Abstract—We consider an infinite collection of agents who make decisions, sequentially, about an unknown underlying binary state of the world. Each agent, prior to making a decision, receives an independent private signal whose distribution depends on the state of the world. Moreover, each agent also observes the decisions of its last $K$ immediate predecessors. We study conditions under which the agent decisions converge to the correct value of the underlying state.

We focus on the case where the private signals have bounded information content and investigate whether learning is possible, that is, whether there exist decision rules for the different agents that result in the convergence of their sequence of individual decisions to the correct state of the world. We first consider learning in the almost sure sense and show that it is impossible, for any value of $K$. We then explore the possibility of convergence in probability of the decisions to the correct state. Here, a distinction arises: if $K = 1$, learning in probability is impossible under any decision rule, while for $K \geq 2$, we design a decision rule that achieves it.

I. INTRODUCTION

In this paper, we study variations and extensions of a model introduced and studied in Cover’s seminal work [5]. We consider a Bayesian binary hypothesis testing problem over an “extended tandem” network architecture whereby each agent $n$ makes a binary decision $x_n$, based on an independent private signal $s_n$ (with a different distribution under each hypothesis) and on the decisions $x_{n-1}, \ldots, x_{n-K}$ of its $K$ immediate predecessors, where $K$ is a positive integer constant. We are interested in the question of whether learning is achieved, that is, whether the sequence $\{x_n\}$ correctly identifies the true hypothesis (the “state of the world,” to be denoted by $\theta$), almost surely or in probability, as $n \to \infty$. For $K = 1$, this coincides with the model introduced by Cover [5] under a somewhat different interpretation, in terms of a single memory-limited agent who acts repeatedly but can only remember its last decision.

At a more broad level, our work is meant to shed light on the question whether distributed information held by a large number of agents can be successfully aggregated in a decentralized and bandwidth-limited manner. Consider a situation where each of a large number of agents has a noisy signal about an unknown underlying state of the world $\theta$. This state of the world may represent an unknown parameter monitored by decentralized sensors, the quality of a product, the applicability of a therapy, etc. If the individual signals are independent and the number of agents is large, collecting these signals at a central processing unit would be sufficient for inferring (“learning”) the underlying state $\theta$. However, because of communication or memory constraints, such centralized processing may be impossible or impractical. It then becomes of interest to inquire whether $\theta$ can be learned under a decentralized mechanism where each agent communicates a finite-valued summary of its information (e.g., a purchase or voting decision, a comment on the success or failure of a therapy, etc.) to a subset of the other agents, who then refine their own information about the unknown state.

Whether learning will be achieved under the model that we study depends on various factors, such as the ones discussed next:

(a) As demonstrated in [5], the situation is qualitatively different depending on certain assumptions on the information content of individual signals. We will focus exclusively on the case where each signal has bounded information content, in the sense that the likelihood ratio associated with a signal is bounded away from zero and infinity — the so called Bounded Likelihood Ratio (BLR) assumption. The reason for our focus is that in the opposite case (of unbounded likelihood ratios), the learning problem is much easier; indeed, [5] shows that almost sure learning is possible, even if $K = 1$.

(b) An aspect that has been little explored in the prior literature is the distinction between different learning modes, learning almost surely or in probability. We will see that the results can be different for these two modes.

(c) The results of [5] suggest that there may be a qualitative difference depending on the value of $K$. Our work will shed light on this dependence.

We next provide here a summary of our main results, together with comments on their relation to prior works. In what follows, we use the term decision rule to refer to the mapping from an agent’s information to its decision and the term decision profile to refer to the collection of the agents’ decision rules. Unless there is a statement to the contrary, all results mentioned below are derived under the Bounded Likelihood Ratio assumption.

(a) Almost sure learning is impossible (Theorem 1). For any $K \geq 1$, we prove that there exists no decision profile that guarantees almost sure convergence of the sequence $\{x_n\}$ of decisions to the state of the world $\theta$. This provides an interesting contrast with the case where the BLR assumption does not hold; in the latter
case, almost sure learning is actually possible [5].

