T}$ as the updating mechanism by implicitly assuming $d(m) \in d^*(m, \mathscr{T})$. Furthermore, for the ease of exposition, I restrict attention to degenerate decision rules throughout the paper unless it is stated otherwise.[19] Denote M^ω as the set of memory states in which the DM will take action ω. The asymptotic utility of an updating mechanism $\mathscr{T}$ is equal to:

$$U_N(\mathscr{T}) = \sum_{\omega=1}^{N} \left[u_N^\omega p_N^\omega \left(\sum_{m \in M^\omega} \mu_{Nm}^\omega \right) \right] \tag{5}$$

and the asymptotic utility loss of an updating mechanism $\mathscr{T}$ is equal to:

$$L_N(\mathscr{T}) = \sum_{\omega=1}^{N} \left[u_N^\omega p_N^\omega \left(1 - \sum_{m \in M^\omega} \mu_{Nm}^\omega \right) \right]. \tag{6}$$

The DM maximizes asymptotic utility or, by duality, minimizes the asymptotic utility loss associated with the updating mechanism. In the paper, I will mostly refer the optimal design of updating mechanism as the minimization of L_N. In general, as shown in Hellman and Cover (1970), an optimal $\mathscr{T}$ may not exist.[20] Therefore, the rest of the paper focuses on ϵ-optimal updating mechanisms that are defined as follows. Denote $L_{NM}^* = \inf_{\mathscr{T}} L_N(\mathscr{T})$. An updating mechanism $\mathscr{T}$ is ϵ-optimal if and only if $L_N(\mathscr{T}) \leq L_{NM}^* + \epsilon$.

2.1 Small vs Big Worlds

In the subsequent sections, I will compare the learning behaviors, that is, the characteristics of the ϵ-optimal updating mechanisms (for small ϵ) in small and big worlds by answering the following questions.

Is learning close to Bayesian? The first question I ask is whether learning behavior of the DM is close to that of a Bayesian individual in small and big worlds. As we know, in the current setup, a Bayesian individual will (almost) perfectly learn the true state of the world asymptotically, no matter how big N is. Thus this question is equivalent to whether L_{NM}^* is close to 0 in small and big worlds.

Definition 2. *The asymptotic learning behavior of the DM is close to Bayesian if*

[19] Some proofs are by contradiction with proposed deviations to non-degenerate decision rules. Obviously, if there is a beneficial deviation to updating mechanisms with non-degenerate decision rules, there must be (weakly) better deviations with degenerate decision rules.

[20] This is true because signals are continuously distributed. Roughly speaking, consider $N = M = 2$, suppose the DM could improve his asymptotic utility by switching from $m = 1$ to $m = 2$ with signals that strongly support state 1, i.e., with large $\frac{F_2^1(S_1)}{F_2^2(S_1)}$. For any set of signal realizations S_1 with strictly positive measure $F_2^1(S_1), F_2^2(S_1)$, the DM can always find another set of signal realizations S_1' with strictly positive measure $F_2^1(S_1')$ and $F_2^2(S_1')$, and higher likelihood ratio $\frac{F_2^1(S_1')}{F_2^2(S_1')}$, thus improves his asymptotic utility.

and only if $L^*_{NM} = 0$ (or $\lim_{M\to\infty} L^*_{NM} = 0$ or $\lim_{N,M\to\infty} L^*_{NM} = 0$ depending on the situation in question as summarized in table 2).

The answer to this question sheds light on how well the Bayesian model approximates individual's learning and decision making in different problems, which would give us an idea on how robust theoretical results are to the setting of bounded memory.

Is ignorance (close to) optimal? Behavioral economics and psychology literature has documented different types of ignorance learning behaviors, including the use of heuristics (Kahneman et al. (1982)), correlation neglect (Enke and Zimmermann (2019)), persistent over-confidence (Hoffman and Burks (2017)), ignorance of informational content of others' strategic behaviors (Eyster and Rabin (2005), Jehiel (2005, 2018)) or in general ignorance of some relevant variables (Graeber (2019)). Consider the classical example of availability heuristic. In the experiment of Tversky and Kahneman (1973), the majority of participants reported that there were more words in the English language that start with the letter K than for which K was the third letter, while the correct answer is the reverse. The proposed explanation is that individuals use availability heuristic: they pay attention only to the easiness of recalling the two types of words but ignore the fact that it is easier to recall words starting with K than words with K as the third letter. These ignorance behaviors are sometimes thought to be naive or inattentive, while sometimes viewed as efficient given our limited ability (Gigerenzer and Goldstein (1996), Gigerenzer and Brighton (2009), Gigerenzer and Gaissmaier (2011)).

Applying to the setting in this paper, define ω_1 as the event where there are more words that start with K and ω_1' as the event where there are more words with K as the third letter, and define ω_2 as the event where the position of the K letter affects the readiness of recall and ω_2' as the event where the position of the K letter does not affect recall probability. Define the state space as $\{\omega_1, \omega_1'\} \times \{\omega_2, \omega_2'\}$. In Tversky and Kahneman (1973), individuals ignore the states (ω_1, ω_2) and (ω_1', ω_2) as they infer whether ω_1 or ω_1' is true while implicitly fixing ω_2'. As I assume an one-to-one mapping from states to actions, I define ignorance as the behavior in which the DM never picks some of the N actions.

Definition 3. *An updating mechanism ignores state ω if the DM never picks action ω under all states of the world as $t \to \infty$, i.e., either $M^\omega = \emptyset$ or $\sum_{m\in M^\omega} \mu^{\omega'}_{Nm} = 0$ for all ω'.*[21]

[21] One may wonder why the DM would assign memory states to action ω if he never takes the action. This could happen for example when the optimal updating mechanism involves always picking action 1, e.g., when $u^1_N p^1_N$ is big enough and the information structures are very noisy. In that case, the DM can set $M^\omega = \{1\}$, $\mathscr{T}(1, s) = 1$ for all s and the initial state as memory state 1 to achieve the optimal (always choosing action 1), and the decision rule $d(m)$ for all $m \neq 1$ has no impact on the asymptotic utility loss.

It is important to note that I define ignorance based on the DM's actions instead of his "beliefs". It is because different from the Bayesian setting, the DM with bounded memory does not track his belief in all possible states but merely transit between the memory states in M. Theses memory states do not have to be associated with confidence levels of any of the N states. While one could interpret that the DM transits between a countable subset of beliefs in the N-dimensional probability simplex and always pays attention to his confident levels of all the N states, I believe that this interpretation goes against the concept of bounded memory as it requires unnecessary cognitive resources of the DM. In particular, as I assume an one-to-one mapping from states to optimal action, if the DM never chooses a particular action ω, he has no incentive to track his confidence level of state ω. Never choosing an action ω is thus effectively equivalent to never being aware of the possibility of state ω.

By definition, any updating mechanism is ignorant if $M < N$: if the DM lacks the cognitive resources to consider all possible states, he has to be ignorant. In the rest of the paper, I will focus on the more interesting case where $M \geq N$ and analyze whether an ignorant updating mechanism could be ϵ-optimal for small ϵ.

Does disagreement persist among different individuals? The third question relates to learning among heterogeneous individuals, in particular on whether they will agree asymptotically. As discussed after definition 3 of ignorance behaviors, it is tricky and arbitrary to define the distance of beliefs among different individuals with bounded memory. Thus, I define agreement and disagreement based on actions. Consider two individuals A and B with potentially different prior beliefs (p^ω_{NA}) and (p^ω_{NB}), different signals structures (f^ω_{NA}) and (f^ω_{NB}), or different levels of cognitive ability M_A, M_B, they agree with each other asymptotically if they choose the same action as $t \to \infty$:

Definition 4. *Two individuals A and B are bound (resp. almost bound) to agree with each other asymptotically if and only if their actions a^A_t and a^B_t satisfy*

$$\lim_{t\to\infty} Pr_N\{a^A_t = a^B_t \mid \omega\} = 1$$

for all ω (resp. almost all ω measured with both (p^ω_{NA}) and (p^ω_{NB})). They are bound (resp. almost bound) to disagree with each other asymptotically if and only if,

$$\lim_{t\to\infty} Pr_N\{a^A_t \neq a^B_t \mid \omega\} = 1.$$

for all ω (resp. almost all ω measured with both (p^ω_{NA}) and (p^ω_{NB})).

And if $\lim_{t\to\infty} \Pr_N\{a^A_t \neq a^B_t \mid \omega\} \in (0, 1)$, asymptotic disagreement happens probabilistically. I study the question on whether individuals with different prior beliefs but observe a long sequence of public information with eventually agree with each other, i.e., whether disagreement persists when individuals receive a large amount of public information. I also ask the question on whether two individuals who start with the same prior but receive many private information and have different abilities of information acquisition (f_N^ω), or individuals with levels of cognitive ability (M), will eventually disagree with each other in small and big worlds. The latter contrasts with the existing literature that explain asymptotic disagreement based on differences in prior beliefs or subjective information structures among different individuals.

3 Small World

I first analyze the optimal learning mechanisms in small worlds where $\frac{N}{M}$ is small. The following two propositions show that asymptotic learning behavior is close to Bayesian.

Proposition 1. *Fix a finite N and $M \to \infty$, asymptotic learning behavior is close to Bayesian, i.e.,* $\lim_{M\to\infty} L^*_{NM} = 0$.

Proposition 2. *Suppose $N, M \to \infty$ and $N = O(M^h)$ where $h < 1$, asymptotic learning behavior is close to Bayesian, i.e.,* $\lim_{N,M\to\infty} L^*_{NM} = 0$.

Proposition 1 and 2 show that in small world, (almost) optimal asymptotic learning mechanisms is very close to the Bayesian counterpart, i.e., the DM (almost) always takes the action matches with the true state. The proof of the two propositions, along with other proofs in the paper, are shown in appendix C. In the following, I will roughly describe the proof by showing a simple learning mechanism, illustrated in figure 3, achieves (almost) perfect asymptotic learning as M (or N, M) goes to ∞. The proposed simple mechanism tracks his favorable action and the corresponding confidence level over time. At any period t, the DM believes one of the N actions or no action is favorable, while the confidence level, if he has a favorable action, is an integer ranging from 1 to $\lfloor \frac{M-1}{N} \rfloor$. The memory states could thus be represented by $m_t \in \{0\} \cup \{1, \cdots, N\} \times \{1, \cdots, \lfloor \frac{M-1}{N} \rfloor\}$ where memory state 0 stands for no favorable action. The decision rule is such that he takes the favorable action if he has one, and randomly takes one of the N action with equal probability if he does not have a favorable action. The transition rule is described as follows.

First, the DM starts with no favorable action, i.e., the red memory state named "W" in figure 3.[22] If he receives a confirmatory signal for a state ω, he changes his favorable action to action ω with confidence level 1; if he receives signals that is not confirmatory for any states, he stays in the same memory state W that he has no favorable action.

[22]The starting memory state has no impact on the stationary distribution over the memory states and does not affect the asymptotic payoff.

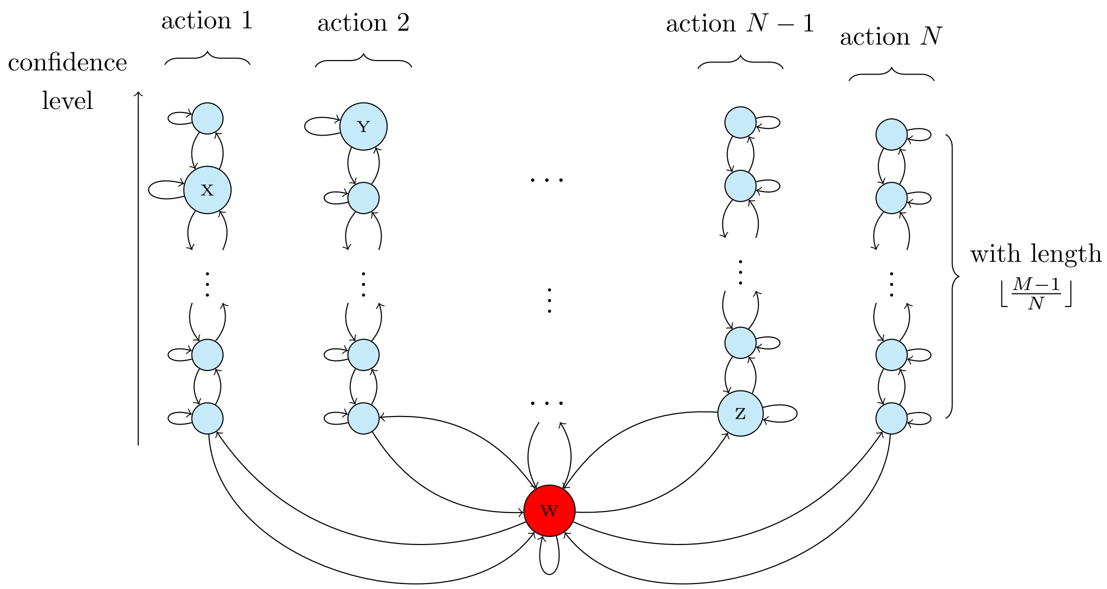

Figure 3: A simple updating mechanism that achieves perfect learning in small worlds.

Second, suppose at some period t the DM's favorable action is action ω with confidence level K. If $K < \lfloor \frac{M-1}{N} \rfloor$, e.g., in memory state X and Z, and he receives a confirmatory signal for state ω, he revises upwards his confidence level to $K+1$; if $K = \lfloor \frac{M-1}{N} \rfloor$, e.g., in memory state Y, and he receives a confirmatory signal for state ω , he stays in his current memory state with the same favorable action ω and the confidence level $K = \lfloor \frac{M-1}{N} \rfloor$. On the other hand, if $K > 1$, e.g., in memory state X and Y, and he receives a confirmatory signal for some state $\omega' \neq \omega$, he keeps action ω as his favorable action but revises downwards his confidence level to $K-1$ with some probability $\frac{1}{\delta} < 1$ and keep his confidence level at K otherwise;[23] if $K = 1$, e.g., in memory state Z, and he receives a confirmatory signal for some state $\omega' \neq \omega$, he transits to the red memory state W with no favorable actions with some probability $\frac{1}{\delta} < 1$ and otherwise stays in his current memory state with the same favorable action ω and confidence level $K = 1$. Lastly, if he receives signals that is not "confirmatory" for any states, he stays in his current memory state with favorable action ω and confidence level K.

The proof involves carefully defining the set of confirmatory signals for each state, such that it is more likely to receive a confirmatory signal for state ω than a confirmatory signal for any other states when the true state is ω. It also involves choosing a big enough δ such that it is more likely to adjust upwards than to adjust downwards the confidence level of action ω when the true state is ω. Crucially, with this simple mechanism, when $\frac{M}{N} \to \infty$, the upper bound of the confidence level goes to infinite which means that the DM can memorize infinitely many confirmatory signals for any state. Because asymptotically, the DM receives infinitely

[23] δ is constant across all K.

more confirmatory signals for the true state than the confirmatory signals for the any other states, as $t \to \infty$, the DM should almost surely memorize infinitely many confirmatory signals for the true state and thus almost surely learn perfectly the true state.

As the DM learns perfectly for all states of the world, intuitively he has no incentive to ignore any of the N states. Thus, ignorance behavior is not optimal in small worlds.

Corollary 1. *Fix a finite N, there exists some $\bar{M}$ such that if $M > \bar{M}$ (e.g., when $M \to \infty$), there exists some $\epsilon > 0$ such that no ignorant updating mechanism is ϵ-optimal.*

Corollary 2. *Suppose $N, M \to \infty$ and $N = O(M^h)$ where $h < 1$, a sequence of updating mechanism $\mathscr{T}_N$ is ϵ-optimal only if it ignores at most $\frac{\epsilon}{\underline{u}}$ measures of states at the limit.*

Note that I have showed a stronger result in corollary 1: not only when $M \to \infty$ but when M is big enough (or equivalently when $\frac{N}{M}$ is small enough), ignorant learning behavior is never optimal.

I now turn to the question on whether disagreement could persist asymptotically. Consider two individuals A and B who have different prior beliefs $(p^\omega_{NA})^N_{\omega=1}$ and $(p^\omega_{NB})^N_{\omega=1}$, or different abilities of information acquisition captured by $(f^\omega_{NA})^N_{\omega=1}$ and $(f^\omega_{NB})^N_{\omega=1}$.[24] As proposition 1 and 2 hold for all $(p^\omega_N)^N_{\omega=1}$ and $(f^\omega_N)^N_{\omega=1}$,[25] different individuals will be bound to choose the same action asymptotically if they adopt an ϵ-optimal updating mechanism with $\epsilon \to 0$.

