Teaching and learning in uncertainty

Varun Jog
Department of Electrical and Computer Engineering
University of Wisconsin - Madison, WI 53706
vjog@wisc.edu

***Abstract*—We investigate a simple model for social learning with two agents: a teacher and a student. The teacher's goal is to teach the student the state of the world Θ, however, the teacher herself is not certain about Θ and needs to simultaneously learn it and teach it to the student. We model the teacher's and the student's uncertainty via binary symmetric channels, and employ a simple heuristic decoder at the student's end. We focus on two teaching strategies: a "low effort" strategy of simply forwarding information, and a "high effort" strategy of communicating the teacher's current best estimate of Θ at each time instant. Using tools from large deviation theory, we calculate the exact learning rates for these strategies and demonstrate regimes where the low effort strategy outperforms the high effort strategy. Our primary technical contribution is a detailed analysis of the large deviation properties of the sign of a transient Markov random walk on $\mathbb{Z}$.**

***Index Terms*—Social learning, large deviations, Markov chains**

I. Introduction

Individuals in a society learn about their environments not only though their own experiences but also from communicating with other members of the society. This interaction drives the exchange of ideas, technologies, news, opinions, and is critically important to the social and economic processes in a society. Understanding and predicting the effects of social interaction on society is a hard problem: each individual's opinion is dynamic and depends on their biases, observations, and social interactions. The question of how agents learn through social interactions has received much attention in the past few decades, and a number of mathematical models have been proposed to analyze social learning phenomena [1], [2].

Given a mathematical model for social learning, one is primarily interested in analyzing the following questions: (a) Convergence: Does an agent's opinion eventually converge?; (b) Agreement: Given convergence, do the agents agree? ; (c) Learning: Given agreement, is the unanimous opinion the true state of the world?; and (d) Given learning, how fast does learning take place? Our work in this paper focuses on (d) and is most closely related to the papers [3]–[6]. The specific learning model is motivated by the work in [6], which analyzed the speed of learning in a two agent model. In a Bayesian setting, the authors demonstrated the counterintuitive result that more interaction among agents can in fact impede learning.

The model is this paper also considers a two agent setting as in [6], with two key differences. First, the second agent (whose speed of learning we wish to analyze) does not have any private observations that allow them to learn. Any information they get is via a noisy interaction with the first agent. And second, the agents are not Bayesian but instead perform a heuristic calculation to form their opinions. There is a rich history of studying both Bayesian and non-Bayesian models in social learning, since the assumption of rationality in the former is not always suitable for human agents. Similar to the results in [6], we also demonstrate that certain counterintuitive phenomena occur in our model as well. In particular, we observe that "helpful" social interactions actually slow down the speed of learning.

The main mathematical tools we use in the paper derive from the theory of large deviations. Our primary technical contribution is analyzing the large deviation properties of the sign of a transient Markov random walk on $\mathbb{Z}$. We show the rate function of this process can be explicitly calculated, and moreover it has a surprisingly neat closed-form expression.

The structure of the paper is as follows. In Section II we describe our model in detail and state the problem. In Sections III and IV, we relate the stochastic process generated by social interaction in our model to the sign of a Markov process and analyze this process in detail. Finally, we conclude the paper in Section V with some open problems and discussions.

II. Model description and problem statement

In this paper, we consider a simple model of social learning with two agents: a teacher and a student. Both agents are trying to learn an unknown binary random variable Θ which is called the state of the world. We assume Θ takes values in the set $\{-1, +1\}$ uniformly at random. At each time $i \geq 1$ the teacher observes a noisy version of Θ through a binary symmetric channel with a parameter $p \in [0, 1/2)$; i.e.,

$$\mathbb{P}(O_i = \Theta) = 1 - p, \quad \text{and} \quad \mathbb{P}(O_i = -\Theta) = p.$$

Conditioned on Θ, the random observations $\{O_i\}_{i\geq 1}$ are independent and identically distributed as above. The student does not make any direct observations (noisy or otherwise) of Θ, and may only learn it from the teacher. At each time i, the teacher communicates a binary random variable $\hat{X}_i$ which is a (possibly random) function of the history of observations $\{O_j\}_{1\leq j\leq i}$ and the student receives a noisy version of $\hat{X}_i$, which we call Z_i. The communication channel from the teacher to the student is assumed to be another binary symmetric channel with parameter $q \in [0, 1/2)$. The student's estimate of Θ after observing $\{Z_j\}_{j\leq i}$ is denoted by $\hat{\Theta}_i \in \{-1, +1\}$. We refer to the sequence of random variables