(b) Learning in probability is impossible if $K = 1$ (Theorem 2). This strengthens a result of Koplowitz [10] who showed the impossibility of learning in probability for the case where $K = 1$ and the private signals $s_n$ are i.i.d. Bernoulli random variables.

(c) Learning in probability is possible if $K \geq 2$ (Theorem 3). For the case where $K \geq 2$, we provide a fairly elaborate decision profile that yields learning in probability.

Our paper contributes to a large and growing literature on decentralized information aggregation. Two seminal contributions to this literature are [5] and [10], which studied information aggregation by an infinite population of agents arranged in a tandem. Each sensor receives a noisy signal relevant to the underlying state of the world and summarizes its information in a message to the next agent that can take finitely many values. These papers show that for the case of the tandem configuration and under the Bounded Likelihood ratio assumption a ternary message is necessary to establish convergence of the decisions of a subsequence of the agents to the correct state of the world. The tandem configuration was also analyzed in [9] where a lower bound for the eventual probability of error for homogeneous decision profiles is obtained, and also in [12], [11] for the case of “myopically” acting agents, for the cases of binary and ternary messages, respectively.

Another class of decentralized information aggregation problems was introduced in [16], which studies a fusion center solving a binary hypothesis testing problem based on finite-valued messages received from decentralized sensors. Other papers in this area, including [8] and [7], study the same problem on more general network structures. Another strand of literature, in economics, focuses on sequential learning by strategic Bayesian agents (see [4], [3] and [14]). These papers typically focus on the specific observation structure, where each agent observes the decisions of all previous agents. An exception is the recent work [1] which studies sequential learning by strategic agents under general network topologies (see also [15] and [2]).

II. THE MODEL AND PRELIMINARIES

In this section we present the observation model (illustrated in Figure 1) and introduce our basic terminology and notation.

A. The observation model

We consider an infinite sequence of agents, indexed by $n \in \mathbb{N}$, where $\mathbb{N}$ is the set of natural numbers. There is an underlying state of the world $\theta \in \{0, 1\}$, which is modeled as a random variable whose value is unknown to the agents. To simplify notation, we assume that both of the underlying states are a priori equally likely, that is, $P(\theta = 0) = P(\theta = 1) = 1/2$.

Each agent $n$ forms posterior beliefs about this state based on a private signal that takes values in a set $S$, and also by observing the decisions of its $K$ immediate predecessors. We denote by $s_n$ the random variable representing agent $n$’s private signal, while we use $s$ to denote specific values in $S$. Conditional on the state of the world $\theta$ being equal to zero (respectively, one), the private signals are independent random variables distributed according to a probability measure $F_0$ (respectively, $F_1$) on the set $S$. Throughout the paper, the following two assumptions will always remain in effect. First, $F_0$ and $F_1$ are absolutely continuous with respect to each other, implying that no signal value can be fully revealing about the correct state. Second, $F_0$ and $F_1$ are not identical, so that the private signals can be informative.

Each agent $n$ is to make a decision, denoted by $x_n$, which takes values in $\{0, 1\}$. The information available to agent $n$ consists of its private signal $s_n$ and the random vector

$$V_n = (x_{n-K}, \ldots, x_{n-1})$$

of decisions of its $K$ immediate predecessors. (For notational convenience, an agent $i$ with index $i \leq 0$ is identified with agent 1.) The decision $x_n$ is made according to a decision rule $d_n : \{0, 1\}^K \times S \rightarrow \{0, 1\}$:

$$x_n = d_n(V_n, s_n).$$

A decision profile is a sequence $d = \{d_n\}_{n \in \mathbb{N}}$ of decision rules. Given a decision profile $d$, the sequence $X = \{x_n\}_{n \in \mathbb{N}}$ of agent decisions is a well defined stochastic process, described by a probability measure to be denoted by $P_d$, or simply by $P$ if $d$ has been fixed. For notational convenience, we also use $P^j(\cdot)$ to denote the conditional measure under the state of the world $j$, that is

$$P^j(\cdot) = P(\cdot \mid \theta = j).$$

It is also useful to consider randomized decision rules, whereby the decision $x_n$ is determined according to $x_n = d_n(z_n, v_n, s_n)$, where $z_n$ is an exogenous random variable which is independent for different $n$ and also independent of $\theta$ and $(v_n, s_n)$. (The construction in Section V will involve a randomized decision rule.)