Corollary 3. *Fix a finite N and $M \to \infty$, different individuals with different prior beliefs and/or information acquisition abilities are bound to agree asymptotically in small worlds. That is, for all $(u^\omega_{NA}, p^\omega_{NA}, f^\omega_{NA})^N_{\omega=1}$ and $(u^\omega_{NB}, p^\omega_{NB}, f^\omega_{NB})^N_{\omega=1}$, fixing N and $M \to \infty$, if the two individuals adopt ϵ-optimal mechanisms with $\epsilon \to 0$,*

$$\lim_{M\to\infty} \lim_{\epsilon\to 0} \lim_{t\to\infty} Pr_N\{a^A_t = a^B_t \mid \omega\} = 1 \text{ for all } \omega.$$

Corollary 4. *Suppose $N, M \to \infty$ and $N = O(M^h)$ where $h < 1$, different individuals with different prior beliefs and/or information acquisition abilities are almost bound to agree asymptotically in small worlds as long as they agree on 0 probability events. That is, for all $(u^\omega_{NA}, p^\omega_{NA}, f^\omega_{NA})^N_{\omega=1}$ and $(u^\omega_{NB}, p^\omega_{NB}, f^\omega_{NB})^N_{\omega=1}$ such that $\lim_{N\to\infty} p^\omega_{NA} > 0$ if and only if $\lim_{N\to\infty} p^\omega_{NB} > 0$ for all ω, if the two individuals*

[24] For example, individual A could receive noisier signals than individual B, i.e., $f^\omega_{NA} = \gamma + (1-\gamma) f^\omega_{NB}$ for some $\gamma \in (0,1)$; or individual A could have different learning advantages, identifying some states better but other states worse than individual B, i.e., $\sup_s f^\omega_{NA}(s)/f^{\omega'}_{NA}(s) > \sup_s f^\omega_{NB}(s)/f^{\omega'}_{NB}(s)$ but $\sup_s f^{\omega''}_{NA}(s)/f^{\omega'''}_{NA}(s) < \sup_s f^{\omega''}_{NB}(s)/f^{\omega'''}_{NB}(s)$ for some $\omega, \omega', \omega'', \omega'''$.

[25] Note that I do not assume that individuals have the "correct" prior beliefs. As long as their prior beliefs satisfy the full support assumption 2 or 2', the results in this paper would hold.

adopt ϵ-optimal mechanisms with $\epsilon \to 0$,

$$\lim_{N,M\to\infty}\lim_{\epsilon\to 0}\sum_{\omega}\left[\mathbb{1}\left\{\lim_{t\to\infty} Pr_N\{a_t^A = a_t^B \mid \omega\} = 1\right\} p_I^{\omega}\right] = 1 \textit{ for } I = A, B.$$

Thus, corollary 3 and 4 show that different individuals with different priors and/or information acquisition abilities who adopt (almost) optimal updating mechanism are bound to agree with each other if they receive a large amount of (public or private) information.

It is important and interesting to note that the assumption that the DM is able design and adopt an optimal updating mechanism is not as far-stretched as one may think. Put it differently, there is a large set of updating mechanisms that would achieve $\lim_{\frac{M}{N}\to\infty} L_N(\mathscr{T}) = 0$ and guarantee asymptotic agreement. I illustrate this "robustness" result in the appendix C.7. Roughly speaking, I consider the aforementioned simple updating mechanism illustrated in figure 3 but assume that the DM mistakenly transits to a neighboring memory state with some probability γ in each period regardless of the signal realization s_t: in each period if the DM has a confidence level K where $\lfloor\frac{M-1}{N}\rfloor > K > 1$, he adjusts upwards or downwards one unit of his confidence level mistakenly with equal probability $\frac{\gamma}{2}$; if $K = \lfloor\frac{M-1}{N}\rfloor$, he adjusts downwards one unit of his confidence level mistakenly with probability γ; if $K = 1$, he adjusts upwards one unit of his confidence level or transit to no favorable action (the red memory state) with equal probability $\frac{\gamma}{2}$; if he has no favorable action, he changes his favorable action to action ω with confidence level 1 with equal probability $\frac{\gamma}{N}$ or all ω; with probability $1-\gamma$, the DM follows the simple updating mechanism illustrated in figure 3.

Such (local) mistake could be induced by mistakes in the perception of signals or imperfect tracking (local fluctuation) of memory states. Appendix C.7 show that the results of (almost) perfect learning, i.e., proposition 1 and 2, and asymtotic agreement, i.e., corollary 3 and 4, hold for all $\gamma \in [0, 1)$. Local mistake, however likely it is, does not break down the results of (almost) perfect learning and asymptotic agreement in small worlds.

4 Big world

Now I proceed to the analysis of the big world and show that the three implications in the small world, i.e., asymptotic learning is close to Bayesian, ignorance is never optimal, disagreement does not persist, do not hold. Before I present the results, it would be useful to discuss a simple example of $N = M = 2$, to understand how bounded memory affects asymptotic learning and to introduce variables that capture important features of an updating mechanism.

An example of $\mathbf{N = M = 2}$. Suppose $M^1 = \{1\}$ and $M^2 = \{2\}$, i.e., the DM takes action 1 in memory state 1 and action 2 in memory state 2. An important feature of the updating mechanism is the state likelihood ratios in the memory states that measures how likely the DM will be in a memory state m under state ω vs that under state ω'.

Definition 5. *The state $\omega - \omega'$ likelihood ratio at memory state m is defined as* $\frac{\mu^{\omega}_{Nm}}{\mu^{\omega'}_{Nm}}$.

The higher the state $\omega - \omega'$ likelihood ratio at memory state m is, given that the DM is in memory state m, the more confident he is that the true state is ω instead of ω'. In this simple example, a good updating mechanism induces a high $\frac{\mu^1_{21}}{\mu^2_{21}}$ and a low $\frac{\mu^1_{22}}{\mu^2_{22}}$. In particular, perfect learning requires that the state $1 - 2$ likelihood ratio goes to infinity at memory state 1 but equals 0 at memory state 2. In other words, the DM has to be almost sure that state 1 is true when he takes action 1 and almost sure that state 2 is true when he takes action 2. However, as shown in Hellman and Cover (1970), the ratio of the likelihood ratio between two memory states is constrained by the information structures and the bounded memory. Intuitively, to be almost sure about state ω, the DM has to have the ability to record/memorize almost perfect information supporting state ω, either through recording one (almost) perfect confirmatory signal or infinitely many imperfect confirmatory signal supporting state ω. However, in the current setting the former is constrained by the information structure, i.e., assumption 3, and the latter is constrained by the bounded memory. To illustrate this constraint, define the state $\omega - \omega'$ spread as:

Definition 6. *Denote the state $\omega - \omega'$ spread as $\Upsilon^{\omega\omega'}_N$ which is given by the following equation:*

$$\Upsilon^{\omega\omega'}_N = \frac{\max_{m\in M^{\omega}} \frac{\mu^{\omega}_{Nm}}{\mu^{\omega'}_{Nm}}}{\min_{m\in M^{\omega'}} \frac{\mu^{\omega}_{Nm}}{\mu^{\omega'}_{Nm}}}$$

In this simple example, $\Upsilon^{12}_2 = \frac{\mu^1_{21}}{\mu^2_{21}}/\frac{\mu^1_{22}}{\mu^2_{22}}$. I will show it the following that Υ^{12}_2 is bounded above, such that the DM can never be sure both when he takes action ω and ω'. In an irreducible automaton, i.e., $\mu^{\omega}_{Nm} > 0$ for all m, the probability mass moving from $m = 1$ to $m = 2$ must be equal to the probability mass moving to the opposite direction. Suppose the DM updates his belief from $m = 1$ to $m = 2$ given some signals S_2 and updates in the opposite direction given some signals S_1, we have in the stationary distribution

$$\mu^1_{21}F^1_2(S_2) = \mu^1_{22}F^1_2(S_1)$$
$$\mu^2_{21}F^2_2(S_2) = \mu^2_{22}F^2_2(S_1)$$

Thus,

$$\Upsilon^{12}_2 = \frac{\mu^1_{21}}{\mu^2_{21}}/\frac{\mu^1_{22}}{\mu^2_{22}} = \frac{F^1_2(S_1)}{F^2_2(S_1)}/\frac{F^1_2(S_2)}{F^2_2(S_2)} \le \bar{l}^{12}_2\bar{l}^{21}_2 \tag{7}$$

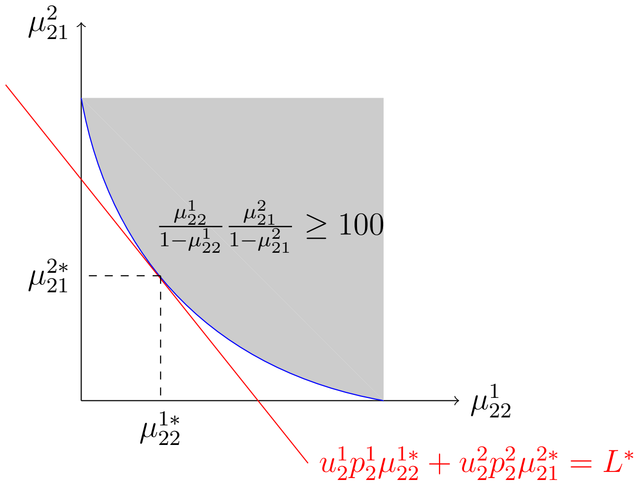

Figure 4: This example illustrates the trade-off between inference of the two states, where $\bar{l}_2^{12}\bar{l}_2^{21} = 100$. The feasibility set of μ_{22}^1 and μ_{21}^2 is shown as the gray area (which may not include the boundary) and the infimum of utility loss is given by the point where the objective function touches the feasibility set.

where $\bar{l}_N^{\omega\omega'} = \sup_{F_N^{\omega'}(S')>0} \frac{F_N^{\omega}(S')}{F^{\omega'}(S')}$. This bound on the state $1-2$ spread, imposed by the (maximum) informativeness of the signals and the bounded memory, induces a trade-off between the inference the two states. Suppose $\bar{l}_2^{12}\bar{l}_2^{21} = 100$. If the DM wants to design an automaton such that he chooses action 1 with probability 99% in state 1 , then

$$\frac{99}{\mu_{21}^2} / \frac{1}{\mu_{22}^2} \leq 100$$
$$\frac{\mu_{21}^2}{\mu_{22}^2} \geq \frac{99}{100},$$
$$\mu_{21}^2 \geq \frac{99}{199},$$

i.e., the DM has to make mistake in state 2 (chooses action 1) more than almost half of the time. If however he decreases his quality of decision making in state 1 such that he chooses action 1 in state 1 with probability 90%, then $\mu_{21}^2 \geq \frac{9}{109}$ and he could make mistake in state 2 with probability as low as 8.3%.

How much the DM would trade-off between the inference of the two states depend on the prior attractiveness of the two actions: $u_2^1 p_2^1$ vs $u_2^2 p_2^2$. When $u_2^1 p_2^1$ is bigger than $u_2^2 p_2^2$, the DM is willing to sacrifice inference in state 2 to improve inference in state 1. The trade-off is illustrated in figure 4. As I will show later, this trade-off implies that ignorant learning behavior could be optimal.

As shown in Hellman and Cover (1970), when M increases, the upper bound on Υ_2^{12} increases and the DM can sacrifice less on inference in one state to improve inference in another state. In particular, when $M \to \infty$, the bound goes to infinity and the DM can achieve almost perfect learning in both states, as previously shown in the cases of small worlds. In contrast, this paper shows that what matter is not the absolute value of M, but the fraction $\frac{M}{N}$. ♦

The following two propositions show that in the big world asymptotic learning is imperfect, and thus significantly different from Bayesian. It thus generalizes the aforementioned result in Hellman and Cover (1970) with $N = 2$, and more importantly, shows that what matter is the ratio $\frac{M}{N}$ rather than the absolute value of M. It implies that fixing the cognitive resources of individuals, learning behaviors in a complicated problem are more likely to differ from Bayesian learning compared to that in a simple problem.

Proposition 3. *Suppose both N and M are finite, asymptotic learning differs significantly from Bayesian learning, i.e., $L^*_{NM} > 0$. Moreover, fixing N, L^*_{NM} decreases in M, i.e., asymptotic learning becomes closer to Bayesian learning as M increases.*

Proposition 4. *Suppose $N, M \to \infty$ and $N = O(M^h)$ where $h \geq 1$, asymptotic learning differs significantly from Bayesian learning, i.e., $\lim_{N,M\to\infty} L^*_{NM} > 0$.*

Comparing proposition 3 and 4 with proposition 1 and 2 gives us the first difference of learning behavior in small and big worlds. In small world problems, asymptotic learning could be well approximated by Bayesian updating; while in big worlds, asymptotic learning behavior significantly differs from Bayesian updating. Moreover, the second part of proposition 3 implies that fixing N, learning behavior gets closer to the Bayesian benchmark when M increases or equivalently when the relative complexity $\frac{N}{M}$ decreases.[26]

To understand the intuition of proposition 3 and 4, it would be useful to revisit the simple updating mechanism proposed in the small world (see figure 3). Different from the case of small worlds, in big worlds, the range of confidence levels are bounded above for all (or almost all measured in fraction when $N, M \to \infty$) states of the world as $\frac{M}{N} < \infty$, as illustrated in figure 5. As $\frac{M}{N}$ is bounded above, no (or negligible fraction of) actions could be allocated with infinite number of memory states. As a result, in states of the world where confidence level is bounded above, the DM is not able to record infinite (imperfect) signals supporting that state against the other and thus will not be almost sure about his action. As the DM is never almost sure about choosing action ω, it implies that there is a positive probability that he chooses action ω in other states $\omega' \neq \omega$, which leads to utility loss. Thus $L^*_{NM} > 0$.

The second difference between small and big worlds concerns the optimality of ignorance learning behaviors. In a big world, as mentioned above, the DM cannot allocate infinite cognitive resources to all states of the world. As the DM is bound to make mistakes when he take actions, he thus faces meaningful trade-off in the

[26]Note that it is difficult to directly compute the comparative statics with respect to $\frac{M}{N}$ because when N changes, prior beliefs and the signal structures, i.e., the nature of the inference problem, also change.

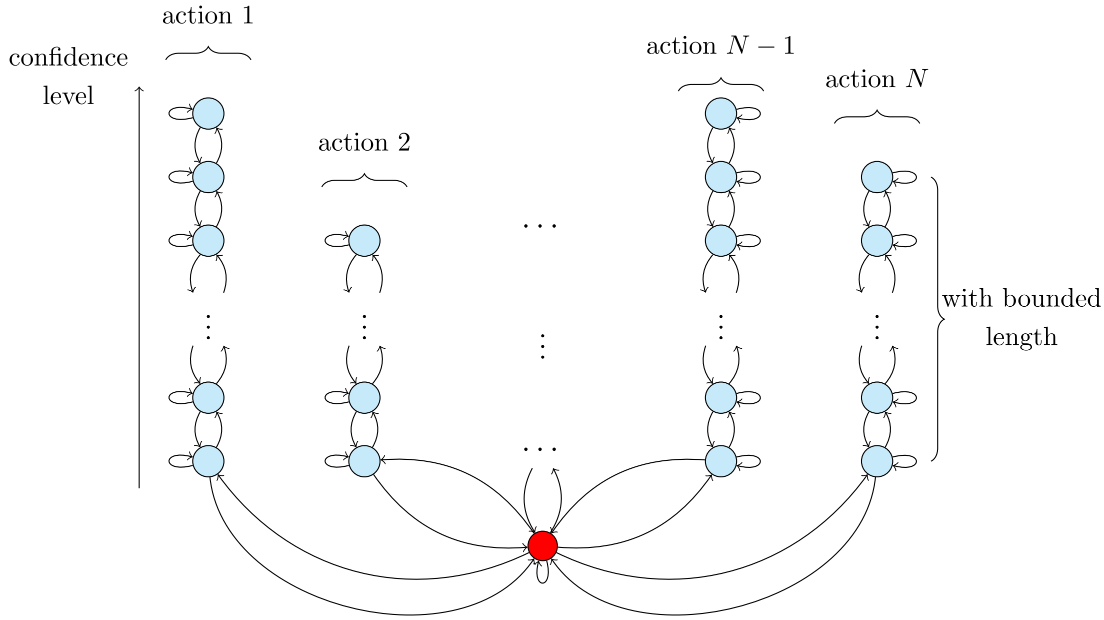

Figure 5: The simple updating mechanism, revisited, in a big world and with different range of confidence level across different actions. In particular, confidence level of all actions $1, 2, \cdots, N$ are bounded above.

allocation of cognitive resources as he trades off between the probability of mistakes in different states of the world.[27] Intuitively, when the prior probability of a state is very low, the DM would rather allocate cognitive resources to infer about other states of the world and ignore that à priori unlikely state.[28] The following two propositions show that such ignorance behavior could indeed be optimal in big worlds.