$\{\hat{X}_i\}$ as the teacher's strategy, and the decoding rules $\{\hat{\Theta}_i\}$ as the student's learning strategy. For fixed teaching and learning strategies, the student's rate of learning is defined as follows:

$$\mathcal{R} = \limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left(\hat{\Theta}_n \neq \Theta\right).$$

Notice that the teacher is assured to the learn the state of the world eventually owing to her repeated noisy observations of Θ. The student's learning, on the other hand, depends on both the teacher's strategy as well as her own decoding strategy. In this paper, we will largely focus on a fixed student's strategy called the *majority rule*, defined simply as

$$\hat{\Theta}_n = \begin{cases} +1 & \text{if } \frac{\sum_{i=1}^n \mathbb{1}(Z_i=+1)}{n} \geq \frac{1}{2}, \\ -1 & \text{otherwise.} \end{cases}$$

If the student knows the teacher's strategy, it is optimal for the student to use the maximum likelihood decoder to arrive at her estimate of $\hat{\Theta}_n$. However, as is well-documented in the literature on social learning, a fully rational model often places unreasonable computational demands on Bayesian agents [1]. In such models, assuming non-Bayesian agents serves two goals: it makes the model more realistic by reducing its complexity, and in some cases it also helps make the model mathematically tractable.

In a majority learning model, the student will learn what they hear most often, and therefore the teacher should try to teach the "correct lesson" more often than the "wrong lesson". What are be some natural strategies that the teacher might employ?

Don't teach: A lazy strategy for the teacher is to put $\hat{X}_i = O_i$; i.e., simply forward the teacher's observation to the student. The student then receives Z_n which is effectively a noisy observation of Θ through a BSC($p \star q$), where $p \star q = p(1-q) + q(1-p) = p\bar{q} + q\bar{p}$. The optimal decision rule for the student in this case is simply using the majority rule and declaring $\hat{\Theta}_n$ to be $+1$ if there were more $+1$'s in $\{Z_i\}_{1\leq i\leq n}$, and declare -1 otherwise. The learning rate for this strategy is given by $D(1/2||p\star q)$, where $D(a||b) = a\log(a/b) + \bar{a}\log(\bar{a}/\bar{b})$ is the Kullback Leibler divergence between the Bernoulli(a) and the Bernoulli(b) distributions [7]. The rate of learning for this "lazy" or "low-effort" strategy will be the benchmark against which we want to compare the "helpful" or "high-effort" strategy described below.

Teach the current best guess: In contrast to the lazy strategy of not teaching, the teacher might follow a strategy of always teaching the teacher's current best estimate of Θ, which is obtained by applying the majority rule to her observations $\{O_i\}_{1\leq i\leq n}$. Analyzing the learning rate for this strategy is the main problem we tackle in this paper.

Notice that the high-effort strategy clearly satisfies the property that after some finite time, the process $\{\hat{X}_n\}$ is identically equal to Θ. The lazy-strategy never converges in such a manner, and thus the teacher is correct more often in the high-effort strategy. This is the intuitive reason one might expect the high-effort strategy to dominate. In what follows, we calculate the exact learning rate for this strategy.

III. Analyzing the sign of a random walk

Assume that $\Theta = +1$. The teacher's strategy is a majority rule applied to her observations $\{O_i\}_{1\leq i\leq n}$; i.e., her response $\hat{X}_n$ equals the sign of $\sum_{i=1}^n O_i$. The random process $X_n := \sum_{i=1}^n O_i$ may be modeled as a random walk $\mathbb{Z}$ with transition probabilities as follows:

$$p(X_{n+1} = i+1 | X_n = i) = 1-p, \text{and}$$
$$p(X_{n+1} = i-1 | X_n = i) = p,$$

where $p < 1/2$. Notice that this random walk is transient; i.e. for every $i \in \mathbb{Z}$, the random walk visits state i finitely many times with probability 1. Since $p < 1/2$, the random walk eventually runs off to $+\infty$. Denote the process $\hat{X}_n$ as follows:

$$\hat{X}_n = \begin{cases} +1 & \text{if } X_n > 0, \\ -1 & \text{if } X_n < 0, \\ \hat{X}_{n-1} & \text{if } X_n = 0. \end{cases}$$