B. An assumption and the definition of learning

As mentioned in the Introduction, we focus on the case where every possible private signal value has bounded information content. The assumption that follows will remain in
effect throughout the paper and will not be stated explicitly in our results.

Assumption 1: (Bounded Likelihood Ratios — BLR).
There exist some \( m > 0 \) and \( M < \infty \), such that the Radon-Nikodym derivative \( \frac{d\mathbb{F}_0}{d\mathbb{F}_1} \) satisfies
\[
m < \frac{d\mathbb{F}_0}{d\mathbb{F}_1}(s) < M,\]
for almost all \( s \in S \) under the measure \( (\mathbb{F}_0 + \mathbb{F}_1)/2 \).

We study two different types of learning. As will be seen in the sequel, the results for these two types are, in general, different.

Definition 1: We say that a decision profile \( d \) achieves \textbf{almost sure learning} if
\[
\lim_{n \to \infty} x_n = \theta, \quad \mathbb{P}_d\text{-almost surely,}
\]
and that it achieves \textbf{learning in probability} if
\[
\lim_{n \to \infty} \mathbb{P}_d(x_n = \theta) = 1.
\]

III. IMPOSSIBILITY OF ALMOST SURE LEARNING

In this section, we show that almost sure learning is impossible, for any value of \( K \).

Theorem 1: For any given number \( K \) of observed immediate predecessors, there exists no decision profile that achieves almost sure learning.

The rest of this section is devoted to the proof of Theorem 1. We note that the proof does not use anywhere the fact that each agent only observes the last \( m \) predecessors. Furthermore, while the proof is given for the case of deterministic decision rules, the reader can verify that it also applies to the case where randomized decision rules are allowed.

The following lemma is a simple consequence of the BLR assumption and its proof is omitted.

Lemma 1: For any \( u \in \{0, 1\}^K \) and any \( j \in \{0, 1\} \), we have
\[
m \cdot \mathbb{P}^1(x_n = j \mid v_n = u) < \mathbb{P}^0(x_n = j \mid v_n = u)
\leq M \cdot \mathbb{P}^1(x_n = j \mid v_n = u),
\]
where \( m \) and \( M \) are as in Definition 1.

Lemma 1 states that (under the BLR assumption) if under one state of the world some agent \( n \), after observing \( u \), decides 0 with positive probability, then the same must be true with proportional probability under the other state of the world. This proportional dependence of decision probabilities for the two possible underlying states is central to the proof of Theorem 1.

Before proceeding with the proof of the main result we mention the following algebraic fact that will be of use.

Lemma 2: Consider a sequence \( \{q_n\}_{n \in \mathbb{N}} \) of real numbers, with \( q_n \in [0, 1] \), for all \( n \in \mathbb{N} \). Then,
\[
1 - \sum_{n \in V} q_n \leq \prod_{n \in V} (1 - q_n) \leq e^{-\sum_{n \in V} q_n},
\]
for any \( V \subseteq \mathbb{N} \).

We are now ready to prove the main result of this section.

Proof: [Proof of Theorem 1] Let \( U \) denote the set of all binary sequences with a finite number of zeros (equivalently, the set of binary sequences that converge to one). Suppose, to derive a contradiction, that we have almost sure learning. Then, \( \mathbb{P}^1(x \in U) = 1 \). The set \( U \) is easily seen to be countable, which implies that there exists an infinite binary sequence \( u = \{u_n\}_{n \in \mathbb{N}} \) such that \( \mathbb{P}^1(x = u) > 0 \). In particular,
\[
\mathbb{P}^0(x_k = u_k, \text{ for all } k < n) > 0, \quad \text{for all } n \in \mathbb{N}.
\]