Proposition 5. *Suppose N and M are finite. There exists some $\xi > 0$ such that when $p_N^\omega < \xi$ and $\epsilon \to 0$, all ϵ-optimal updating mechanisms ignore state ω, i.e., $\lim_{\epsilon\to 0} M^\omega = \emptyset$ or $\lim_{\epsilon\to 0} \sum_{m\in M^\omega} \mu_{Nm}^{\omega'} = 0$ for all $\omega' \in \Omega$.*

On the other hand, suppose for some $(u_N^{\omega'}, p_N^{\omega'})_{\omega'=1}^N$, when $\epsilon \to 0$, all ϵ-optimal updating mechanisms ignore state ω. Then for $(\tilde{u}_N^{\omega'}, \tilde{p}_N^{\omega'})_{\omega'=1}^N$ such that

$$\frac{\tilde{u}_N^{\omega'}\tilde{p}_N^{\omega'}}{\tilde{u}_N^{\omega''}\tilde{p}_N^{\omega''}} = \frac{u_N^{\omega'}p_N^{\omega'}}{u_N^{\omega''}p_N^{\omega''}} \text{ for all } \omega', \omega'' \neq \omega;$$
$$\frac{\tilde{u}_N^{\omega}\tilde{p}_N^{\omega}}{\tilde{u}_N^{\omega'}\tilde{p}_N^{\omega'}} < \frac{u_N^{\omega}p_N^{\omega}}{u_N^{\omega'}p_N^{\omega'}} \text{ for all } \omega' \neq \omega,$$

when $\epsilon \to 0$, all ϵ-optimal updating mechanisms ignore state ω.

[27] In small world, the DM can allocate infinite memory states to all actions. Adding or taking away a finite number of memory states for each action thus does not affect stationary distribution and utility. Thus there is no meaningful trade-off in small worlds.

[28] Note that the trade-off does not only happen in the allocation of memory states to M^ω for different ω, but also in choosing the asymptotic probability of taking different actions. More specifically, a lower probability of choosing action ω, $\sum_{m\in M^\omega} \mu_{Nm}^{\omega'} = 0$, implies a higher probability of choosing other actions.

Proposition 6. *Suppose $N, M \to \infty$ and $N = O(M^h)$ where $h \geq 1$, all updating mechanisms must ignore almost all (measured in fraction) states.*

Note that the statements in proposition 5 and 6 are different. Proposition 5 says that there exists some prior beliefs such that the DM with those prior beliefs will ignore some states when he adopts an (almost) optimal updating mechanism. In contrast, proposition 6 is "stronger": all learning mechanisms, including the (almost) optimal updating mechanism have to be ignorant. The intuition of the former has been briefly mentioned before the proposition: when the DM cannot allocate infinite cognitive resources and achieve perfect learning in all states of the world, he faces meaningful trade-off between learning across different states of the world. In that case, it might be efficient to focus on a strict subset of the states of the world, i.e., to adopt an ignorant learning mechanism. On the other hand, roughly speaking, the intuition of proposition 6 is as follows: when $M \to \infty$, each probability on each memory state become infinitesimally small. As the DM cannot allocate infinite memory state in almost all M^ω, i.e., the DM take each action ω in a finite subset of memory states, he must picks almost all actions with 0 probability.

Nonetheless, comparing proposition 5 and 6 with corollary 3 and 4 show that ignorance behavior could be optimal only in big worlds but not in small worlds. As will be discussed in details in next section, this explains inattentive/neglecting/heuristics learning behavior documented in Tversky and Kahneman (1973), Enke and Zimmermann (2019) and Graeber (2019), and provides micro-foundation to the equilibrium concepts proposed in Eyster and Rabin (2005) and Jehiel (2005). Importantly, it also speaks to in what circumstances these behavioral abnormalities are likely to be observed, i.e., in big worlds where the relative complexity of inference problems is high.

As different individuals with different prior beliefs and utility matrix would adopt different updating mechanism, they could focus their learning on, or ignore, different subsets of states of the world and thus disagree asymptotically. Consider two individuals with different utility functions and prior beliefs $(u^\omega_{NA}, p^\omega_{NA})^N_{\omega=1}$ and $(u^\omega_{NB}, p^\omega_{NB})^N_{\omega=1}$, but receive a long sequence of public signal where $(f^\omega_{NA})^N_{\omega=1} = (f^\omega_{NB})^N_{\omega=1}$ for all ω, or two individuals with the same utility functions and prior beliefs, but receive private signals that are generated by different signal structures, or two individuals with different utility functions, prior beliefs and signals structures, etc., the following two corollaries show that they might disagree with each other with certainty as $t \to \infty$ even when they adopt an (almost) optimal updating mechanism.

Corollary 5. *Suppose N and M are finite. There exists some $(u^\omega_{NA}, p^\omega_{NA}, f^\omega_{NA})^N_{\omega=1}$ and $(u^\omega_{NB}, p^\omega_{NB}, f^\omega_{NB})^N_{\omega=1}$ such that the two individuals are bound to disagree when* *they adopt ϵ-optimal updating mechanism with ϵ small enough:*

$$\lim_{\epsilon\to 0}\lim_{t\to\infty} Pr_N(a_t^A \neq a_t^B \mid \omega) = 1 \text{ for all } \omega.$$

Corollary 6. *Suppose $N, M \to \infty$ and $N = O(M^h)$ where $h \geq 1$. For all $\epsilon > 0$, there exists some $(u_{NA}^\omega, p_{NA}^\omega, f_{NA}^\omega)_{\omega=1}^N$ and $(u_{NB}^\omega, p_{NB}^\omega, f_{NB}^\omega)_{\omega=1}^N$ and ϵ-optimal updating mechanisms $(\mathscr{T}_{NA}, d_{NA})$ and $(\mathscr{T}_{NB}, d_{NB})$ such that the two individuals adopting $(\mathscr{T}_{NA}, d_{NA})$ and $(\mathscr{T}_{NB}, d_{NB})$ are bound to disagree:*[29]

$$\lim_{N,M\to\infty}\lim_{t\to\infty} Pr_N(a_t^A \neq a_t^B \mid \omega) = 1 \text{ for all } \omega.$$

It is important to note that the disagreement are certain in the two corollaries, instead of probabilistic: the two individuals would never agree with each other as they would focus on learning some disjoint sets of states. Again, comparing corollary 3 and 4 with corollary 5 and 6, we can conclude that asymptotic disagreement only happens in big worlds but not in small worlds.

4.1 Disagreement driven by differences in cognitive ability

Corollary 5 and 6 show that individuals could be bound to disagree asymptotically, if they have different utility functions, priors beliefs and/or information structures. This subsection shows that disagreement could also be driven by differences in cognitive ability. I illustrate the result in the following example. Consider a setting with $N = 3$ and two individuals, A and B, who share the same prior beliefs and the same objective signal structure:

$$\begin{aligned} p_3^1 &= \frac{1}{3} + 2\nu \\ p_3^2 &= \frac{1}{3} - \nu \\ p_3^3 &= \frac{1}{3} - \nu \end{aligned} \tag{8}$$

$$\begin{aligned} \sup_s \frac{f_3^1(s)}{f_3^n(s)} &= \sup_s \frac{f_3^n(s)}{f_3^1(s)} = \sqrt{1+\tau} \text{ for } n = 2, 3 \\ \sup_s \frac{f_3^2(s)}{f_3^3(s)} &= \sup_s \frac{f_3^3(s)}{f_3^2(s)} = \sqrt{1+\Upsilon} \text{ where } \Upsilon > \tau. \end{aligned}$$

with $1 + \tau \geq \frac{\frac{1}{3}+2\nu}{\frac{1}{3}-\nu}$.[30] Moreover, to simplify things, assume $u_3^1 = u_3^2 = u_3^3 = 1$. The only difference the two individuals have is their levels of cognitive ability, in which individual A has $M = 1$ and individual B has $M = 2$. I present the following result.

[29] Note that corollary 6 is weaker than corollary 5 in a sense that there exists some, but not all, ϵ-optimal updating mechanisms that lead to disagreement.

[30] It ensures that if $M \geq 2$, the DM never pick action 1 with probability 1 and he can achieve a lower utility loss compared to the benchmark of no information.

Proposition 7. *There exists ν, τ, Υ such that individual A and B adopting ϵ-optimal mechanisms are bound to disagree for small ϵ, i.e.,*

$$\exists \nu, \tau, \Upsilon, \tilde{\epsilon} > 0 \text{ such that } \lim_{t\to\infty} Pr_3\{a_t^A \neq a_t^B \mid \omega\} = 1 \text{ for all } \omega \text{ and } \epsilon < \tilde{\epsilon}.$$

The intuition is as follows: individual A obviously always choose action 1 as he has not enough cognitive resources to learn. On the other hand, when ν is small enough, and when Υ is much bigger than τ, it is more beneficial individual B to focus on learning state 2 and 3 as the signals supporting the two states are more informative. In this case, he never chooses action 1 and thus never agrees with individual A.[31]

Proposition 7 shows that even when individuals start from the same prior and has the same objective signal structures, they can be bound to disagree with each other after receiving a large amount of public information.[32] This result proposes a different channel of asymptotic disagreement that is in contrast with the existing explanations that assume different individuals have different prior beliefs or perceive signals in different ways.

5 Discussion and Conclusion

Limited ability and behavioral abnormalities This paper contributes to a growing set of theoretical literature that explains behavioral abnormalities as optimal or efficient strategies assuming individuals have limited ability. Wilson (2014) shows that in the case where $N = 2$, individuals with bounded memory exhibits confirmation bias; Steiner and Stewart (2016) shows that an optimal response to noises in perceiving details of lotteries leads to phenomenon of probability weighting in prospect theory (Kahneman and Tversky (1979)); Jehiel and Steiner (2019) and Leung (2020)) shows that a capacity constraint in belief updating drives confirmation bias and other biases in belief formation.

Different from the literature, this paper explains a larger set of behavioral abnormalities including ignorant learning behaviors, departures from Bayesian learning and disagreement among individuals. In particular, different phenomena, such as use of heuristics, correlation neglect, inattentive learning, persistent over-confidence and other model misspecification can be modeled as ignorant learning behavior in

[31]One may argue that after seeing individual B choosing action 2 or 3, individual A should change his action. Note that however this is not possible as he has only one unit of memory capacity $M = 1$ and thus have to effectively committed to one action. In particular, one can generate this framework to which the two individuals also see each others' action as signals and proposition 7 would still hold.

[32]Note that although this example imposes strong assumptions in particular on the size of bounded memory of individual A, it generates a strong form of disagreement in which the two individuals disagree asymptotically with certainty. Similar intuition implies that even when the assumption is relaxed, difference in M would lead to asymptotic disagreement at least probabilistically.

this setting.[33] Moreover, this paper also connects the relative complexity of inference problems to behavioral abnormalities and derives testable predictions on their prominence: the prominence of these behavioral abnormalities should increases in the relative complexity of the inference problem.

Asymptotic disagreement/polarization Next, the results on asymptotic disagreement contributes to the large literature that explains the phenomenon. In the existing literature, asymptotic disagreement is driven by differences in signal distributions across state or exogenous learning mechanisms (Morris (1994), Mailath and Samuelson (2020), Gilboa et al. (2020)), the lack of identification (or uncertainty in signal distributions) (Acemoglu et al. (2016), confirmation bias (Rabin and Schrag (1999)), or model misspecification (Berk (1966), Freedman (1963, 1965)).

Differently, this paper looks into the connection between limited ability and disagreement, and show when asymptotic disagreement could arise and when it will not happen, depending on the relative complexity of the inference problem. Moreover, proposition 7 shows that disagreement could arise solely because of differences in cognitive abilities, even when two individuals share the same prior without model misspecification, perceive signals in the same way with objective signal structures across states, adopt an (almost) optimal updating mechanism and observe a large amount of public information. To the best of my knowledge, this novel mechanism that disagreement can be driven by differences in cognitive ability has not been studied in the theoretical literature.

Heuristic(ignorance) learning As previously mentioned, the results of ignorance behavior explain the inattentive/neglecting/heuristic inference behavior documented in Tversky and Kahneman (1973), Enke and Zimmermann (2019) and Graeber (2019), and micro-found those modeled in the concept of cursed equilibrium (Eyster and Rabin (2005)), analogy-based equilibrium (Jehiel (2005)), selection neglect (Jehiel (2018)) or persistent over-confidence (Heidhues et al. (2018)). Importantly, the comparison of small and big worlds illustrates a link between the (relative) complexity of the inference problem and the ignorance behavior, which is supported by results in Enke and Zimmermann (2019) and Graeber (2019). The former shows in section 2.4.3 that inattentive learning negatively correlates with the cognitive ability of subjects and in section 3.1 that "an extreme reduction in the environment's complexity eliminates the bias", while the latter shows that a reduction in the complexity of the problem by removing a decipher stage of signals reduces inattentive learning behavior.

Interestingly, Enke and Zimmermann (2019) and Graeber (2019) also shows that simply reminding subjects about the neglected variables reduces inattentive

[33] The optimality/efficiency of ignorant learning behaviors is also studied in the psychology literature, notably in Gigerenzer and Goldstein (1996) and Gigerenzer and Gaissmaier (2011).

learning and improves inference. It seems to contradict the result in this paper that shows inattentive learning could be optimal. However, this "reminder effect" could be interpreted in the current setup via a change in the state space. Consider the behavior of inattentive inference in Graeber (2019). Denote the variable that subjects are asked to infer as A and the ignored variable as B. Before being reminded about the ignored variable, the state state is $\text{supp}(A) \times \text{supp}(B) \times \{B \text{ affects the signal distribution}, B \text{ does not affect the signal distribution}\}$, in which subjects might ignore the states that say "B affects the signal distribution", i.e., they adopt an ignorance learning mechanism. After being reminded about the effect of B, the set of states of the world reduces effectively to $\text{supp}(A) \times \text{supp}(B) \times \{B \text{ affects the signal distribution}\}$ and subjects adopt another learning mechanism that does not involve ignorance learning.

The mechanism mentioned in the previous paragraph thus brings forth an open question that is not answered in this paper. In reality, individuals face different (set of) inference problems and are likely endowed with different learning mechanisms for different sets of states of the world. Like in the example mentioned in the last paragraph, some states space could be nested in another, and individuals could transit from one learning mechanism to another given some information that triggers him to revise the states space. This is also related to the question of how individuals construct the state space given an inference problem. Arguably, there are infinitely many variables that might affect the signal distributions, thus their realizations could be incorporated in the set of possible states. Roughly speaking, the result of ignorance seems to suggest that individuals may only include the most "important" or "à priori probable" states, while the "reminder effect" suggests that the construct of the state space also depends on the information received by the individual. Moving forward, I believe that the question of how individuals construct their perceived state space deserves more in depth and careful analysis as it is fundamental to individuals' learning behavior.

References

Acemoglu, Daron, Victor Chernozhukov, and Muhamet Yildiz (2016) "Fragility of asymptotic agreement under Bayesian learning," *Theoretical Economics*, Vol. 11, No. 1, pp. 187–225.

Basu, Pathikrit and Kalyan Chatterjee (2015) "On interim rationality, belief formation and learning in decision problems with bounded memory," *Working Paper.*

Berk, Robert H (1966) "Limiting behavior of posterior distributions when the model is incorrect," *The Annals of Mathematical Statistics*, pp. 51–58.

Blackwell, David and Lester Dubins (1962) "Merging of opinions with increasing information," *The Annals of Mathematical Statistics*, Vol. 33, No. 3, pp. 882–886.

Compte, Olivier and Andrew Postlewaite (2012) "Belief formation," *Unpublished manuscript, University of Pennsylvania.*

Enke, Benjamin and Florian Zimmermann (2019) "Correlation neglect in belief formation," *The Review of Economic Studies*, Vol. 86, No. 1, pp. 313–332.

Eyster, Erik and Matthew Rabin (2005) "Cursed equilibrium," *Econometrica*, Vol. 73, No. 5, pp. 1623–1672.

Freedman, David A (1963) "On the asymptotic behavior of Bayes' estimates in the discrete case I," *The Annals of Mathematical Statistics*, pp. 1386–1403.

——— (1965) "On the asymptotic behavior of Bayes estimates in the discrete case II," *The Annals of Mathematical Statistics*, Vol. 36, No. 2, pp. 454–456.

Gigerenzer, Gerd and Henry Brighton (2009) "Homo heuristicus: Why biased minds make better inferences," *Topics in cognitive science*, Vol. 1, No. 1, pp. 107–143.

Gigerenzer, Gerd and Wolfgang Gaissmaier (2011) "Heuristic decision making," *Annual review of psychology*, Vol. 62, pp. 451–482.

Gigerenzer, Gerd and Daniel G Goldstein (1996) "Reasoning the fast and frugal way: models of bounded rationality.," *Psychological review*, Vol. 103, No. 4, p. 650.

Gilboa, Itzhak, Larry Samuelson, and David Schmeidler (2020) "Learning (to Disagree?) in Large Worlds," *Working Paper.*

Graeber, Thomas (2019) "Inattentive inference," *Working Paper.*

Heidhues, Paul, Botond Kőszegi, and Philipp Strack (2018) "Unrealistic expectations and misguided learning," *Econometrica*, Vol. 86, No. 4, pp. 1159–1214.

Hellman, Martin E and Thomas M Cover (1970) "Learning with finite memory," *The Annals of Mathematical Statistics*, pp. 765–782.