Let M_n be the number of times the teacher is correct up to time n; i.e., $M_n := \sum_{i=1}^n \mathbb{1}(\hat{X}_i = +1)$. Our goal is to explore the large deviations behavior of M_n. In particular, we are interested in the probability $\mathbb{P}(M_n/n \approx 1-\delta)$, which calculates how often the teacher is correct up to time n. We expect this probability to be approximately equal to $e^{-nf(\delta)}$, for some suitable exponent $f(\delta)$, and we would like to pinpoint $f(\delta)$ in terms of p and δ.

A. Preliminary calculations for $\{X_n\}$

The random walk $\{X_n\}$ is transient, and thus there is some probability of never returning to state i starting from state i. This probability is independent of i, and it is an easy exercise to show that this equals $1-2p$.

The next quantity we focus on is the sojourn time T, which we define as the time of *first* return to 0 starting from 0. We use the convention that T is positive if the random walk is positive during the sojourn, and otherwise T is negative. Notice that sojourn times can only take even values: $T = 2k$ when the random walk takes k positive steps and k negative steps in total, with only the endpoints of the sojourn being at 0. The probability may be calculated by finding all such paths and multiplying the result by $p^k\bar{p}^k$. The number of such paths is seen to be the $k-1$-th Catalan number [8], $C_{k-1} = \frac{1}{k}\binom{2k-2}{k-1}$. The distribution of T is now given by

$$\mathbb{P}(T=2k) = \begin{cases} p^{|k|}\bar{p}^{|k|}\frac{1}{|k|}\binom{2|k|-2}{|k|-1} & \text{if } 0<|k|<\infty, \\ 0 & \text{if } k=0, \\ 1-2p & \text{if } k=+\infty, \\ 0 & \text{if } k=-\infty. \end{cases}$$

We are interested in the random variable T *conditioned on the event that* $|T| < \infty$, which we call $\tilde{T}$. It is easy to see that $\mathbb{E}\tilde{T} = 0$, and

$$\mathbb{E}|\tilde{T}| = \sum_{k=1}^{\infty} 2\cdot\frac{p^k\bar{p}^k}{2p}C_{k-1}\cdot 2k = 2\sum_{k=0}^{\infty}\frac{p^{k+1}\bar{p}^{k+1}}{p}C_k\cdot(k+1).$$

The generating function of Catalan numbers is given by

$$f(x) = \sum_{k=0}^{\infty} C_k x^k = \frac{1-\sqrt{1-4x}}{2x}.$$

Thus, $\sum_{k=0}^{\infty} C_k(k+1)x^k = \frac{d}{dx}(xf(x)) = \frac{1}{\sqrt{1-4x}}$. Substituting $x = p\bar{p}$, we conclude that $\mathbb{E}|\tilde{T}| = \frac{2\bar{p}}{(\bar{p}-p)}$. Another important quantity we will need is the moment generating function of the random vector $(\tilde{T}, |\tilde{T}|)$. Using the generating function as above, this can be explicitly calculated:

$$\mathbb{E}e^{\lambda_1\tilde{T}+\lambda_2|\tilde{T}|} = \frac{2-\sqrt{1-4p\bar{p}e^{2(\lambda_1+\lambda_2)}}-\sqrt{1-4p\bar{p}e^{2(\lambda_2-\lambda_1)}}}{4p}.$$

Define the log moment generating function as $L(\lambda_1,\lambda_2) = \log \mathbb{E}e^{\lambda_1\tilde{T}+\lambda_2|\tilde{T}|}$. The domain of L is given by the set $\mathcal{D} = \{(x_1,x_2) \mid |x_1| + x_2 < D(1/2||p)\}$. Outside of this set, the function takes the value $+\infty$. We state a lemma that can be easily proved:

Lemma 1: For $(\lambda_1,\lambda_2) \in \mathcal{D}$, we have $L(\lambda_1,\lambda_2) \le \log\frac{1}{2p}$.