Since \( (x_1, x_2, \ldots, x_n) \) is determined by \( (s_1, s_2, \ldots, s_n) \) and since the distributions of \( (s_1, s_2, \ldots, s_n) \) under the two hypotheses are absolutely continuous with respect to each other, it follows that
\[
\mathbb{P}^0(x_k = u_k, \text{ for all } k \leq n) > 0, \quad \text{for all } n \in \mathbb{N}. \tag{2}
\]

We define
\[
a_n^0 = \mathbb{P}^0(x_n \neq u_n \mid x_k = u_k, \text{ for all } k < n),
\]
\[
a_n^1 = \mathbb{P}^1(x_n \neq u_n \mid x_k = u_k, \text{ for all } k < n).
\]

Lemma 1 implies that
\[
m a_n^1 < a_n^0 < Ma_n^1, \tag{3}
\]
because for \( j \in \{0, 1\} \), \( \mathbb{P}^j(x_n \neq u_n \mid x_k = u_k, \text{ for all } k < n) = \mathbb{P}^j(x_n \neq u_n \mid x_k = u_k, \text{ for } k = n - K, \ldots, n - 1) \).

Suppose that
\[
\sum_{n=1}^{\infty} a_n^1 = \infty.
\]

Then, a variant of the Borel-Cantelli Lemma implies that the event \( \{x_k \neq u_k, \text{ for some } k \} \) has probability 1, under \( \mathbb{P}^1 \). Therefore, \( \mathbb{P}^1(x = u) = 0 \), which contradicts the definition of \( u \).

Suppose now that \( \sum_{n=1}^{\infty} a_n^1 < \infty \). Then,
\[
\sum_{n=1}^{\infty} a_n^0 < M \sum_{n=1}^{\infty} a_n^1 < \infty,
\]
and
\[
\lim_{N \to \infty} \sum_{n=N}^{\infty} \mathbb{P}^0(x_n \neq u_n \mid x_k = u_k, \text{ for all } k < n)
= \lim_{N \to \infty} \sum_{n=N}^{\infty} a_n^0 = 0.
\]

Choose some \( \hat{N} \) such that
\[
\sum_{n=\hat{N}}^{\infty} \mathbb{P}^0(x_n \neq u_n \mid x_k = u_k, \text{ for all } k < n) < \frac{1}{2}.
\]

7439
Then,
\[ P^0(x = u) = P^0(x_k = u_k, \text{ for all } k < N) \cdot \prod_{n=N}^{\infty} (1 - P^0(x_n \neq u_n | x_k = u_k, \text{ for all } k < n)). \]
The first term on the right-hand side is positive by (2), while
\[ \prod_{n=N}^{\infty} (1 - P^0(x_n \neq u_n | x_k = u_k, \text{ for all } k < n)) \]
\[ \geq 1 - \sum_{n=N}^{\infty} P^0(x_n \neq u_n | x_k = u_k, \text{ for all } k < n) > \frac{1}{2}. \]
Combining the above, we obtain \( P^0(x = u) > 0 \) and
\[ \lim_{n \to \infty} P^0(x_n = 1) \geq P^0(x = u) > 0, \]
which contradicts almost sure learning and completes the proof.

Given Theorem 1, in the rest of the paper we concentrate exclusively on the weaker notion of learning in probability, as defined in Section II-B.

IV. NO LEARNING IN PROBABILITY WHEN \( K = 1 \)

In this section, we consider the case where \( K = 1 \), so that each agent only observes the decision of its immediate predecessor. Our main result, stated next, shows that learning in probability is not possible.

Theorem 2: If \( K = 1 \), there exists no decision profile that achieves learning in probability.

We fix a decision profile and use a Markov chain to represent the evolution of the decision process under a particular state of the world. In particular, we consider a two-state Markov chain whose state is the observed decision \( x_{n-1} \). A transition from state \( i \) to state \( j \) for the Markov chain associated with \( \theta = l \), where \( i, j \in \{0, 1\} \), corresponds to agent \( n \) taking the decision \( j \) given that its immediate predecessor \( n - 1 \) decided \( i \), under the state \( \theta = l \). The Markov property is satisfied because the decision \( x_n \), conditional on the immediate predecessor’s decision, is determined by \( s_n \) and hence is (conditionally) independent from the history of previous decisions. Since a decision profile \( d \) is fixed, we can again suppress \( d \) from our notation and define the transition probabilities of the two chains by
\[ a_{ij}^n = P^0(x_n = j | x_{n-1} = i) \]
\[ \bar{a}_{ij}^n = P^1(x_n = j | x_{n-1} = i), \]
where \( i, j \in \{0, 1\} \). The two chains are illustrated in Fig. 2.