Hellman, Martin Edward (1969) *Learning with finite memory* Ph.D. dissertation, Standford University.

Hoffman, Mitchell and Stephen V Burks (2017) "Worker overconfidence: Field evidence and implications for employee turnover and returns from training," Technical report, National Bureau of Economic Research.

Jehiel, Philippe (2005) "Analogy-based expectation equilibrium," *Journal of Economic theory*, Vol. 123, No. 2, pp. 81–104.

——— (2018) "Investment strategy and selection bias: An equilibrium perspective on overoptimism," *American Economic Review*, Vol. 108, No. 6, pp. 1582–97.

Jehiel, Philippe and Jakub Steiner (2019) "Selective sampling with information-storage constraints," *The Economic Journal*, DOI: https://doi.org/10.1093/ej/uez068.

Kahneman, Daniel (2011) *Thinking, fast and slow*: Macmillan.

Kahneman, Daniel, Paul Slovic, and Amos Tversky (1982) *Judgment under uncertainty: Heuristics and biases*: Cambridge university press.

Kahneman, Daniel and Amos Tversky (1979) "Prospect Theory: An Analysis of Decision under Risk," *Econometrica*, Vol. 47, No. 2, pp. 263–292.

Leung, Benson Tsz Kin (2020) "Limited cognitive ability and selective information processing," *Games and Economic Behavior*, Vol. 120, pp. 345–369.

Mailath, George J and Larry Samuelson (2020) "Learning under Diverse World Views: Model-Based Inference," *American Economic Review*, Vol. 110, No. 5, pp. 1464–1501.

Monte, Daniel and Maher Said (2014) "The value of (bounded) memory in a changing world," *Economic Theory*, Vol. 56, No. 1, pp. 59–82.

Morris, Stephen (1994) "Trade with heterogeneous prior beliefs and asymmetric information," *Econometrica*, pp. 1327–1347.

Rabin, Matthew (1998) "Psychology and economics," *Journal of Economic Literature*, Vol. 36, No. 1, pp. 11–46.

Rabin, Matthew and Joel L Schrag (1999) "First impressions matter: A model of confirmatory bias," *The Quarterly Journal of Economics*, Vol. 114, No. 1, pp. 37–82.

Savage, Leonard J (1972) *The foundations of statistics*: Courier Corporation.

Steiner, Jakub and Colin Stewart (2016) "Perceiving prospects properly," *American Economic Review*, Vol. 106, No. 7, pp. 1601–31.

Tversky, Amos and Daniel Kahneman (1973) "Availability: A heuristic for judging frequency and probability," *Cognitive psychology*, Vol. 5, No. 2, pp. 207–232.

Wilson, Andrea (2014) "Bounded memory and biases in information processing," *Econometrica*, Vol. 82, No. 6, pp. 2257–2294.

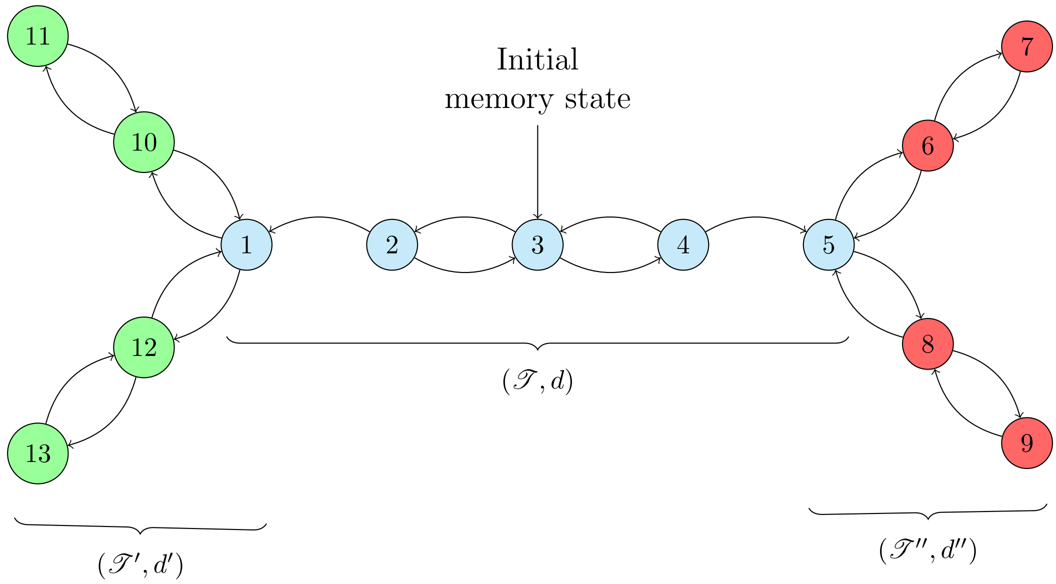

Figure 6: An example of a 13-memory states automaton that involving switching between three 5-memory states automatons. The DM starts at memory state 3 with $(\mathscr{T}, d)$, and switch to $(\mathscr{T}', d')$ and $(\mathscr{T}'', d'')$ once he transits to memory state 1 and 5 respectively.

A Switching between different automatons

This section illustrates how the setup in this paper allows the DM to switch between automatons with size smaller than M.

I first argue that assuming the DM could switch between multiple learning mechanism with M memory states implicitly implies that he has a larger memory capacity than M. Suppose a DM starts with $(\mathscr{T}, d)$ at memory state m_0 and switch to $(\mathscr{T}', d')$ and $(\mathscr{T}'', d'')$ once he transit to memory state m_1 and m_2 respectively. As $(\mathscr{T}, d) \neq (\mathscr{T}', d') \neq (\mathscr{T}'', d'')$, when the DM receives a signal s at memory state m and decides to which memory state he transit to, or when he decides which action he takes at memory state m, he has to remember whether he has once transited to memory state m_1 and m_2. In other words, the DM has to track not only his current memory states, but have to memorize the (incomplete) history of his previous memory states. Switching between multiple automatons thus implicitly implies a larger memory capacity.

Now I illustrate by an example that a M memory states automaton could be designed to involve switching between automatons with smaller sizes. The example is illustrated in figure 6. In the example, the DM starts at memory state 3 with a learning mechanism $(\mathscr{T}, d)$ that involves 5 memory states 1 to 5. Once he transit to memory state 1, he switch to another learning mechanism $(\mathscr{T}', d')$ that involves 5 memory states 1, $10, 11, 12, 13$. On the other hand, once he transit to memory state 5, he switch to the learning mechanism $(\mathscr{T}', d')$ that involves 5 memory states $5, 6, 7, 8, 9$. Thus, the proposed 13 memory states automaton can be interpretation as a mechanism that involve switching between three 5 memory states automatons.

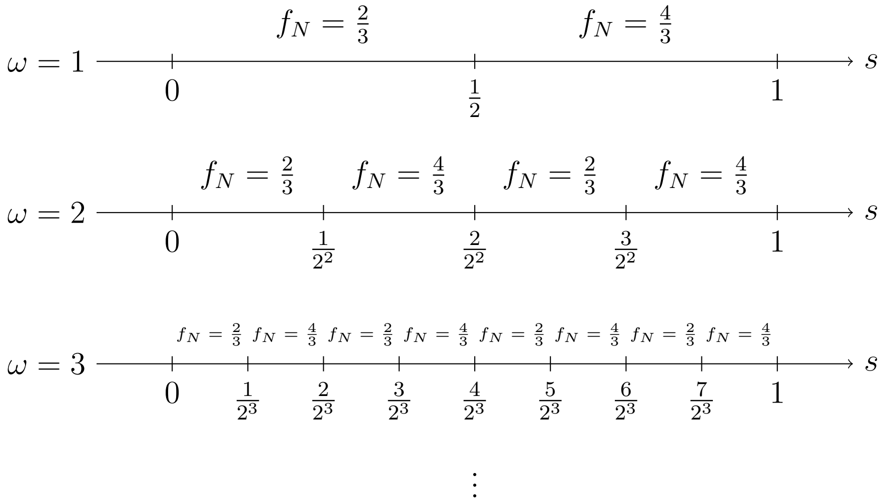

Figure 7: Signal structures that satisfies assumption 3 and 4' as $N \to \infty$. The signal structures comprise of low and high density alternatively in 2^ω equal-sized intervals.

B Example of a sequence of signal structures that satisfies assumption 3 and 4'

In this section, I provide an example of a sequence of signal structures that satisfies assumption 3 and 4'. Consider the following (set of) signal structures $(f^\omega)_{\omega=1}^N$ with $S = [0, 1)$:

$$f_N^1(s) = \begin{cases} \frac{2}{3} \text{ for } s \in [\frac{i}{2^1}, \frac{i+1}{2^1}) \text{ where } i = 0 \\ \frac{4}{3} \text{ for } s \in [\frac{i}{2^1}, \frac{i+1}{2^1}) \text{ where } i = 1 \end{cases}$$

$$f_N^\omega(s) = \begin{cases} \frac{2}{3} \text{ for } s \in [\frac{i}{2^\omega}, \frac{i+1}{2^\omega}) \text{ where } i = 0, 2, \cdots, 2^\omega - 2 \\ \frac{4}{3} \text{ for } s \in [\frac{i}{2^\omega}, \frac{i+1}{2^\omega}) \text{ where } i = 1, 3, \cdots, 2^\omega - 1 \end{cases}$$

The signal structures are illustrated in figure 7. First, it satisfies assumption 3 as $\frac{f_N^\omega(s)}{f_N^{\omega'}(s)} \geq \frac{1}{2}$ for all ω, ω' and s. Second, given any ω and $\omega' \neq \omega$ and for any N, the Cauchy-Schwarz distance is equal to:

$$\begin{aligned} -\log \frac{\int f_N^\omega(s) f_N^{\omega'}(s) ds}{\sqrt{\int (f_N^\omega(s))^2 ds} \sqrt{\int (f_N^{\omega'}(s))^2 ds}} &= -\log \frac{\frac{1}{4}(\frac{2}{3})^2 + \frac{1}{2}\frac{2}{3}\frac{4}{3} + \frac{1}{4}(\frac{4}{3})^2}{\frac{1}{2}(\frac{2}{3})^2 + \frac{1}{2}(\frac{4}{3})^2} \\ &= -\log \frac{1}{10/9} > 1 + \xi \end{aligned}$$

where $\xi < \log \frac{10}{9} - 1$ and thus satisfies assumption 4'.

C Omitted Proofs and Results

C.1 Proof of proposition 1

Proof. Before I prove the proposition, I present and prove the following lemma. With slight abuse of notations, I use F to denote the probability mass on any lotteries of signals, i.e., for any lotteries of signal $S' = \sum_s g(s) \times s$ where $g \in \triangle(S \cup \{\emptyset\})$, $F_N^\omega(S') \equiv \int f_N^\omega(s') g(s')\, ds'$.

Lemma C.1. *There exists $\delta > 1$ and a set of lotteries of signals $\{S_N^\omega\}_{\omega=1}^N$ such that*

$$\begin{aligned} &\frac{\delta F_N^\omega(S_N^{\omega'})}{\sum_{\omega'' \neq \omega'} F_N^\omega(S_N^{\omega''})} > 1 \text{ for all } \omega, \omega' \in \Omega; \\ &\frac{\delta F_N^\omega(S_N^{\omega})}{\sum_{\omega'' \neq \omega} F_N^\omega(S_N^{\omega''})} > \frac{\delta F_N^\omega(S_N^{\omega'})}{\sum_{\omega'' \neq \omega'} F_N^\omega(S_N^{\omega''})} \text{ for all } \omega \in \Omega \text{ and } \omega' \neq \omega. \end{aligned} \tag{C.1}$$

Moreover if $(\delta, \{S_N^\omega\}_{\omega=1}^N)$ satisfies equation (C.1), $(\delta, \{S_N^{\prime\omega}\}_{\omega=1}^N)$ where $S_N^{\prime\omega} = \beta \times \{S_N^{\prime\omega}\} + (1-\beta) \times \{\emptyset\}$ also satisfies equation (C.1) for all $\beta \in (0, 1]$.

Proof. First, note that δ does not affect the second inequality. Thus, for all ω and ω' such that $F_N^\omega(S_N^{\omega'}) > 0$, there always exists a big enough δ that satisfies the first inequality.

I proceed to show that there exists a set of lotteries of signals $\{S_N^\omega\}_{\omega=1}^N$ such that $F_N^\omega(S_N^{\omega'}) > 0$ for all $\omega, \omega' \in \Omega$ and satisfies the second inequality. Note that

$$\frac{\delta F_N^\omega(S_N^{\omega})}{\sum_{\omega'' \neq \omega} F_N^\omega(S_N^{\omega''})} > \frac{\delta F_N^\omega(S_N^{\omega'})}{\sum_{\omega'' \neq \omega'} F_N^\omega(S_N^{\omega''})} \iff F_N^\omega(S_N^\omega) > F_N^\omega(S_N^{\omega'}).$$

Denote $A \equiv \min\left\{1, \min_\omega \frac{\sqrt{\int (f_N^\omega(s))^2\, ds}}{\int (f_N^\omega(s))^2\, ds}\right\}$ and consider[34]

$$S_N^\omega = \sum_{s \in S} \left(\frac{A f_N^\omega(s)}{\sqrt{\int (f_N^\omega(s))^2 ds}} \times s \right) + \left(1 - A\sqrt{\int (f_N^\omega(s))^2 ds}\right) \times \emptyset, \tag{C.2}$$

i.e., a random lottery that put mass $\dfrac{A f_N^\omega(s)}{\sqrt{\int (f_N^\omega(s))^2 ds}}$ on each signal s and the remaining probability on the empty set. We have

$$F_N^\omega(S_N^\omega) = \frac{A \int (f_N^\omega(s))^2 ds}{\sqrt{\int (f_N^\omega(s))^2 ds}} = A\sqrt{\int (f_N^\omega(s))^2 ds} > \frac{A \int f_N^\omega(s) f_N^{\omega'}(s) ds}{\sqrt{\int (f_N^{\omega'}(s))^2 ds}} = F_N^\omega(S_N^{\omega'})$$

[34] Note that the choice of A is only to ensure that the probabilities in the random lottery of signal S_N^ω sum up to be 1. In particular, it does not affect the subsequent proof and the stationary probability distribution among memory states, as will be shown in the second part of the lemma C.1.

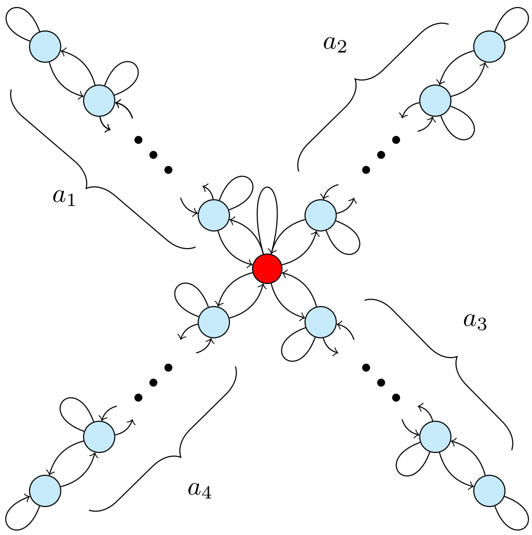

Figure 8: This figure illustrate a "star" updating mechanism with 4 actions. In the central (red) memory state, the DM randomly chooses one of the 4 actions. There are 4 equiv-length branches that correspond to each of the 4 actions.

where the inequality is given by the Cauchy–Schwarz inequality.