Finally, the last ingredient we need is number of returns of the random walk $\{X_n\}$ to 0. This is a geometric random variable with distribution given by

$$\mathbb{P}(G=i) = (2p)^i(1-2p), \quad i \ge 0.$$

B. Large deviation properties of M_n

Let B be the random variable indicating the final visit to state 0. We break up the probability as follows:

$$\mathbb{P}\left(\frac{M_n}{n} \le 1-\delta\right) = \sum_{g=0}^{n/2}\mathbb{P}(M_n \le n(1-\delta), B \le n, G=g) + \mathbb{P}(M_n \le n(1-\delta), B > n).$$

Note that there are less than n terms in this sum, and the largest among these terms will dictate the exponential growth rate of the sum. The final term can be expressed as $\mathbb{P}(B > n)\mathbb{P}(M_n \le n(1-\delta)|B>n)$. Notice that conditioned on the event $\{B>n\}$, the random variable M_n/n has a symmetric distribution around $1/2$. This is because the mirror image of every path up to time n has the exact same probability as the original path when conditioned on the event $\{B>n\}$. We have the bounds $1/2 \le \mathbb{P}(M_n \le n(1-\delta)|B>n) \le 1$, and thus the final term is $\Theta(\mathbb{P}(B>n))$. It is not hard to show that this probability is $e^{-n(D(1/2||p)+o(1))}$, and we skip the details here. We focus on the first $n/2$ terms. We rewrite the probability as follows:

$$\sum_{g=0}^{n/2}\mathbb{P}(M_n \le n(1-\delta), B \le n, G=g) = \left[\sum_{g=0}^{n/2}\mathbb{P}(G=g)\times \left\{\sum_{b=2g}^{n}\sum_{a=-b}^{b}\mathbb{P}\left(\sum_{j=1}^{g}\tilde{T}_j = -a, \sum_{j=1}^{g}|\tilde{T}_j| = b\right)\mathbb{1}\left(\frac{a+b}{2} > n\delta\right)\right\}\right].$$

We explain the indicator function as follows. The total number of $+1$'s received equals $(n-b)$ (the $+1$'s received after time b) plus the total number of $+1$'s received up to time b, which equals $(b-a)/2$. This equals $n-(a+b)/2$, and since we are interested in the event that this quantity is at most $n(1-\delta)$, we introduce the indicator $\mathbb{1}((a+b)/2 > n\delta)$. Substitute $\alpha := a/n, \beta := b/n, \gamma := g/n$. We may rewrite the above as a summation over α, β, and γ, were we implicitly assume that they take values of the form i/n for some integer i:

$$\sum_{\gamma=0}^{1/2}\mathbb{P}\left(\frac{G}{n}=\gamma\right)\times \left\{\sum_{\beta=\max(\delta,2\gamma)}^{1}\sum_{\alpha=2\delta-\beta}^{\beta}\mathbb{P}\left(\frac{\sum_{j=1}^{n\gamma}\tilde{T}_j}{n} = -\alpha, \frac{\sum_{j=1}^{g}|\tilde{T}_j|}{n} = \beta\right)\right\}$$

Our next theorem forms the core result of this paper:

Theorem 1: The following equality holds:

$$\frac{1}{n}\lim_{n\to\infty}\log\mathbb{P}\left(\frac{M_n}{n} \le (1-\delta)\right) = -\delta D(1/2||p).$$

Proof: Define the random vector $Z_n := \left(\frac{G}{n}, \frac{\sum_{j=1}^{G}\tilde{T}_j}{n}, \frac{\sum_{j=1}^{G}|\tilde{T}_j|}{n}\right)$. Define the set $\mathcal{R} := \{\alpha,\beta,\gamma \mid \gamma \in [0,1/2], \beta \in [\max(\delta,2\gamma),1], \alpha \in [2\delta-\beta,\beta]\} \subseteq \mathbb{R}^3$. Note that we are interested in the quantity $\mathbb{P}(Z_n \in \mathcal{R})$. We shall evaluate this probability for large n using the Gartner-Ellis theorem from large deviation theory [7]. The first step is to show that the following limit exists for every $\lambda := (\lambda_1,\lambda_2,\lambda_3) \in \mathbb{R}^3$:

$$\begin{aligned}
\Lambda(\lambda) &:= \lim_{n\to\infty}\frac{1}{n}\log\mathbb{E}e^{n\lambda\cdot Z_n}\\
&= \lim_{n\to\infty}\frac{1}{n}\log\mathbb{E}\exp\left(\lambda_1 G + \lambda_2\left(\sum_{j=1}^{G}\tilde{T}_j\right) + \lambda_3\left(\sum_{j=1}^{G}|\tilde{T}_j|\right)\right)\\
&= \lim_{n\to\infty}\frac{1}{n}\log\mathbb{E}\left[\mathbb{E}\left[\exp\left(\lambda_1 G + \lambda_2\left(\sum_{j=1}^{G}\tilde{T}_j\right) + \lambda_3\left(\sum_{j=1}^{G}|\tilde{T}_j|\right)\right)\Big|G\right]\right]\\
&\overset{(a)}{=} \lim_{n\to\infty}\frac{1}{n}\log\mathbb{E}\left[e^{\lambda_1 G + L(\lambda_2,\lambda_3)G}\right]\\
&= \begin{cases} 0 & \text{if } \lambda_1 + L(\lambda_2,\lambda_3) < \log(1/2p),\\ +\infty & \text{otherwise.}\end{cases}
\end{aligned}$$

where the equality in (a) holds only when $(\lambda_2,\lambda_3) \in \mathcal{D}$, otherwise the limit is $+\infty$. To summarize,

$$\Lambda(\lambda) = \begin{cases} +\infty & \text{if } (\lambda_2,\lambda_3) \in \mathcal{D}^c,\\ +\infty & \text{if } (\lambda_2,\lambda_3) \in \mathcal{D}, \lambda_1 \ge \log(1/2p) - L(\lambda_2,\lambda_3),\\ 0 & \text{otherwise.}\end{cases}$$

Let the domain of Λ be $\mathcal{D}_\Lambda$. Let Λ^* be the convex conjugate of Λ. A direct application of the Gartner-Ellis theorem gives the following upper bound:

$$\limsup_{n\to\infty}\frac{1}{n}\log\mathbb{P}_n(\mathcal{R}) \le -\inf_{z\in\mathcal{R}}\Lambda^*(z) \tag{1}$$

where $\mathbb{P}_n$ is the distribution of Z_n. We evaluate the convex conjugate of Λ at the point $(\gamma,\alpha,\beta) \in \mathbb{R}^3_+$:

$$\begin{aligned}
\Lambda^*(\gamma,\alpha,\beta) &= \sup_{\lambda\in\mathcal{D}_\Lambda}\lambda_1\gamma + \lambda_2\alpha + \lambda_3\beta\\
&\overset{(a)}{=} \sup_{(\lambda_2,\lambda_3)\in\mathcal{D}}\left(\log(1/2p) - L(\lambda_2,\lambda_3)\right)\gamma + \lambda_2\alpha + \lambda_3\beta.
\end{aligned}$$

where in (a) we used that $\gamma \ge 0$ in $\mathcal{R}$. We make the following crucial observations: First, Lemma 1 implies that the coefficient of γ in the above expression is positive, and therefore $\Lambda^*(\gamma,\alpha,\beta)$ is a monotonically increasing function of γ for fixed α and β. Second, the set $\mathcal{R}$ is such that the possible

values of the pair (α, β) can only increase as γ becomes smaller. This implies that if $\gamma_1 < \gamma_2$, then

$$\inf_{(\alpha,\beta):(\gamma_1,\alpha,\beta)\in\mathcal{R}} \Lambda^*(\gamma_1,\alpha,\beta) < \inf_{(\alpha,\beta):(\gamma_2,\alpha,\beta)\in\mathcal{R}} \Lambda^*(\gamma_2,\alpha,\beta),$$

since not only is the left hand smaller for every fixed (α, β), but also the range of possible values of (α, β) on the left hand side contains the range on the right hand side. Thus, the infimum of Λ^* over $\mathcal{R}$ must be when $\gamma = 0$; i.e.,

$$\inf_{(\gamma,\alpha,\beta)\in\mathcal{R}} \Lambda^*(\gamma,\alpha,\beta) = \inf_{\beta\in[\delta,1],\alpha\in[2\delta-\beta,\beta]} \Lambda^*(0,\alpha,\beta).$$