Fig. 2: The Markov chains that model the decision process for \( K = 1 \). States represent observed decisions. The transition probabilities under \( \theta = 0 \) or \( \theta = 1 \) are given by \( a_{ij}^n \) and \( \bar{a}_{ij}^n \), respectively. If learning in probability is to occur, the probability mass needs to become concentrated on the highlighted state.

Lemma 3: If we have learning in probability, then
\[ \sum_{n=1}^{\infty} a_{01}^n = \infty, \]
and
\[ \sum_{n=1}^{\infty} a_{10}^n = \infty. \]

Proof: For the sake of contradiction, assume that \( \sum_{n=1}^{\infty} a_{01}^n < \infty \). By Eq. 6, we also have \( \sum_{n=1}^{\infty} \bar{a}_{01}^n < \infty \). Then, the expected number of transitions from state 0 to state 1 is finite under either state of the world. In particular the (random) number of such transitions is finite, almost surely. This can only happen if \( \{x_n\}_{n=1}^{\infty} \) converges almost surely. However, almost sure convergence together with learning in probability would imply almost sure learning, which would contradict Theorem 1. The proof of the second statement in the lemma is similar.

The next lemma states that if we have learning in probability, then the transition probabilities between different states should converge to zero.

Lemma 4: If we have learning in probability, then
\[ \lim_{n \to \infty} a_{01}^n = 0. \]

Proof: Assume, to arrive at a contradiction that there exists some \( \epsilon \in (0, 1) \) such that
\[ a_{01}^n = P^0(x_n = 1 | x_{n-1} = 0) > \epsilon, \]
for infinitely many values of \( n \). Since we have learning in probability, we also have \( P^0(x_{n-1} = 0) > 1/2 \) when \( n \) is large enough. This implies that for infinitely many values of \( n \),
\[ P^0(x_n = 1) \geq P^0(x_n = 1 | x_{n-1} = 0)P^0(x_{n-1} = 0) \geq \frac{\epsilon}{2}. \]
But this contradicts learning in probability.

We are now ready to complete the proof of Theorem 2, by arguing as follows. Since the transition probabilities from state 0 to state 1 converge to zero, while their sum is infinite, under either state of the world, we can divide the agents...
and combining with Eq. (6), we also have
\[
\frac{m}{2M} \leq \sum_{n \in A_k} \bar{a}_n^{01} \leq \frac{2}{3},
\] (14)
for all \(k\).