Now I prove the second part of the inequality. Note that

$$\begin{aligned}\frac{\delta F_N^{\omega}(S_N^{\prime\omega'})}{\sum_{\omega''\neq\omega'} F_N^{\omega}(S_N^{\prime\omega''})} =& \frac{\beta\delta F_N^{\omega}(S_N^{\omega'})}{\sum_{\omega''\neq\omega'} \beta F_N^{\omega}(S_N^{\omega''})} \\ =& \frac{\delta F_N^{\omega}(S_N^{\omega'})}{\sum_{\omega''\neq\omega'} F_N^{\omega}(S_N^{\omega''})}\end{aligned}$$

for all ω, ω' Thus, $(\delta, \{S_N^{\prime\omega}\}_{\omega=1}^N)$ satisfies equation (C.1). □

Now, I construct a sequence of updating mechanism with asymptotic utility loss converges to 0 as $M \to \infty$. Consider a "star" updating mechanism illustrated in figure 8. In the central memory state of the star, the DM randomly chooses one of the N actions; while there are N equiv-length branches that correspond to each of the N actions. For the ease of exposition, denote $\lambda = \lfloor (M-1)/N \rfloor$, I relabel the memory states in the star as $0, 11, 12, \cdots, 1\lambda, 21, 22, \cdots, 2\lambda, \cdots, N\lambda$ and denote the unused memory states as $N\lambda + 1, N\lambda + 2, \cdots, M-1$. Formally, the decision rule is as follows:[35]

$$\begin{aligned}&d(0) = \frac{1}{N} \times \{1\} + \frac{1}{N} \times \{2\} + \cdots + \frac{1}{N} \times \{N\};\\ &d(ik) = i \text{ for all } i = 1, \ldots, N \text{ and } k = 1, \cdots, \lambda;\\ &d(m) = 1 \text{ for all } m > N\lambda.\end{aligned}$$

The transition function between the memory states is defined as below for some $\delta > 1$ and $\{S_N^i\}_{i=1}^N$ that satisfies the two inequalities in lemma C.1. Denote $S_N^i(s)$

[35] The action chosen in the unused memory states can be assigned to any of the N actions.

as the probability assigned to the realization s in the lottery S_N^i and pick $\{S_N^i\}_{i=1}^N$ such that $\sum_{i=1}^N S_N^i(s) \leq 1$ for all $s \in S$.[36] Suppose the DM receives some signal s, he follows the following transition rule:

$$\mathscr{T}(0,s) = \sum_{i=1}^N S_N^i(s) \times \{i1\} + \left(1 - \sum_{i=1}^N S_N^i(s)\right) \times \{0\}$$

$$\begin{aligned}\mathscr{T}(i1,s) = S_N^i(s) \times \{i2\} + \sum_{j\in\Omega\setminus\{i\}} \frac{S_N^j(s)}{\delta} \times \{0\} \\ + \left[\sum_{j\in\Omega\setminus\{i\}} S_N^j(s)(1-\frac{1}{\delta}) + 1 - \sum_{j=1}^N S_N^j(s)\right] \times \{i1\}\end{aligned}$$

$$\begin{aligned}\mathscr{T}(i\lambda,s) = \sum_{j\in\Omega\setminus\{i\}} \frac{S_N^j(s)}{\delta} \times \{i(\lambda-1)\} \\ + \left[S_N^i(s) + \sum_{j\in\Omega\setminus\{i\}} S_N^j(s)(1-\frac{1}{\delta}) + 1 - \sum_{j=1}^N S_N^j(s)\right] \times \{i\lambda\}\end{aligned}$$

while for $k = 2, 3, \cdots, \lambda - 1$,

$$\begin{aligned}\mathscr{T}(ik,s) = S_N^i(s) \times \{i(k+1)\} + \sum_{j\in\Omega\setminus\{i\}} \frac{S_N^j(s)}{\delta} \times \{i(k-1)\} \\ + \left[\sum_{j\in\Omega\setminus\{i\}} S_N^j(s)(1-\frac{1}{\delta}) + 1 - \sum_{j=1}^N S_N^j(s)\right] \times \{ik\}\end{aligned}$$

Finally, for $m > N\lambda$, $\mathscr{T}(m,s) = m$ for all s. Restricting the initial memory state to one of $0, 11, 12, \cdots, 1\lambda, 21, 22, \cdots, 2\lambda, \cdots, N\lambda$, the DM will never transit to memory states $m > N\lambda$.

Before I prove the proposition, it would be useful to discuss the interpretation of the updating mechanism. It could be interpreted as a two-steps transition rule that involving first labeling the signal as supporting one of the states in Ω or not supporting any states, and then transiting based on the label of the signal. In particular, upon receiving a signal s, the DM labels it as supporting state ω with probability $S_N^\omega(s)$ for all $\omega \in \Omega$ and labels it as supporting no state with the remaining probability.

Next, the transition rule based on the labeling could be interpreted as the DM tracks only the favorable action and his confidence level of that action ranging from 1 to λ, as mentioned in the main text. Say at some period t the favorable action of

[36] Such $\{S_N^i\}_{i=1}^N$ exists as shown in the second part of lemma C.1.

the DM is ω, he revises his confidence level one unit upwards if he receives a signal supporting state ω (belief-confirming signal); he revises his confidence level one unit downwards with probability $\frac{1}{\delta} < 1$ if he receives a signal supporting other states (belief-challenging signal); in all other cases he does not revise his confidence level. It therefore can be interpreted as a learning algorithm with a particular (stochastic) definition of belief-confirming or belief-challenging signals, and under-reaction to belief challenging signals.

Now I compute the stationary probability distribution $\boldsymbol{\mu}_N^\omega$. Fix the state ω, in the stationary probability distribution, we have at the two extreme memory states in branch ω', i.e., memory states $\omega'\lambda$ and $\omega'(\lambda-1)$,

$$\begin{aligned}\mu^\omega_{N\omega'(\lambda-1)} F^\omega_N(S^{\omega'}_N) &= \mu^\omega_{N\omega'\lambda} \frac{1}{\delta} \sum_{\omega''\neq\omega'} F^\omega_N(S^{\omega''}_N) \\ \mu^\omega_{N\omega'(\lambda-1)} &= \mu^\omega_{N\omega'\lambda} \left[\frac{\delta F^\omega_N(S^{\omega'}_N)}{\sum_{\omega''\neq\omega'} F^\omega_N(S^{\omega''}_N)}\right]^{-1}\end{aligned}$$

for all ω'. That is, in the stationary distribution, the probability mass that enters memory state $\omega'\lambda$ equals that leaves it. It also implies at memory state $\omega'(\lambda-1)$,

$$\begin{aligned}\mu^\omega_{N\omega'\lambda} \frac{1}{\delta} \sum_{\omega''\neq\omega'} F^\omega_N(S^{\omega''}_N) + \mu^\omega_{N\omega'(\lambda-2)} F^\omega_N(S_{\omega'}) &= \mu^\omega_{N\omega'(\lambda-1)} \left[\frac{1}{\delta} \sum_{\omega''\neq\omega'} F^\omega_N(S^{\omega''}_N) + F^\omega_N(S^{\omega'}_N)\right] \\ \mu^\omega_{N\omega'(\lambda-2)} F^\omega_N(S^{\omega'}_N) &= \mu^\omega_{N\omega'(\lambda-1)} \frac{1}{\delta} \sum_{\omega''\neq\omega'} F^\omega_N(S^{\omega''}_N) \\ \mu^\omega_{N\omega'(\lambda-2)} &= \mu^\omega_{N\omega'(\lambda-1)} \left[\frac{\delta F^\omega_N(S^{\omega'}_N)}{\sum_{\omega''\neq\omega'} F^\omega_N(S^{\omega''}_N)}\right]^{-1}\end{aligned}$$

Repeating the same procedures implies that for all $k = 1, \cdots, \lambda$

$$\mu^\omega_{N\omega'k} = \mu^\omega_{N\omega'\lambda} \left[\frac{\delta F^\omega_N(S^{\omega'}_N)}{\sum_{\omega''\neq\omega'} F^\omega_N(S^{\omega''}_N)}\right]^{-(\lambda-k)} \tag{C.3}$$

and

$$\mu^\omega_{N0} = \mu^\omega_{N\omega'\lambda} \left[\frac{\delta F^\omega_N(S^{\omega'}_N)}{\sum_{\omega''\neq\omega'} F^\omega_N(S^{\omega''}_N)}\right]^{-\lambda} \tag{C.4}$$

As $\sum_{\omega'=1}^N \sum_{k=1}^\lambda \mu^\omega_{N\omega'k} + \mu^\omega_{N0} = 1$, we have

$$\begin{aligned}&\mu^\omega_{N\omega\lambda} \sum_{k=1}^{\lambda} \left[\frac{\delta F^\omega_N(S^\omega_N)}{\sum_{\omega''\neq\omega} F^\omega_N(S^{\omega''}_N)}\right]^{-(\lambda-k)} + \mu^\omega_{N\omega\lambda} \left[\frac{\delta F^\omega_N(S^\omega_N)}{\sum_{\omega''\neq\omega} F^\omega_N(S^{\omega''}_N)}\right]^{-\lambda} \\ &\quad + \mu^\omega_{N\omega\lambda} \left[\frac{\delta F^\omega_N(S^\omega_N)}{\sum_{\omega''\neq\omega} F^\omega_N(S^{\omega''}_N)}\right]^{-\lambda} \sum_{\omega'\neq\omega} \sum_{k=1}^{\lambda} \left[\frac{\delta F^\omega_N(S^{\omega'}_N)}{\sum_{\omega''\neq\omega'} F^\omega_N(S^{\omega''}_N)}\right]^{(\lambda-k)} = 1\end{aligned} \tag{C.5}$$

The two inequalities in lemma C.1 imply that fixing N, as $\lambda \to \infty$,

$$\mu^{\omega}_{N\omega\lambda}\left[\frac{\delta F^{\omega}_{N}(S^{\omega}_{N})}{\sum_{\omega''\neq\omega} F^{\omega}_{N}(S^{\omega''}_{N})}\right]^{-\lambda} + \mu^{\omega}_{N\omega\lambda}\left[\frac{\delta F^{\omega}_{N}(S^{\omega}_{N})}{\sum_{\omega''\neq\omega} F^{\omega}_{N}(S^{\omega''}_{N})}\right]^{-\lambda}\sum_{\omega'\neq\omega}\sum_{k=1}^{\lambda}\left[\frac{\delta F^{\omega}_{N}(S^{\omega'}_{N})}{\sum_{\omega''\neq\omega'} F^{\omega}_{N}(S^{\omega''}_{N})}\right]^{(\lambda-k)} \to 0 \quad \text{(C.6)}$$

Thus,

$$\sum_{k=1}^{\lambda}\mu^{\omega}_{N\omega k} = \mu^{\mu}_{N\omega\lambda}\sum_{k=1}^{\lambda}\left[\frac{\delta F^{\omega}_{N}(S^{\omega}_{N})}{\sum_{\omega''\neq\omega} F^{\omega}_{N}(S^{\omega''}_{N})}\right]^{-(\lambda-k)} \to 1 \qquad \text{(C.7)}$$

for all ω and the asymptotic utility loss of the proposed non-ignorant updating mechanism is 0, which proves $\lim_{\lambda\to\infty} L^{*}_{NM} = 0$. □

C.2 Proof of proposition 2

Proof. I show that for all $\epsilon > 0$, there exists a sequence of updating mechanism with utility loss converges to smaller than ϵ as $N, M \to \infty$. First, by assumption 4', for all $\frac{\epsilon}{\bar{u}} > 0$, there exists an $\xi > 0$ and a sequence of subset of states N_{ξ} such that $\lim_{N\to\infty}\left(\sum_{\omega\in N_{\xi}} p^{\omega}_{N}\right) > 1 - \frac{\epsilon}{\bar{u}}$ and

$$\lim_{N\to\infty}\inf_{\omega,\omega'\in N_{\xi};\omega'\neq\omega}\left\{-\log\frac{\int f^{\omega}_{N}(s)f^{\omega'}_{N}(s)ds}{\sqrt{\int (f^{\omega}_{N}(s))^{2}ds}\sqrt{\int (f^{\omega'}_{N}(s))^{2}ds}}\right\} > \xi$$

Before proving the proposition, I first prove the following lemma for $(S^{\omega}_{N})^{N}_{\omega=1}$ defined in equation (C.2).

Lemma C.2. *For $\frac{\epsilon}{\bar{u}} > 0$, there exists an $\tilde{\xi} > 0$, a sequence of $\delta_N > 1$ and a sequence of subset of states N_{ξ} with $\lim_{N\to\infty}\left(\sum_{\omega\in N_{\xi}} p^{\omega}_{N}\right) > 1 - \frac{\epsilon}{\bar{u}}$ such that*

$$\begin{gathered}\frac{\delta_N F^{\omega}_{N}(S^{\omega'}_{N})}{\sum_{\omega''\neq\omega'} F^{\omega}_{N}(S^{\omega''}_{N})} > 2 \text{ for all } \omega,\omega'\in\Omega \text{ and } N;\\ \lim_{N\to\infty}\inf_{\omega,\omega'\in N_{\xi};\omega'\neq\omega}\frac{\dfrac{\delta_N F^{\omega}_{N}(S^{\omega}_{N})}{\sum_{\omega''\neq\omega} F^{\omega}_{N}(S^{\omega''}_{N})}}{\dfrac{\delta_N F^{\omega}_{N}(S^{\omega'}_{N})}{\sum_{\omega''\neq\omega'} F^{\omega}_{N}(S^{\omega''}_{N})}} > 1+\tilde{\xi}.\end{gathered} \qquad \text{(C.8)}$$

Proof. As $F^{\omega}_{N}(S^{\omega'}_{N}) > 0$ for all ω, ω', there always exists a δ_N such that the first inequality of equation (C.8) holds. To prove the second inequality, note that

$$
\begin{aligned}
&\lim_{N\to\infty}\inf_{\omega,\omega'\in N_\xi;\omega'\neq\omega}\frac{\dfrac{\delta_N F_N^\omega(S_N^\omega)}{\sum_{\omega''\neq\omega}F_N^\omega(S_N^{\omega''})}}{\dfrac{\delta_N F_N^\omega(S_N^{\omega'})}{\sum_{\omega''\neq\omega'}F_N^\omega(S_N^{\omega''})}}\\
\geq&\lim_{N\to\infty}\inf_{\omega,\omega'\in N_\xi;\omega'\neq\omega}\frac{F_N^\omega(S_N^\omega)}{F_N^\omega(S_N^{\omega'})}\\
=&\lim_{N\to\infty}\inf_{\omega,\omega'\in N_\xi;\omega'\neq\omega}\frac{\sqrt{\int(f_N^\omega(s))^2ds}\sqrt{\int(f_N^{\omega'}(s))^2ds}}{\int f_N^\omega(s)f_N^{\omega'}(s)ds}\\
>&\exp(\xi)\\
=&1+\tilde{\xi}
\end{aligned}
\tag{C.9}
$$

where first inequality of equation (C.9) is implied by the fact that $F_N^\omega(S_N^\omega)\geq F_N^\omega(S_N^{\omega'})$ and thus $\frac{\sum_{\omega''\neq\omega'}F_N^\omega(S_N^{\omega''})}{\sum_{\omega''\neq\omega}F_N^\omega(S_N^{\omega''})}\geq 1$. □

Now consider a sequence of "star" updating mechanism described in the proof of proposition 1, but include only states in N_ξ, i.e., there are only branches correspond to actions in N_ξ and exists no $m\in M$ such that $d(m)=\omega'$ for $\omega'\notin N_\xi$. As in the proof of proposition 1, I compute the stationary distribution under state $\omega\in N_\xi$:

$$
\begin{aligned}
&\lim_{N,M\to\infty}\Bigg\{\mu^\omega_{N\omega\lambda}\sum_{k=1}^{\lambda}\left[\frac{\delta_N F_N^\omega(S_N^\omega)}{\sum_{\omega''\neq\omega}F_N^\omega(S_N^{\omega''})}\right]^{-(\lambda-k)}+\mu^\omega_{N\omega\lambda}\left[\frac{\delta_N F_N^\omega(S_N^\omega)}{\sum_{\omega''\neq\omega}F_N^\omega(S_N^{\omega''})}\right]^{-\lambda}\\
&+\mu^\omega_{N\omega\lambda}\left[\frac{\delta_N F_N^\omega(S_N^\omega)}{\sum_{\omega''\neq\omega}F_N^\omega(S_N^{\omega''})}\right]^{-\lambda}\sum_{\omega'\in N_\xi\setminus\{\omega\}}\sum_{k=1}^{\lambda}\left[\frac{\delta_N F_N^\omega(S_N^{\omega'})}{\sum_{\omega''\neq\omega'}F_N^\omega(S_N^{\omega''})}\right]^{(\lambda-k)}\Bigg\}=1
\end{aligned}
\tag{C.10}
$$

As $\lim_{N,M\to\infty}\frac{M}{N}=\infty$, $\lim_{N,M\to\infty}\lambda=\infty$. Given the first inequality of lemma C.2,