When $\beta \in [\delta, 1]$, and $\alpha \in [\beta - 2\delta, \beta]$, we have

$$\Lambda^*(0,\alpha,\beta) = \sup_{(\lambda_2,\lambda_3)\in\mathcal{D}} \lambda_2\alpha + \lambda_3\beta = \beta D(1/2||p).$$

Hence, we conclude that

$$\begin{aligned}\inf_{(\gamma,\alpha,\beta)\in\mathcal{R}} \Lambda^*(\gamma,\alpha,\beta) &= \inf_{\beta\in[\delta,1],\alpha\in[2\delta-\beta,\beta]} \beta D(1/2||p)\\ &= \delta D(1/2||p).\end{aligned}$$

To complete the proof, we need to establish a lower bound counterpart to inequality (1). This is established via the following lemma, whose proof we skip here:

Lemma 2: The following inequality holds:

$$\liminf_{n\to\infty} \frac{1}{n}\log \mathbb{P}_n(\mathcal{R}) \geq -\delta D(1/2||p).$$

The proof follows by constructing a set of paths that satisfy $M_n < (1-\delta)n$ and explicitly computing the combined probability of these paths. This completes the proof of Theorem 1, and we conclude $\mathbb{P}\left(\frac{M_n}{n} \leq (1-\delta)\right) \approx e^{-n\delta D(1/2||p)}$. ∎

IV. Learning rate for the high-effort strategy

Notice that Theorem 1 already provides the exact learning rate if the teacher to student channel is perfect; i.e., if $q = 0$. In this case, a majority learner student makes an error only if $M_n < n/2$. Substituting $\delta = 1/2$, we see that the student will learn at a rate of $\frac{1}{2}D(1/2||p)$ via high-effort strategy. In contrast, the learning rate for the low-effort strategy is $D(1/2||p)$—more than that of the high-effort strategy!

To evaluate the learning rate with $q > 0$, we first prove the following large deviations result for a mixture of Bernoulli random variables:

Lemma 3: Let $\theta \in [0,1]$ and $q \in [0, 1/2]$. Consider a sequence of i.i.d. Bernoulli$(1-q)$ random variables $\{U_i\}_{1\leq i\leq n-\lfloor n\theta\rfloor}$ and i.i.d. Bernoulli(q) random variables $\{V_j\}_{1\leq j\leq \lfloor n\theta\rfloor}$ such that the U_i's are independent of the V_j's. Define the random variable W_n as follows:

$$W_n := \frac{\sum_{i=1}^{n-\lfloor n\theta\rfloor} U_i + \sum_{j=1}^{\lfloor n\theta\rfloor} V_j}{n}.$$

Then W_n satisfies the large deviation principle with rate function

$$I_\theta(w) = w\log(\eta) - \bar{\theta}\log(\bar{q}\eta + q) - \theta\log(q\eta + \bar{q})$$

where

$$\eta := \frac{-\tau + \sqrt{\tau^2 + 4w\bar{z}}}{2\bar{w}}, \quad \tau := \frac{\bar{q}}{q}(\bar{\theta} - w) + \frac{q}{\bar{q}}(\theta - w).$$

The proof is a direct application of the Gartner-Ellis theorem and we skip it here.

Although we are interested in $\delta = 1/2$, it is not too much work to evaluate the probability that at most $n(1-\delta)$ instances of the student's received sequence $\{Z_n\}$ equal $+1$.

Theorem 2: Let $\delta \in [q, 1-q]$. Suppose we say that the student commits an error if the fraction of received $+1$'s is at most $(1-\delta)$. Then the rate of learning is given by

$$R := \inf_{\theta\in[0,1]} \theta D(1/2||p) + \hat{I}(\theta),$$

where $\hat{I}(\cdot)$ is defined as

$$\hat{I}(\theta) = \inf_{w\in[0,1-\delta]} I_\theta(w),$$

where $I_\theta(\cdot)$ is the rate function from Lemma 3.