We consider two cases for the sum of transition probabilities from state 1 to state 0 during block \(A_k\). We first assume that
\[
\sum_{n \in A_k} a_n^{10} > \frac{1}{2}.
\]
Using Eq. (6), we obtain
\[
\sum_{n \in A_k} \bar{a}_n^{10} > \sum_{n \in A_k} \frac{1}{M} \cdot a_n^{10} > \frac{1}{2M}.
\] (15)
The probability of a transition from state 1 to state 0 during the block \(A_k\), under \(\theta = 1\), is
\[
P^0(\bigcup_{n \in A_k} \{x_n = 0\} \mid x_{rk} = 1) = 1 - \prod_{n \in A_k} (1 - \bar{a}_n^{10}).
\]
Using Eq. (15) and Lemma 2, the product on the right-hand side can be bounded from above,
\[
\prod_{n \in A_k} (1 - \bar{a}_n^{10}) \leq e^{-\sum_{n \in A_k} \bar{a}_n^{10}} \leq e^{-1/(2M)},
\]
which yields
\[
P^0(\bigcup_{n \in A_k} \{x_n = 0\} \mid x_{rk} = 1) \geq 1 - e^{-1/(2M)}.
\]
After a transition to state 0 occurs, the probability of staying at that state until the end of the block is bounded below as follows:
\[
P^1(x_{rk+1} = 0 \mid \bigcup_{n \in A_k} \{x_n = 0\}) \geq \prod_{n \in A_k} (1 - \bar{a}_n^{01}).
\]
The right-hand side can be further bounded using Eq. (14) and Lemma 2, as follows:
\[
\prod_{n \in A_k} (1 - \bar{a}_n^{01}) \geq 1 - \sum_{n \in A_k} \bar{a}_n^{01} \geq \frac{1}{3}.
\]
Combining the above and using (12), we conclude that
\[
P^1(x_{rk+1} = 0) \geq P^1(x_{rk+1} = 0 \mid \bigcup_{n \in A_k} \{x_n = 0\}) \cdot P^1(\bigcup_{n \in A_k} \{x_n = 0\} \mid x_{rk} = 1) P^1(x_{rk} = 1)
\]
\[
\geq \frac{1}{3} \cdot \left(1 - e^{-1/(2M)}\right) \cdot \frac{1}{2}.
\]
We now consider the second case and assume that
\[
\sum_{n \in A_k} a_n^{10} \leq \frac{1}{2}.
\]
The probability of a transition from state 0 to state 1 during the block \(A_k\) is
\[
P^0(\bigcup_{n \in A_k} \{x_n = 1\} \mid x_{rk} = 0) = 1 - \prod_{n \in A_k} (1 - \bar{a}_n^{01}).
\]
The product on the right-hand side can be bounded above using Lemma 2,
\[ \prod_{n \in A_k} (1 - a_{n}^{10}) \leq e^{-\sum_{n \in A_k} a_{n}^{10}} \leq e^{-m/(2M)}, \]
which yields
\[ P^0 \left( \bigcup_{n \in A_k} \{x_n = 1 \mid x_{r_k} = 0 \} \right) \geq 1 - e^{-m/2}. \]
After a transition to state 1 occurs, the probability of staying at that state until the end of the block is bounded from below as follows:
\[ P^0 \left( x_{r_k+1} = 1 \mid \bigcup_{n \in A_k} \{x_n = 1 \} \right) \geq \prod_{n \in A_k} (1 - a_n^{10}). \]
The right-hand side can be bounded using Eq. (14) and Lemma 2, as follows:
\[ \prod_{n \in A_k} (1 - a_n^{10}) \geq 1 - \sum_{n \in A_k} a_n^{10} \geq \frac{1}{2}. \]
Using also Eq. (11), we conclude that
\[ \liminf_{n \to \infty} P_d(x_n \neq \theta) \geq \frac{1}{2} \min \left\{ \frac{1}{6} \left(1 - e^{1/(2M)} \right), \frac{1}{4} \left(1 - e^{-m/2} \right) \right\} > 0 \]
which contradicts learning in probability and concludes the proof.

Once more, we note that the proof and the result remain valid for the case where randomized decision rules are allowed.

The coupling between the Markov chains associated with the two states of the world is central to the proof of Theorem 2. The importance of the BLR assumption is highlighted by the observation that if either \( m = 0 \) or \( M = \infty \), then the lower bound obtained in (16) is zero, and the proof fails. The next section shows that a similar argument cannot be made to work when \( K \geq 2 \). In particular, we construct a decision profile that achieves learning in probability when agents observe the last two immediate predecessors.

V. LEARNING IN PROBABILITY WHEN \( K \geq 2 \)

In this section we show that learning in probability is possible when \( K \geq 2 \), i.e., when each agent observes the decisions of two or more of its immediate predecessors.