$\mu^{\omega}_{N\omega\lambda}\left[\frac{\delta_N F^{\omega}_N(S^{\omega}_N)}{\sum_{\omega''\neq\omega}F^{\omega}_N(S^{\omega''}_N)}\right]^{-\lambda} \to 0$ as $\lambda\to\infty$. On the other hand,

$$
\begin{aligned}
&\lim_{N,\lambda\to\infty}\left[\frac{\delta_N F^{\omega}_N(S^{\omega}_N)}{\sum_{\omega''\neq\omega}F^{\omega}_N(S^{\omega''}_N)}\right]^{-\lambda}\sum_{\omega'\in N_\xi\setminus\{\omega\}}\sum_{k=1}^{\lambda}\left[\frac{\delta_N F^{\omega}_N(S^{\omega'}_N)}{\sum_{\omega''\neq\omega'}F^{\omega}_N(S^{\omega''}_N)}\right]^{(\lambda-k)}\\
&=\lim_{N,\lambda\to\infty}\left[\frac{\delta_N F^{\omega}_N(S^{\omega}_N)}{\sum_{\omega''\neq\omega}F^{\omega}_N(S^{\omega''}_N)}\right]^{-\lambda}\sum_{\omega'\in N_\xi\setminus\{\omega\}}\frac{\left[\frac{\delta_N F^{\omega}_N(S^{\omega'}_N)}{\sum_{\omega''\neq\omega'}F^{\omega}_N(S^{\omega''}_N)}\right]^{\lambda}-1}{\frac{\delta F^{\omega}_N(S^{\omega'}_N)}{\sum_{\omega''\neq\omega'}F^{\omega}_N(S^{\omega''}_N)}-1}\\
&\le\lim_{N,\lambda\to\infty}\sum_{\omega'\in N_\xi\setminus\{\omega\}}\left[\left[\frac{\frac{\delta_N F^{\omega}_N(S^{\omega'}_N)}{\sum_{\omega''\neq\omega'}F^{\omega}_N(S^{\omega''}_N)}}{\frac{\delta_N F^{\omega}_N(S^{\omega}_N)}{\sum_{\omega''\neq\omega}F^{\omega}_N(S^{\omega''}_N)}}\right]^{\lambda}-\left[\frac{\delta_N F^{\omega}_N(S^{\omega}_N)}{\sum_{\omega''\neq\omega}F^{\omega}_N(S^{\omega''}_N)}\right]^{-\lambda}\right]\\
&\le\lim_{N,\lambda\to\infty}\sum_{\omega'\in N_\xi\setminus\{\omega\}}(1+\tilde{\xi})^{-\lambda}\\
&\le\lim_{N,\lambda\to\infty}N(1+\tilde{\xi})^{-\lambda}
\end{aligned}
$$

where the first inequality is implied by the first inequality in lemma C.2. As $(1+\tilde{\xi})^{-\lambda}$ converges to 0 exponentially and N converges to infinite linearly, $\lim_{N,\lambda\to\infty}N(1+\tilde{\xi})^{-\lambda}=0$. To see it formally, note that $N=O(M^h)$ with $h<1$ and $\lambda=\frac{M}{N}$ implies that $N=O(\lambda^{\frac{h}{1-h}})$ where $\frac{h}{1-h}\in(0,\infty)$. We have

$$
\begin{aligned}
\lim_{\lambda\to\infty}\frac{\lambda^{\frac{h}{1-h}}}{(1+\tilde{\xi})^{\lambda}}&=\lim_{\lambda\to\infty}\frac{\frac{h}{1-h}\lambda^{\frac{h}{1-h}-1}}{(\log(1+\tilde{\xi}))(1+\tilde{\xi})^{\lambda}}\\
&=\lim_{\lambda\to\infty}\frac{\frac{h}{1-h}(\frac{h}{1-h}-1)\lambda^{\frac{h}{1-h}-2}}{(\log(1+\tilde{\xi}))^2(1+\tilde{\xi})^{\lambda}}\\
&\cdots\\
&=0
\end{aligned}
$$

Thus,

$$
\begin{aligned}
0\le&\lim_{N,\lambda\to\infty}\left[\frac{\delta_N F^{\omega}_N(S^{\omega}_N)}{\sum_{\omega''\neq\omega}F^{\omega}_N(S^{\omega''}_N)}\right]^{-\lambda}\sum_{\omega'\in N_\xi\setminus\{\omega\}}\sum_{k=1}^{\lambda}\left[\frac{\delta_N F^{\omega}_N(S^{\omega'}_N)}{\sum_{\omega''\neq\omega'}F^{\omega}_N(S^{\omega''}_N)}\right]^{(\lambda-k)}\\
&\le\lim_{N,\lambda\to\infty}N(1+\tilde{\xi})^{-\lambda}=0
\end{aligned}
$$

and $\lim_{N,\lambda\to\infty}\left[\frac{\delta_N F^{\omega}_N(S^{\omega}_N)}{\sum_{\omega''\neq\omega}F^{\omega}_N(S^{\omega''}_N)}\right]^{-\lambda}\sum_{\omega'\in N_\xi\setminus\{\omega\}}\sum_{k=1}^{\lambda}\left[\frac{\delta_N F^{\omega}_N(S^{\omega'}_N)}{\sum_{\omega''\neq\omega'}F^{\omega}_N(S^{\omega''}_N)}\right]^{(\lambda-k)}=0$, which implies

$$
\lim_{N,\lambda\to\infty}\sum_{k=1}^{\lambda}\mu^{\omega}_{N\omega k}=\lim_{N,\lambda\to\infty}\mu^{\mu}_{N\omega\lambda}\sum_{k=1}^{\lambda}\left[\frac{\delta_N F^{\omega}_N(S^{\omega}_N)}{\sum_{\omega''\neq\omega}F^{\omega}_N(S^{\omega''}_N)}\right]^{-(\lambda-k)}=1 \tag{C.11}
$$

for all $\omega \in N_\xi$. As $\lim_{N\to\infty} \sum_{\omega \in N_\xi} p_N^\omega > 1 - \frac{\epsilon}{\bar{u}}$, the asymptotic utility loss of the proposed non-ignorant updating mechanism is bounded above by $\bar{u} \times \frac{\epsilon}{\bar{u}} = \epsilon$. □

C.3 Proof of corollary 1

Proof. Note that an ignorant updating mechanism induces utility loss weakly greater than $\min_\omega u_N^\omega p_N^\omega$ which is invariant in M. On the other hand, as shown in proposition 1, L_{NM}^* converges to 0 as $M \to \infty$. It implies there exists some big enough $\bar{M}$ such that for $M > \bar{M}$, $L_{NM}^* < \min_\omega u_N^\omega p_N^\omega$. Consider $\epsilon < \min_\omega u_N^\omega p_N^\omega - L_{N\bar{M}}^*$, if an updating mechanism $\mathscr{T}$ ignores some state ω', for we have for $M > \bar{M}$

$$L_N(\mathscr{T}) \geq u_N^{\omega'} p_N^{\omega'} \geq \min_\omega u_N^\omega p_N^\omega > L_{N\bar{M}}^* + \epsilon$$

which proves the result. □

C.4 Proof of corollary 2

Proof. Suppose in contrary the sequence of updating mechanism $\mathscr{T}_N$ ignores strictly more than $\frac{\epsilon}{\underline{u}}$ measures of states at the limit. Denoted the set of states that is ignored by $\tilde{N}$, the utility loss is

$$\lim_{N,M\to\infty} L_N(\mathscr{T}_N) \geq \lim_{N\to\infty} \sum_{\omega\in\tilde{N}} u_N^\omega p_N^\omega \geq \underline{u} \lim_{N\to\infty} \sum_{\omega\in\tilde{N}} p_N^\omega > \underline{u}\frac{\epsilon}{\underline{u}} = \epsilon = \lim_{N,M\to\infty} L_{NM}^* + \epsilon$$

which proves the result. □

C.5 Proof of corollary 3

Proof. By proposition 1, we have for all ω

$$\lim_{M\to\infty} \lim_{\epsilon\to 0} \lim_{t\to\infty} \Pr_N(a_t^I = \omega \mid \omega) = 1 \text{ for } I = A, B.$$

which proves the result. □

C.6 Proof of corollary 4

Proof. By proposition 2, we have for individual I:

$$\lim_{N,M\to\infty} \lim_{\epsilon\to 0} \lim_{t\to\infty} \sum_{\omega=1}^{N} \left[\mathbb{1}\left\{ \lim_{t\to\infty} \Pr_N(a_t^I = \omega \mid \omega) = 1 \right\} p_I^\omega \right] = 1$$

which is equivalent to

$$\lim_{N,M\to\infty} \lim_{\epsilon\to 0} \lim_{t\to\infty} \sum_{\omega=1}^{N} \left[\mathbb{1}\left\{ \lim_{t\to\infty} \Pr_N(a_t^I = \omega \mid \omega) < 1 \right\} p_I^\omega \right] = 0,$$

i.e., individual I would only take sub-optimal actions in probability 0 events measured by $(p^{\omega}_{NI})^{N}_{\omega=1}$. As $\lim_{N\to\infty} p^{\omega}_{NA} > 0$ if and only if $\lim_{N\to\infty} p^{\omega}_{NB} > 0$ for all ω, combined with assumption 2' implies that individual A and B agree on the probability 0 events. That is, for any sequence of subset of states $\hat{N}$ where $\lim_{N\to\infty} \sum_{\omega\in\hat{N}} p^{\omega}_{NA} = 0$, we have $\lim_{N\to\infty} \sum_{\omega\in\hat{N}} p^{\omega}_{NB} = 0$, which implies the result. □

C.7 Robustness of the results in small world to updating mistakes

In below I show that the behavior implications is small world, i.e., learning is close to Bayesian, and that disagreement does not persist, hold even when individuals make "updating mistakes". In other words, the results are robust to individuals' limited ability to design and follow an "optimal" updating mechanism.

Consider two individuals A and B. Individual A adopts the star updating mechanism described in the proof of proposition 1 while individual B "attempts" to adopt the same updating mechanism but makes local mistakes as he randomly transits to neighbor memory states with some probability $\gamma \in (0, 1)$. Formally, the transition rule of individual B, denoted as $\mathscr{T}'(m, s)$, is as follows:

$$\mathscr{T}'(0, s) = (1-\gamma) \times \mathscr{T}(0, s) + \sum_{j=1}^{N} \frac{\gamma}{N} \times \{j1\}$$

$$\mathscr{T}'(i1, s) = (1-\gamma) \times \mathscr{T}(i1, s) + \frac{\gamma}{2} \times \{i2\} + \frac{\gamma}{2} \times \{0\}$$
$$\mathscr{T}'(i\lambda, s) = (1-\gamma) \times \mathscr{T}(i\lambda, s) + \gamma \times \{i(\lambda-1)\}$$

while for $k = 2, 3, \cdots, \lambda - 1$,

$$\mathscr{T}'(ik, s) = (1-\gamma) \times \mathscr{T}(ik, s) + \frac{\gamma}{2} \times \{i(k-1)\} + \frac{\gamma}{2} \times \{i(k+1)\}$$

where $\mathscr{T}(m, s)$ is defined in the proof of proposition 1. Such updating mistakes could be induced by memory imperfection, i.e., the DM's memory state is subject to local fluctuations, or imperfect perception on signals, e.g., the DM may mistakenly perceive any signal as a signal that support state ω.

Proposition C.1. *Consider individual A who adopts the star updating mechanism and individual B who makes local mistakes with some probability $\gamma \in (0, 1)$, characterized by $\mathscr{T}'(m, s)$. Fix a finite N, for all $\gamma \in (0, 1)$, utility loss of individual B converges to 0 as $M \to \infty$, i.e.,*

$$\lim_{M\to\infty} L_N(\mathscr{T}') = 0.$$

Individual A and B are bound to agree in small worlds, i.e., for all $(u^{\omega}_{NA}, p^{\omega}_{NA}, f^{\omega}_{NA})^{N}_{\omega=1}$,

$(u^{\omega}_{NB}, p^{\omega}_{NB}, f^{\omega}_{NB})^{N}_{\omega=1}$, *fixing* N *and* $M \to \infty$,

$$\lim_{M\to\infty} \lim_{\epsilon\to 0} \lim_{t\to\infty} Pr_N(a^A_t = a^B_t \mid \omega) = 1 \text{ for all } \omega.$$

Proof. The proof follows closely the proof of proposition 1. For individual B, fixing the state ω, in the stationary probability distribution, we have at the two extreme memory states in branch ω',

$$\mu^{\omega}_{N\omega'\lambda}\left[\frac{1-\gamma}{\delta}\sum_{\omega''\neq\omega'} F^{\omega}_{N}(S^{\omega''}_{N}) + \gamma\right] = \mu^{\omega}_{N\omega'(\lambda-1)}\left[(1-\gamma)F^{\omega}_{N}(S^{\omega'}_{N}) + \frac{\gamma}{2}\right]$$

$$\mu^{\omega}_{N\omega'(\lambda-1)} = \mu^{\omega}_{N\omega'\lambda}\left[\frac{(1-\gamma)F^{\omega}_{N}(S^{\omega'}_{N}) + \frac{\gamma}{2}}{\frac{1-\gamma}{\delta}\sum_{\omega''\neq\omega'} F^{\omega}_{N}(S^{\omega''}_{N}) + \gamma}\right]^{-1}$$

for all ω'. Similarly, at memory state $\omega'(\lambda-1)$,

$$\mu^{\omega}_{N\omega'\lambda}\left[\frac{1-\gamma}{\delta}\sum_{\omega''\neq\omega'} F^{\omega}_{N}(S^{\omega''}_{N}) + \gamma\right] + \mu^{\omega}_{N\omega'(\lambda-2)}\left[(1-\gamma)F^{\omega}_{N}(S^{\omega'}_{N}) + \frac{\gamma}{2}\right] = \mu^{\omega}_{N\omega'(\lambda-1)}\left[\frac{1-\gamma}{\delta}\sum_{\omega''\neq\omega'} F^{\omega}_{N}(S^{\omega''}_{N}) + (1-\gamma)F^{\omega}_{N}(S^{\omega'}_{N}) + \gamma\right]$$

$$\mu^{\omega}_{N\omega'(\lambda-2)}\left[(1-\gamma)F^{\omega}_{N}(S^{\omega'}_{N}) + \frac{\gamma}{2}\right] = \mu^{\omega}_{N\omega'(\lambda-1)}\left[\frac{1-\gamma}{\delta}\sum_{\omega''\neq\omega'} F^{\omega}_{N}(S^{\omega''}_{N}) + \frac{\gamma}{2}\right]$$

$$\mu^{\omega}_{N\omega'(\lambda-2)} = \mu^{\omega}_{N\omega'(\lambda-1)}\left[\frac{(1-\gamma)F^{\omega}_{N}(S^{\omega'}_{N}) + \frac{\gamma}{2}}{\frac{1-\gamma}{\delta}\sum_{\omega''\neq\omega'} F^{\omega}_{N}(S^{\omega''}_{N}) + \frac{\gamma}{2}}\right]^{-1}$$

Repeating the same procedures implies that for all $k = 1, \cdots, \lambda - 1$

$$\mu^{\omega}_{N\omega'k} = \mu^{\omega}_{N\omega'\lambda}\left[\frac{(1-\gamma)F^{\omega}_{N}(S^{\omega'}_{N}) + \frac{\gamma}{2}}{\frac{1-\gamma}{\delta}\sum_{\omega''\neq\omega'} F^{\omega}_{N}(S^{\omega''}_{N}) + \gamma}\right]^{-1}\left[\frac{(1-\gamma)F^{\omega}_{N}(S^{\omega'}_{N}) + \frac{\gamma}{2}}{\frac{1-\gamma}{\delta}\sum_{\omega''\neq\omega'} F^{\omega}_{N}(S^{\omega''}_{N}) + \frac{\gamma}{2}}\right]^{-(\lambda-k-1)} \tag{C.12}$$

and

$$\mu^{\omega}_{N0} = \mu^{\omega}_{N\omega'\lambda}\left[\frac{(1-\gamma)F^{\omega}_{N}(S^{\omega'}_{N}) + \frac{\gamma}{2}}{\frac{1-\gamma}{\delta}\sum_{\omega''\neq\omega'} F^{\omega}_{N}(S^{\omega''}_{N}) + \gamma}\right]^{-1}\left[\frac{(1-\gamma)F^{\omega}_{N}(S^{\omega'}_{N}) + \frac{\gamma}{2}}{\frac{1-\gamma}{\delta}\sum_{\omega''\neq\omega'} F^{\omega}_{N}(S^{\omega''}_{N}) + \frac{\gamma}{2}}\right]^{-(\lambda-2)}\left[\frac{(1-\gamma)F^{\omega}_{N}(S^{\omega'}_{N}) + \frac{\gamma}{N}}{\frac{1-\gamma}{\delta}\sum_{\omega''\neq\omega'} F^{\omega}_{N}(S^{\omega''}_{N}) + \frac{\gamma}{2}}\right] \tag{C.13}$$

Note that for all $\gamma \in (0, 1)$

$$\frac{(1-\gamma)F^{\omega}_{N}(S^{\omega}_{N}) + \frac{\gamma}{2}}{\frac{1-\gamma}{\delta}\sum_{\omega''\neq\omega} F^{\omega}_{N}(S^{\omega''}_{N}) + \frac{\gamma}{2}} > 1.$$

$$\frac{(1-\gamma)F^{\omega}_{N}(S^{\omega}_{N}) + \frac{\gamma}{2}}{\frac{1-\gamma}{\delta}\sum_{\omega''\neq\omega} F^{\omega}_{N}(S^{\omega''}_{N}) + \frac{\gamma}{2}} > \frac{(1-\gamma)F^{\omega}_{N}(S^{\omega'}_{N}) + \frac{\gamma}{2}}{\frac{1-\gamma}{\delta}\sum_{\omega''\neq\omega} F^{\omega}_{N}(S^{\omega''}_{N}) + \frac{\gamma}{2}}$$

which is the analogue of lemma C.1. Then following the same steps in the proof of proposition 1 proves the result. □

Next I prove the analogue result in the case where $N, M \to \infty$ and $N = O(M^h)$ where $h < 1$. Assume that individual A adopts a star updating mechanism with a sequence of δ_N that satisfies

$$\frac{(1-\gamma)F_N^{\omega}(S_N^{\omega'})+\frac{\gamma}{2}}{\frac{1-\gamma}{\delta_N}\sum_{\omega''\neq\omega'}F_N^{\omega}(S_N^{\omega''})+\frac{\gamma}{2}}>2 \text{ for all } \omega,\omega'\in\Omega \text{ and } N.$$

Fixing $\gamma\in(0,1)$, such sequence of δ_N always exists as $F_N^{\omega}(S_N^{\omega'})>0$ for all $\omega,\omega'\in\Omega$ and N. It also implies that δ_N goes to ∞ and

$$\begin{aligned}
&\lim_{N\to\infty}\inf_{\omega,\omega'\in N_\xi;\omega'\neq\omega}\frac{\frac{(1-\gamma)F_N^{\omega}(S_N^{\omega})+\frac{\gamma}{2}}{\frac{1-\gamma}{\delta_N}\sum_{\omega''\neq\omega}F_N^{\omega}(S_N^{\omega''})+\frac{\gamma}{2}}}{\frac{(1-\gamma)F_N^{\omega}(S_N^{\omega'})+\frac{\gamma}{2}}{\frac{1-\gamma}{\delta_N}\sum_{\omega''\neq\omega'}F_N^{\omega}(S_N^{\omega''})+\frac{\gamma}{2}}}\\
=&\lim_{N\to\infty}\inf_{\omega,\omega'\in N_\xi;\omega'\neq\omega}\frac{\frac{(1-\gamma)F_N^{\omega}(S_N^{\omega})+\frac{\gamma}{2}}{(1-\gamma)\sum_{\omega''\neq\omega}F_N^{\omega}(S_N^{\omega''})+\frac{\gamma\delta_N}{2}}}{\frac{(1-\gamma)F_N^{\omega}(S_N^{\omega'})+\frac{\gamma}{2}}{(1-\gamma)\sum_{\omega''\neq\omega'}F_N^{\omega}(S_N^{\omega''})+\frac{\gamma\delta_N}{2}}}\\
\geq&\lim_{N\to\infty}\inf_{\omega,\omega'\in N_\xi;\omega'\neq\omega}\frac{F_N^{\omega}(S_N^{\omega})+\frac{\gamma}{2(1-\gamma)}}{F_N^{\omega}(S_N^{\omega'})+\frac{\gamma}{2(1-\gamma)}}
\end{aligned}$$

where the last inequality is implied by the fact that $\sum_{\omega''\neq\omega}F_N^{\omega}(S_N^{\omega''})\leq\sum_{\omega''\neq\omega'}F_N^{\omega}(S_N^{\omega''})$. In lemma C.2, we know that for all $\frac{\epsilon}{u}>0$ there exists $\tilde{\xi}>0$ and a sequence of subset of state N_ξ with $\lim_{N\to\infty}\left(\sum_{N\in N_\xi}p_N^{\omega}\right)>1-\frac{\epsilon}{u}$ such that

$$\lim_{N\to\infty}\inf_{\omega,\omega'\in N_\xi;\omega'\neq\omega}\frac{F_N^{\omega}(S_N^{\omega})}{F_N^{\omega}(S_N^{\omega'})}>1+\tilde{\xi}.$$

Thus, fixing $\gamma<1$, there also exists some $\tilde{\tilde{\xi}}>0$ such that

$$\lim_{N\to\infty}\inf_{\omega,\omega'\in N_\xi;\omega'\neq\omega}\frac{F_N^{\omega}(S_N^{\omega})+\frac{\gamma}{2(1-\gamma)}}{F_N^{\omega}(S_N^{\omega'})+\frac{\gamma}{2(1-\gamma)}}>1+\tilde{\tilde{\xi}}.$$

Then following the same steps in the proof of proposition 2 gives the following result.

Proposition C.2. *Consider individual A who adopts the star updating mechanism and individual B who makes local mistakes with some probability $\gamma\in(0,1)$, characterize by $\mathscr{T}'(m,s)$. Suppose $N,M\to\infty$ and $N=O(M^h)$ where $h<1$,*

$$\lim_{N,M\to\infty}L_N(\mathscr{T}')=0.$$

Moreover, the two individuals are almost bound to agree asymptotically in small worlds if they agree on the probability 0 events. For all $(u_{NA}^{\omega},p_{NA}^{\omega},f_{NA}^{\omega})_{\omega=1}^{N}$ and $(u_{NB}^{\omega},p_{NB}^{\omega},f_{NB}^{\omega})_{\omega=1}^{N}$ such that $\lim_{N\to\infty}p_{NA}^{\omega}>0$ if and only if $\lim_{N\to\infty}p_{NB}^{\omega}>0$ for all ω, then

$$\lim_{N,M\to\infty}\lim_{\epsilon\to 0}\sum_{\omega}\left[\mathbb{1}\left\{\lim_{t\to\infty}Pr_N(a_t^A=a_t^B\mid\omega)=1\right\}p_{NI}^{\omega}\right]=1 \text{ for } I=A,B.$$

The result illustrates the robustness of agreement in small world. In particular, even if the individual makes local mistakes with probability close to 1, he will (almost) learn perfectly the true state of the world asymptotically. Combining corollary 3, 4 and proposition C.1 and C.2, we therefore expect disagreement to vanish over time in small world among different individuals with different prior beliefs, abilities of information acquisition or abilities to adopt an "good" belief updating mechanism.

C.8 Proof of proposition 3

Proof. When N and M is finite, consider two states ω, ω', with similar argument in Hellman and Cover (1970), we have

$$\Upsilon_N^{\omega\omega'} \leq (\bar{l}_N^{\omega\omega'} \bar{l}_N^{\omega'\omega})^{M-1}$$
$$\min_{m\in M^{\omega'}} \frac{\mu_{Nm}^{\omega}}{\mu_{Nm}^{\omega'}} \geq (\bar{l}_N^{\omega\omega'} \bar{l}_N^{\omega'\omega})^{-(M-1)} \max_{m\in M^{\omega}} \frac{\mu_{Nm}^{\omega}}{\mu_{Nm}^{\omega'}}$$

where $\bar{l}_N^{\omega\omega'} = \sup_s \frac{f_N^{\omega}(s)}{f_N^{\omega'}(s)}$. Suppose the DM chooses action ω in state ω with probability $1-\varepsilon$ and chooses action ω in state ω' with probability ε', i.e.,

$$\sum_{m\in M^{\omega}} \mu_{Nm}^{\omega} = 1-\varepsilon \text{ and } \sum_{m\in M^{\omega}} \mu_{Nm}^{\omega'} = \varepsilon'.$$

This implies that $\min\max_{m\in M^{\omega}} \frac{\mu_{Nm}^{\omega}}{\mu_{Nm}^{\omega'}} = \frac{1-\varepsilon}{\varepsilon'}$ and

$$\min_{m\in M^{\omega'}} \frac{\mu_{Nm}^{\omega}}{\mu_{Nm}^{\omega'}} \geq (\bar{l}_N^{\omega\omega'} \bar{l}_N^{\omega'\omega})^{-(M-1)} \frac{1-\varepsilon}{\varepsilon'}.$$

Moreover, as $\sum_{m\in M^{\omega'}} \mu_{Nm}^{\omega} + \sum_{m\in M^{\omega}} \mu_{Nm}^{\omega} \leq 1$, we have $\sum_{m\in M^{\omega'}} \mu_{Nm}^{\omega} \leq \varepsilon$ and

$$\frac{\epsilon}{\max_{m\in M^{\omega'}} \mu_{Nm}^{\omega'}} \geq \min_{m\in M^{\omega'}} \frac{\mu_{Nm}^{\omega}}{\mu_{Nm}^{\omega'}} \geq (\bar{l}_N^{\omega\omega'} \bar{l}_N^{\omega'\omega})^{-(M-1)} \frac{1-\varepsilon}{\varepsilon'}$$
$$\max_{m\in M^{\omega'}} \mu_{Nm}^{\omega'} \leq (\bar{l}_N^{\omega\omega'} \bar{l}_N^{\omega'\omega})^{M-1} \frac{\varepsilon\varepsilon'}{1-\varepsilon}$$

As $(\bar{l}_N^{\omega\omega'} \bar{l}_N^{\omega'\omega})^{M-1}$ is bounded above, for ε and ε' small enough, we must have

$$\sum_{m\in M^{\omega'}} \mu_{Nm}^{\omega'} < M \max_{m\in M^{\omega'}} \mu_{Nm}^{\omega'} < M (\bar{l}_N^{\omega\omega'} \bar{l}_N^{\omega'\omega})^{M-1} \frac{\varepsilon\varepsilon'}{1-\varepsilon} < 1.$$

Thus, if the DM chooses ε and ε' close to 0, we must have $\sum_{m\in M^{\omega'}} \mu_{Nm}^{\omega'}$ close to 0 and the utility loss is bigger than $u_N^{\omega'} p_N^{\omega'}$. Therefore, $L^* > 0$.

To prove the second part of the proposition, note that the DM can always "throw away" memory states. Formally, consider an updating mechanism $(\mathscr{T}, d)$ with M, and suppose the DM's memory capacity increases to $M' > M$. He can design an updating mechanism $(\mathscr{T}', d')$ with $\mathscr{T}'(m, s) = \mathscr{T}(m, s)$ and $d'(m) = d(m)$ for all $m \leq M$. By choosing an initial memory state in $m \leq M$, the DM will never transit to memory states $m > M$. The stationary distribution and thus utility loss does not change. Therefore, the DM can always secure a weakly lower L^* when M increases, i.e., L^* weakly decreases in M. □

C.9 Proof of proposition 4

Proof. First, note that for $\lim_{N,M\to\infty} L^*_{NM} = 0$, we must have

$$\lim_{N,M\to\infty} \sum_{m\in M^\omega} \mu^\omega_{Nm} = 1$$

for almost all ω, i.e., there must exists a sequence of subset of states $\hat{N}$ where $\lim_{N\to\infty} \sum_{\omega\in\hat{N}} p^\omega_N = 1$ and

$$\lim_{N,M\to\infty} \sum_{m\in M^\omega} \mu^\omega_{Nm} = 1.$$

for all $\omega \in \hat{N}$. Moreover, assumption 2' implies that there must exists a sequence of subset of states $\hat{N}$ where $\lim_{N\to} \frac{|\hat{N}|}{N} = 1$ and

$$\lim_{N,M\to\infty} \sum_{m\in M^\omega} \mu^\omega_{Nm} = 1,$$

for all $\omega \in \hat{N}$. That is, the DM chooses the optimal action in almost all states, measured in both prior probability or in fraction. It implies that for all ω in $\hat{N}$, there must exist a set of memory state $\hat{M}^\omega \subseteq M^\omega$ such that

$$\begin{aligned}
&\lim_{N,M\to\infty} \sum_{m\in\hat{M}^\omega} \mu^\omega_{Nm} = 1 \\
&\lim_{N,M\to\infty} \sum_{m\in\hat{M}^\omega} \mu^{\omega'}_{Nm} = 0 \text{ for all } \omega' \in \hat{N}\setminus\{\omega\} \\
&\lim_{N,M\to\infty} \max_{m\in\hat{M}^\omega} \frac{\mu^\omega_{Nm}}{\mu^{\omega'}_{Nm}} = \infty \text{ for all } \omega' \in \hat{N}\setminus\{\omega\}
\end{aligned} \tag{C.14}$$

In the following I prove that for equation (C.14) to hold, $\frac{M}{N}$ has to go to ∞. First consider an irreducible automaton. Fix a $\omega' \in \hat{N}\setminus\{\omega\}$, without loss of generality rearrange the memory states such that $\frac{\mu^\omega_{Nm}}{\mu^{\omega'}_{Nm}}$ is weakly decreasing in m. By lemma 2 of Hellman and Cover (1970), we have for all $m < M$,

$$\begin{aligned}\frac{\mu^{\omega}_{N(m+1)}}{\mu^{\omega'}_{N(m+1)}} &\geq (\bar{l}_N^{\omega\omega'}\bar{l}_N^{\omega'\omega})^{-1}\frac{\mu^{\omega}_{Nm}}{\mu^{\omega'}_{Nm}} \\ &\geq \varsigma^2\frac{\mu^{\omega}_{Nm}}{\mu^{\omega'}_{Nm}}.\end{aligned} \tag{C.15}$$

As there must exists some m with $\frac{\mu^{\omega}_{Nm}}{\mu^{\omega'}_{Nm}} \leq 1$, if $\max\frac{\mu^{\omega}_{Nm}}{\mu^{\omega'}_{Nm}} = \frac{\mu^{\omega}_{N1}}{\mu^{\omega'}_{N1}} > K$, equation (C.15) implies that

$$\frac{\mu^{\omega}_{Nm'}}{\mu^{\omega'}_{Nm'}} \geq \varsigma^{2(m'-1)}\frac{\mu^{\omega}_{N1}}{\mu^{\omega'}_{N1}} > \varsigma^{2(m'-1)}K$$

and $\varsigma^{2(m'-1)}K \geq 1$ for all $m'-1 \leq \frac{\log K}{2\log(\varsigma)^{-1}}$. In other words, there must exists at least $\frac{\log K}{2\log(\varsigma)^{-1}}+1$ memory states with a state $\omega-\omega'$ likelihood ratio $\frac{\mu^{\omega}_{Nm}}{\mu^{\omega'}_{Nm}} \geq 1$. Repeating the same analysis for other $\omega'' \in \hat{N}\setminus\{\omega\}$ implies that if $\max\frac{\mu^{\omega}_{Nm}}{\mu^{\omega''}_{Nm}} > K$ for all $\omega'' \in \hat{N}\setminus\{\omega'\}$, there must exists at $\frac{\log K}{2\log(\varsigma)^{-1}}+1$ memory states with a likelihood ratio $\min_{\omega'\in\hat{N}}\frac{\mu^{\omega}_{Nm}}{\mu^{\omega'}_{Nm}} \geq 1$.

With similar arguments, if $\max\frac{\mu^{\omega}_{Nm}}{\mu^{\omega'}_{Nm}} > K$ for all $\omega \in \hat{N}$ and all $\omega' \in \hat{N}\setminus\{\omega\}$, $\frac{M}{|\hat{N}|}$ must be weakly greater than $\frac{\log K}{2\log(\varsigma)^{-1}}+1$. It implies that $\frac{M}{|\hat{N}|}$ goes to ∞ as K goes to ∞. It contradicts the fact that $\lim_{N,M\to\infty}\frac{M}{|\hat{N}|} = \lim_{N,M\to\infty}\frac{M}{N} < \infty$ in big worlds.

Now I analysis the case of reducible automatons. Denote the recurrent communicating classes as $\mathscr{R}_1,\cdots,\mathscr{R}_r$, and the set of transient memory states as $\mathscr{R}_0$. The analysis above applies in the cases where there is only one recurrent communicating class or where the initial memory state is in one of the recurrent communicating classes.

Now consider the case where $r > 1$, i.e., there are more than one recurrent communicating class, and the initial memory state denoted by i is in $\mathscr{R}_0$. I first compute the probability of absorption by $\mathscr{R}_j$ under state ω, denoted by $\mathscr{P}^{\omega}_N(\mathscr{R}_j)$. Consider a new transition rule $\mathscr{T}'$ where all transitions from $m \in \mathscr{R}_0$ to another $m' \in \mathscr{R}_0$ is the same as before. However, $\mathscr{T}'$ differs from $\mathscr{T}$ that all transitions from $m \in \mathscr{R}_0$ to $m' \notin \mathscr{R}_0$ are changed to transition from m to i. Given such a transition rule $\mathscr{T}'$, obviously only memory states in $\mathscr{R}_0$ are reachable. Denote $\mu^{0\omega}_{Nm}$ as the stationary distribution of this new transition rule $\mathscr{T}'$.