A. Reduction to the case of binary observations

We will construct a decision profile that leads to learning in probability, for the special case where the signals \( s_n \) are binary (Bernoulli) random variables with a different parameter under each state of the world. This readily leads to a decision profile that learns, for the case of general signals. Indeed, if the \( s_n \) are general random variables, each agent can quantize its signal, to obtain a quantized signal \( s_n' = h(s_n) \) that takes values in \{0, 1\}. Then, the agents can apply the decision profile for the binary case. The only requirement is that the distribution of \( s_n' \) be different under the two states of the world. This is straightforward to enforce by proper choice of the quantization rule \( h \): for example, we may let \( h(s_n) = 1 \) if and only if \( P(\theta = 1 \mid s_n) > P(\theta = 0 \mid s_n) \). It is not hard to verify that with this construction and under our assumption that the distributions \( P_0 \) and \( P_1 \) are not identical, the distributions of \( s_n' \) under the two states of the world will be different.

We also note that it suffices to construct a decision profile for the case where \( K = 2 \). Indeed, if \( K > 2 \), we can have the agents ignore the actions of all but their two immediate predecessors and employ the decision profile designed for the case where \( K = 2 \).

B. The decision profile

As just discussed, we assume that the signal \( s_n \) is binary. For \( i = 0, 1 \), we let \( p_i = P^0(s_n = 1) \) and \( q_i = 1 - p_i \). We also use \( p \) to denote a random variable that is equal to \( p_i \) if and only if \( \theta = i \). Finally, we let \( \overline{p} = (p_0 + p_1)/2 \) and \( \overline{q} = 1 - \overline{p} = (q_0 + q_1)/2 \). We assume, without loss of generality, that \( p_0 < p_1 \), in which case we have \( p_0 < \overline{p} < p_1 \) and \( q_0 > \overline{q} > q_1 \).

Let \( \{k_m\}_{m \in \mathbb{N}} \) and \( \{r_m\}_{m \in \mathbb{N}} \) be two sequences of positive integers that we will define later in this section. We divide the agents into segments that consist of S-blocks, R-blocks, and transient agents, as follows. We do not assign the first two agents to any segment (and the first segment starts with agent \( n = 3 \)). For segment \( m \in \mathbb{N} \):

(i) the first \( 2k_m - 1 \) agents belong to the block \( S_m \);
(ii) the next agent is an SR transient agent;
(iii) the next \( 2r_m - 1 \) agents belong to the block \( R_m \);
(iv) the next agent is an RS transient agent.

An agent’s information consists of the last two decisions, denoted by \( v_n = (x_{n-2}, x_{n-1}) \), and its own signal \( s_n \). The decision profile is constructed so as to enforce that if \( n \) is the first agent of either an S or R block, then \( v_n = (0, 0) \) or \( (1, 1) \).

(i) Agents 1 and 2 choose 0, irrespective of their private signal.
(ii) During block \( S_m \), for \( m \geq 1 \):

a) If the first agent of the block, denoted by \( n \), observes \((1, 1)\), it chooses 1, irrespective of its private signal. If it observes \((0, 0)\) and its private signal is 1, then
\[ x_n = z_n, \]
where \( z_n \) is an independent Bernoulli random variable with parameter \( 1/m \). If \( z_n = 1 \) we say that a searching phase is initiated. (The cases of observing \((1, 0)\) or \((1, 0)\) will not be allowed to occur.)

b) For the remaining agents in the block:

i) Agents who observe \((0, 1)\) decide 0 for all private signals.
The following fact is used in the proof that follows. Table 1: A decision profile that achieves learning in probability. Note that this is a symmetrical process takes place during block \(S_m\). If agent \(d\) is the first one in an block, as well as for the subsequent SR transient agent, which is agent \(n+2k_m-1\). The latter agent also decides 1, so that the first agent of the block, \(R_m\), observes \(v_{n+2k_m} = (1,1)\).

1) If \(v_n = (1,1)\), then \(v_i = (1,1)\) for all agents \(i\) in the block, which is the case of a block \(S_m\) is entirely symmetrical. Let \(n\) be the first agent of the block, and note that the last agent of the block is \(n+2k_m-2\).

2) If \(v_n = (0,0)\) and \(x_n = 0\), then \(v_i = (0,0)\) for all agents \(i\) in the block, which is not the subsequent SR transient agent, which is agent \(n+2k_m-1\). The latter agent also decides 0, so that the first agent of the next block, \(R_m\), observes \(v_{n+2k_m} = (0,0)\).

3) The interesting case occurs when \(v_n = (0,0)\), \(s_n = 1\), and \(z_n = 1\), so that a search phase is initiated and \(x_n = 1\). \(v_{n+1} = (0,1)\), \(x_{n+1} = 0\), \(v_{n+2} = (1,0)\). Here there are two possibilities:

a) Suppose that for every \(i > n\) in the block \(S_m\), for which \(i-n\) is even (and with \(i\) not the last agent in the block), we have \(s_i = 1\). Then, for \(i-n\) even, we will have \(v_i = (1,0)\), \(x_i = 1\), \(v_{i+1} = (0,1), x_{i+1} = 0, v_{i+2} = (1,0)\), etc. When \(i\) is the last agent of the block, then \(i = n+2k_m-2\), so that \(i-n\) is even, \(v_i = (1,0)\), and \(x_i = 1\). The subsequent SR transient agent, agent \(n+2k_m-1\), sets \(x_{n+2k_m-1} = 1\), so that the first agent of the next block, \(R_m\), observes \(v_{n+2k_m} = (1,1)\).

b) Suppose that for some \(i > n\) in the block \(S_m\), for which \(i-n\) is even, we have \(s_i = 0\). Let \(i\) be the first agent in the block with this property. We have \(v_i = (1,0)\) (as in the previous case), but \(x_i = 0\), so that \(v_{i+1} = (0,0)\). Then, all subsequent decisions in the block, as well as by the next SR transient agent are 0, and the first agent of the next block, \(R_m\), observes \(v_{n+2k_m} = (0,0)\).

**Theorem 3:** Under the decision profile described above there exist parameter choices such that

\[
\lim_{n \to \infty} \mathbb{P}(x_n = \theta) = 1.
\]

**Proof:** [Sketch of proof] To understand the overall effect of our construction, we consider a (non-homogeneous) Markov chain representation of the evolution of decisions. We focus on the subsequence of agents consisting of the first agent of each S- and R-block. By the construction of the decision profile, the state \(v_n\), restricted to this subsequence, can only take values \((0,0)\) or \((1,1)\), and its evolution can be represented by a 2-state Markov chain. The transition probabilities between the states in this Markov chain is given by a product of terms, the number of which is related to the size of the S- and R-blocks. For learning to occur, there has to be an infinite number of switches between the two states in the Markov chain (otherwise getting trapped in an incorrect decision would have positive probability).
Moreover, the probability of these switches should go to zero (otherwise there would be a probability of switching to the incorrect decision that is bounded away from zero). We obtain these features by allowing switches from state $(0,0)$ to state $(1,1)$ during S-blocks and switches from state $(1,1)$ to state $(0,0)$ during R-blocks. By suitably defining blocks of increasing size, we can ensure that the probabilities of such switches remain positive but decay at a desired rate. This can be accomplished by suitable parameter choices.

VI. CONCLUSIONS

We have obtained sharp results on the fundamental limitations of learning by a sequence of agents who only observe the decisions of a fixed number $K$ of immediate predecessors, under the assumption of Bounded Likelihood Ratios. Specifically, we have shown that almost sure learning is impossible whereas learning in probability is possible if and only if $K > 1$.

The scheme in Section V is only of theoretical interest, because the rate at which the probability of error decays to zero is extremely slow. This is quite unavoidable, even for the much more favorable case of unbounded likelihood ratios [13], and we do not consider the problem of improving the convergence rate a promising one.

Finally, one may consider extensions to the case of $m > 2$ hypotheses and $m$-valued decisions by the agents. Our negative results are expected to hold, and the construction of a decision profile that learns when $K \geq m$, is also expected to go through, paralleling a similar extension in [10].

Finally, whether learning will be achieved or not, depends on the way that agents make their decisions. In an engineering setting, one can assume that the agents decision rules are chosen (through an offline centralized process) by a system designer. In contrast, in game-theoretic models, each agent is assumed to be a Bayesian maximizer of an individual objective, based on the available information. We shed light on this dichotomy by considering a special class of individual objectives that incorporate a certain degree of altruism in [6], an expanded, journal version of this paper.

REFERENCES