As is known in the theory of Markov chain (see appendix 2 of Hellman (1969)), $\mathscr{P}^{\omega}(\mathscr{R}_j)$ is given by:

$$\mathscr{P}^{\omega}_N(\mathscr{R}_j) = \sum_{m\in\mathscr{R}_0}\mu^{0\omega}_{Nm}\sum_{m'\in\mathscr{R}_j}q^{\omega}_{mm'}$$

Also denote $\mu^{j\omega}_{Nm}$ as the stationary distribution of the recurrent communicating class

$\mathscr{R}_j$, we have for $m \in \mathscr{R}_j$

$$\begin{aligned}\frac{\mu^{\omega}_{Nm}}{\mu^{\omega'}_{Nm}} =& \frac{\mathscr{P}^{\omega}_N(\mathscr{R}_j)}{\mathscr{P}^{\omega'}_N(\mathscr{R}_j)} \times \frac{\mu^{j\omega}_{Nm}}{\mu^{j\omega'}_{Nm}} \\ =& \frac{\sum_{m\in\mathscr{R}_0} \mu^{0\omega}_{Nm} \sum_{m'\in\mathscr{R}_j} q^{\omega}_{mm'}}{\sum_{m\in\mathscr{R}_0} \mu^{0\omega'}_{Nm} \sum_{m'\in\mathscr{R}_j} q^{\omega'}_{mm'}} \times \frac{\mu^{j\omega}_{Nm}}{\mu^{j\omega'}_{Nm}} \\ \leq& \varsigma^{-2} \max_{m\in\mathscr{R}_0} \frac{\mu^{0\omega}_{Nm}}{\mu^{0\omega}_{Nm}} \times \frac{\mu^{j\omega}_{Nm}}{\mu^{j\omega'}_{Nm}}.\end{aligned}$$

Thus, if $\max \frac{\mu^{\omega}_{Nm}}{\mu^{\omega'}_{Nm}} > K$, we must have

$$\varsigma^{-2} \max_{m\in\mathscr{R}_0} \frac{\mu^{0\omega}_{Nm}}{\mu^{0\omega}_{Nm}} \times \max_{j\in\{1,2,\cdots,r\}} \max_{m\in\mathscr{R}_j} \frac{\mu^{j\omega}_{Nm}}{\mu^{j\omega'}_{Nm}} > K$$

Thus, when $K \to \infty$, we must have either

$$\max_{m\in\mathscr{R}_0} \frac{\mu^{0\omega}_{Nm}}{\mu^{0\omega}_{Nm}} \to \infty, \text{ or;}$$
$$\max_{j\in\{1,2,\cdots,r\}} \max_{m\in\mathscr{R}_j} \frac{\mu^{j\omega}_{Nm}}{\mu^{j\omega'}_{Nm}} \to \infty.$$

Then the result follows similar arguments in the case of irreducible automatons.

□

C.10 Proof of proposition 5

Proof. I first prove the first statement. Following Hellman and Cover (1970), we have for all ω'

$$\begin{aligned}\min_{m\in M} \frac{\mu^{\omega'}_{Nm}}{\mu^{\omega}_{Nm}} &\geq (\bar{l}^{\omega\omega'}_N \bar{l}^{\omega'\omega}_N)^{-(M-1)} \max_{m\in M} \frac{\mu^{\omega'}_{Nm}}{\mu^{\omega}_{Nm}} \\ \min_{m\in M} \frac{\mu^{\omega'}_{Nm}}{\mu^{\omega}_{Nm}} &\geq \varsigma^{2(M-1)} \max_{m\in M} \frac{\mu^{\omega'}_{m}}{\mu^{\omega}_{m}} \\ \min_{m\in M} \frac{\mu^{\omega'}_{Nm}}{\mu^{\omega}_{Nm}} &\geq \varsigma^{2(M-1)} \min \max_{m\in M} \frac{\mu^{\omega'}_{Nm}}{\mu^{\omega}_{Nm}} \\ \min_{m\in M} \frac{\mu^{\omega'}_{Nm}}{\mu^{\omega}_{Nm}} &\geq \varsigma^{2(M-1)} \\ \frac{u^{\omega'}_N p^{\omega'}_N \mu^{\omega'}_{Nm}}{u^{\omega}_N p^{\omega}_N \mu^{\omega}_{Nm}} &\geq \varsigma^{2(M-1)} \frac{u^{\omega'}_N p^{\omega'}_N}{u^{\omega}_N p^{\omega}_N} \text{ for all } m.\end{aligned}$$

Given p^{ω}_N, $\min\max_{\omega'} \frac{u^{\omega'}_N p^{\omega'}_N}{u^{\omega}_N p^{\omega}_N} = \frac{\underline{u}}{\bar{u}} \frac{\frac{1-p^{\omega}_N}{N-1}}{p^{\omega}_N}$. Thus, when p^{ω}_N is small enough, there exists some $\omega' \neq \omega$

$$\frac{u_N^{\omega'} p_N^{\omega'} \mu_{Nm}^{\omega'}}{u_N^{\omega} p_N^{\omega} \mu_{Nm}^{\omega}} \geq \varsigma^{2(M-1)} \min \max_{\omega'} \frac{u_N^{\omega'} p_N^{\omega'}}{u_N^{\omega} p_N^{\omega}} \geq \varsigma^{2(M-1)} \frac{\underline{u}}{\overline{u}} \frac{\frac{1-p_N^{\omega}}{N-1}}{p_N^{\omega}} > 1 + A \text{ for all } m \in M.$$

for some $A > 0$. Suppose $M^\omega \neq \emptyset$, it implies that if the DM chooses action ω' instead of action ω memory state m, his asymptotic utility loss decreases by $u_N^{\omega'} p_N^{\omega'} \mu_{Nm}^{\omega'} - u_N^{\omega} p_N^{\omega} \mu_{Nm}^{\omega} \geq 0$. Thus when $\epsilon \to 0$, we have either $M^\omega = \emptyset$ or $\max_{\omega'} u_N^{\omega'} p_N^{\omega'} \mu_{Nm}^{\omega'} - u_N^{\omega} p_N^{\omega} \mu_{Nm}^{\omega} \to 0$. As $\max_{\omega'} u_N^{\omega'} p_N^{\omega'} \mu_{Nm}^{\omega'} - u_N^{\omega} p_N^{\omega} \mu_{Nm}^{\omega} > A u_N^{\omega} p_N^{\omega} \mu_{Nm}^{\omega}$, $\max_{\omega'} u_N^{\omega'} p_N^{\omega'} \mu_{Nm}^{\omega'} - u_N^{\omega} p_N^{\omega} \mu_{Nm}^{\omega} \to 0$ also implies $u_N^{\omega} p_N^{\omega} \mu_{Nm}^{\omega} \to 0$ and $\max_{\omega'} u_N^{\omega'} p_N^{\omega'} \mu_{Nm}^{\omega'} \to 0$. Therefore, $u_N^{\omega'} p_N^{\omega'} \mu_{Nm}^{\omega'} \to 0$ for all ω'. As M^ω is finite and $u_N^{\omega'} p_N^{\omega'} > 0$ for ω', $\sum_{m \in M^\omega} \mu_{Nm}^{\omega'} \to 0$ for all ω'.

I now prove the second statement. Note that L_{NM}^* is given by the following minimization problem. Denote $\beta_N^\omega = 1 - \sum_{m \in M^\omega} \mu_{Nm}^\omega$,

$$L_{NM}^* = \min \sum_{\omega=1}^{N} u_N^\omega p_N^\omega \beta_N^\omega$$
$$\text{subject to } (\beta_N^\omega)_{\omega=1}^N \in \text{cl}(\mathscr{M})$$

where $\mathscr{M}$ is the feasibility set and $\text{cl}(\mathscr{M})$ is the closure of $\mathscr{M}$. As an example, when $N = M = 2$, $\mathscr{M}$ is characterized by equation 7. We could also represent the problem as a maximization problem:

$$\max \sum_{\omega=1}^{N} u_N^\omega p_N^\omega \alpha_N^\omega$$
$$\text{subject to } (\alpha_N^\omega)_{\omega=1}^N \in \text{cl}(\tilde{\mathscr{M}})$$

where $(\alpha_N^\omega)_{\omega=1}^N \in \text{cl}(\tilde{\mathscr{M}})$ if and only if $(1 - \alpha_N^\omega)_{\omega=1}^N \in \text{cl}(\mathscr{M})$.

Suppose for some $(u_N^{\omega'}, p_N^{\omega'})_{\omega'=1}^N$, ϵ-optimal updating mechanisms must ignore state ω as $\epsilon \to 0$. Denote $(\alpha_N^{\omega' *})_{\omega'=1}^N$ as the solution of the maximization problem. It implies that $\alpha_N^{\omega *} = 0$ and

$$\sum_{\omega' \neq \omega} u_N^{\omega'} p_N^{\omega'} \alpha_N^{\omega' *} > \sum_{\omega'=1}^{N} u_N^{\omega'} p_N^{\omega'} \alpha_N^{\omega'}$$

for all $(\alpha_N^{\omega'})_{\omega'=1}^N \in \text{cl}(\mathscr{M})$ where $\alpha_N^\omega > 0$. Rearranging the inequality gives us

$$\sum_{\omega' \neq \omega} \frac{u_N^{\omega'} p_N^{\omega'}}{u_N^\omega p_N^\omega} (\alpha_N^{\omega' *} - \alpha_N^{\omega'}) > \alpha_N^\omega.$$

Now, without loss of generality assume $\omega \neq 1$,

$$\begin{aligned}\sum_{\omega'\neq\omega}\frac{\tilde{u}_N^{\omega'}\tilde{p}_N^{\omega'}}{\tilde{u}_N^{\omega}\tilde{p}_N^{\omega}}(\alpha_N^{\omega'*}-\alpha_N^{\omega'}) &= \sum_{\omega'\neq\omega}\frac{\tilde{u}_N^{\omega'}\tilde{p}_N^{\omega'}}{\tilde{u}_N^{1}\tilde{p}_N^{1}}\times\frac{\tilde{u}_N^{1}\tilde{p}_N^{1}}{\tilde{u}_N^{\omega}\tilde{p}_N^{\omega}}(\alpha_N^{\omega'*}-\alpha_N^{\omega'})\\ &= \sum_{\omega'\neq\omega}\frac{u_N^{\omega'}p_N^{\omega'}}{u_N^{1}p_N^{1}}\times\frac{\tilde{u}_N^{1}\tilde{p}_N^{1}}{\tilde{u}_N^{\omega}\tilde{p}_N^{\omega}}(\alpha_N^{\omega'*}-\alpha_N^{\omega'})\\ &= \frac{\tilde{u}_N^{1}\tilde{p}_N^{1}}{\tilde{u}_N^{\omega}\tilde{p}_N^{\omega}}\times\frac{u_N^{\omega}p_N^{\omega}}{u_N^{1}p_N^{1}}\times\sum_{\omega'\neq\omega}\frac{u_N^{\omega'}p_N^{\omega'}}{u_N^{1}p_N^{1}}\times\frac{u_N^{1}p_N^{1}}{u_N^{\omega}p_N^{\omega}}(\alpha_N^{\omega'*}-\alpha_N^{\omega'})\\ &> \sum_{\omega'\neq\omega}\frac{u_N^{\omega'}p_N^{\omega'}}{u_N^{1}p_N^{1}}\times\frac{u_N^{1}p_N^{1}}{u_N^{\omega}p_N^{\omega}}(\alpha_N^{\omega'*}-\alpha_N^{\omega'})\\ &> \alpha_N^{\omega}.\end{aligned}$$

It implies that

$$\sum_{\omega'\neq\omega}\tilde{u}_N^{\omega'}\tilde{p}_N^{\omega'}\alpha_N^{\omega'*} > \sum_{\omega'=1}^{N}\tilde{u}_N^{\omega'}\tilde{p}_N^{\omega'}\alpha_N^{\omega'}$$

for all $(\alpha_N^{\omega'})_{\omega'=1}^N \in \mathrm{cl}(\mathscr{M})$ where $\alpha_N^{\omega} > 0$. Thus ϵ'-optimal updating mechanisms must ignore state ω as $\epsilon' \to 0$. □

C.11 Proof of proposition 6

Proof. Suppose to the contrary that there exists a sequence of subset of states $\tilde{N}$ where $\lim_{N\to\infty}\frac{|\tilde{N}|}{N} > 0$ such that for all $\omega \in \tilde{N}$, the DM takes action ω with some strictly positive probability in some state ω' (where ω' could be different for different ω). Formally, it implies that for all $\omega \in \tilde{N}$

$$\lim_{N.M\to\infty}\sum_{m\in M^\omega}\mu_{Nm}^{\omega'} \geq \xi > 0$$

for some ω'. As M^ω must be finite for almost all ω measured in fraction, it implies that there exists some $\tilde{N}'$ where $\lim_{N\to\infty}\frac{|\tilde{N}'|}{N} > 0$ for all $\omega \in \tilde{N}$ there exists some $\omega' \in \Omega$,

$$\lim_{N.M\to\infty}\mu_{Nm}^{\omega'} \geq \xi' > 0 \text{ for some } m \in M^\omega.$$

Now denote the set of ω', the set of states of the world where the DM picks some action $\omega \in \tilde{N}'$ with strictly positive probability as $\tilde{N}''$. We must have $\lim_{N,M\to\infty}\frac{|\tilde{N}''|}{N} > 0$. Otherwise, there must exist some $\omega' \in \tilde{N}''$ such that there are infinitely many memory states with strictly positive probability and it contradicts the fact that $\lim_{N,M\to\infty}\sum_{m=1}^M \mu_{Nm}^{\omega'} = 1$. With similar arguments, for almost all $\omega \in \tilde{N}'$, there must exists some $m \in M^\omega$ such that $\lim_{N,M\to\infty}\mu_{Nm}^{\omega'} \geq \xi'$ for some $\omega' \in \tilde{N}''$ but $\lim_{N,M\to\infty}\mu_{Nm}^{\omega''} = 0$ for almost all states $\omega'' \in \tilde{N}\setminus\{\omega'\}$. It implies that there exists some sequence of subset of states $\tilde{\tilde{N}}$ where $\lim_{N\to\infty}\frac{|\tilde{\tilde{N}}|}{N}>0$ such that

$$\lim_{N,M\to\infty}\max_{m}\frac{\mu^{\omega'}_{Nm}}{\mu^{\omega''}_{Nm}}=\infty$$

for all $\omega'\in\tilde{\tilde{N}}$ and all $\omega''\in\tilde{\tilde{N}}\setminus\{\omega'\}$. But that is proved to be impossible in the proof of proposition 4. The result thus follows.

□

C.12 Proof of corollary 5

Proof. By proposition 5, if p^{ω}_{NA} is small enough for all $\omega\in N_A\subset N$ and p^{ω}_{NB} is small enough for all $\omega\in N_B=N\setminus N_A$, individual A never picks action ω for all $\omega in N_A$ and individual B never picks action ω for all $\omega\in N\setminus N_A$. Therefore, they must disagree with each other. □

C.13 Proof of corollary 6

Proof. Consider an example where $\lim_{N\to\infty}p^{\omega}_{N}=0$ for all ω. As shown in the proposition 6, the DM must ignore almost all actions when N,M goes to infinite. Thus, $\lim_{N,M\to\infty}L^{*}_{NM}=\lim_{N\to\infty}\sum_{\omega}u^{\omega}_{N}p^{\omega}_{N}$ and all updating mechanism is ϵ-optimal for any $\epsilon\geq 0$. Thus, if individual A adopts an updating mechanism with $d(m)=1$ for all m and individual B adopts an updating mechanism with $d(m)=2$ for all m, then they must disagree with each other under all ω. □

C.14 Proof of proposition 7

Proof. First, as $M=1$ for individual A, his action is constant in all periods for all signal realizations. The optimal automaton is thus $M^2=M^3=\emptyset$ and $a^A_t=1$ for all t. Now I characterize the(almost) optimal updating mechanism of individual B. With some abuse of notations, denote $L^{*}_{32}(nn')$ as optimal utility loss where the DM chooses action n in memory state 1 and action n' in memory state 2. Building on results in Hellman and Cover (1970), we have

$$\begin{aligned}
L^{*}_{32}(11)&=\frac{2}{3}-2\nu,\\
L^{*}_{32}(22)=L^{*}_{32}(33)&=\frac{2}{3}+\nu,\\
L^{*}_{32}(12)=L^{*}_{32}(13)&=\frac{1}{3}-\nu+\frac{2\sqrt{(1+\tau)\left(\frac{1}{3}+2\nu\right)\left(\frac{1}{3}-\nu\right)}-\left(\frac{2}{3}-\nu\right)}{\tau},\\
L^{*}_{32}(23)&=\frac{1}{3}+2\nu+\frac{2\left(\frac{1}{3}-\nu\right)\sqrt{1+\Upsilon}-\left(\frac{2}{3}-2\nu\right)}{\Upsilon}.
\end{aligned}$$

where $L^*_{32}(22) = L^*_{32}(33) > L^*_{32}(11) \geq L^*_{32}(12) = L^*_{32}(13)$. I first prove $L^*_{32}(12) > L^*_{32}(23)$ if and only if ν is small enough. First, $L^*_{32}(12) > L^*_{32}(23)$ if and only if

$$\triangle L_{32}(12{-}23) = 3\nu + \frac{2(\frac{1}{3}-\nu)\sqrt{1+\Upsilon}}{\Upsilon} - \frac{\frac{2}{3}-2\nu}{\Upsilon} - \frac{2\sqrt{(1+\tau)(\frac{1}{3}+2\nu)(\frac{1}{3}-\nu)}}{\tau} + \frac{\frac{2}{3}+\nu}{\tau} < 0.$$

When $\nu = 0$,

$$\triangle \triangle L_{32}(12-23) = \frac{2}{3}\left(\frac{\sqrt{1+\Upsilon}}{\Upsilon} - \frac{\sqrt{1+\tau}}{\tau}\right) - \frac{2}{3}\left(\frac{1}{\Upsilon} - \frac{1}{\tau}\right).$$

As both $\frac{\sqrt{1+x}}{x}$ and $\frac{1}{x}$ decreases in x, and $\Upsilon > \tau$, $\triangle L_{32}(12-23) < 0$, i.e., $L^*_{32}(12) > L^*_{32}(23)$, which by continuity proves the result. $\square$