Proof: Let $\mathcal{E}_n$ be the error event; i.e. the event when the student receives at most $n(1-\delta)$ number of $+1$'s. Consider an integer $N > 0$, whose value will be decided later. Divide the interval $[0,1]$ into intervals L_i, for $0 \leq i \leq N-1$ such that $L_i := [i/N, i+1/N]$. The probability of error can be written as follows:

$$\mathbb{P}(\mathcal{E}_n) = \sum_{i=0}^{N-1} \mathbb{P}\left(\frac{M_n}{n} \in L_i\right)\mathbb{P}\left(\mathcal{E}_n \middle| \frac{M_n}{n} \in L_i\right).$$

Note that we have the following bounds on the first term:

$$\mathbb{P}\left(\frac{M_n}{n} \in L_i\right) = \mathbb{P}\left(\frac{M_n}{n} \leq \frac{i+1}{N}\right) - \mathbb{P}\left(\frac{M_n}{n} \leq \frac{i}{N}\right).$$

Let $\epsilon_1 > 0$ be an arbitrarily small constant. From Theorem 1, we have that for all $0 \leq i \leq N-1$ and for all large enough n, the following inequality holds:

$$\left| -\frac{1}{n}\log\mathbb{P}\left(\frac{M_n}{n} \in L_i\right) - \frac{N-i-1}{N} D\left(1/2||p\right)\right| < \epsilon_1. \quad (2)$$

Turning to the second term, we note that the probability of error is a monotonically decreasing function of M_n/n. Thus, we may write the following bounds:

$$\mathbb{P}\left(\mathcal{E}_n \middle| \frac{M_n}{n} = \frac{i+1}{N}\right) \leq \mathbb{P}\left(\mathcal{E}_n \middle| \frac{M_n}{n} \in L_i\right) \leq \mathbb{P}\left(\mathcal{E}_n \middle| \frac{M_n}{n} = \frac{i}{N}\right).$$

Let $\epsilon_2 > 0$ be an arbitrarily small constant. Using the large deviations principle from Lemma 3, we have that for all large enough n and for all $0 \leq i \leq N-1$, we have the bounds

$$\left|\frac{1}{n}\log\mathbb{P}\left(\mathcal{E}_n \middle| \frac{M_n}{n} = \frac{i}{N}\right) - \hat{I}((N-i)/N)\right| < \epsilon_2. \quad (3)$$

Combining the bounds from inequalities (2) and (3), we obtain

$$\mathbb{P}(\mathcal{E}_n) \leq \sum_{i=0}^{N-1} e^{-n\left(\frac{N-i-1}{N}D(1/2||p) + \hat{I}((N-i)/N) - \epsilon_1 - \epsilon_2\right)}$$

$$\mathbb{P}(\mathcal{E}_n) \geq \sum_{i=0}^{N-1} e^{-n\left(\frac{N-i-1}{N}D(1/2||p) + \hat{I}((N-i-1)/N) + \epsilon_1 + \epsilon_2\right)}.$$

Let $\epsilon_3 > 0$ be an arbitrarily small constant. Define the three quantities

$$u := \inf_{x\in[0,1]} xD(1/2||p) + \hat{I}(x),$$

$$\bar{u} := \inf_{0\leq i\leq N-1} \frac{N-i-1}{N}D(1/2||p) + \hat{I}((N-i)/N),$$

$$\underline{u} := \inf_{1\leq i\leq N-1} \frac{N-i-1}{N}D(1/2||p) + \hat{I}((N-i-1)/N).$$

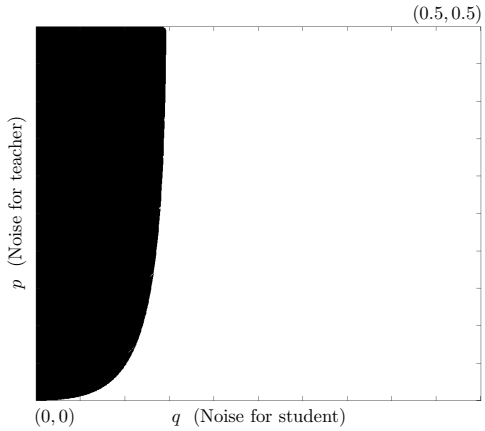

Fig. 1. The dark region indicates where the low-effort strategy outperforms the high-effort strategy.

Using the continuity of $\hat{I}$, we now pick N such that

$$\max(|u-\bar{u}|, |u-\underline{u}|) < \epsilon_3.$$

Then for all large enough n, we have the bounds

$$\mathbb{P}(\mathcal{E}_n) \leq N e^{-n(u-\epsilon_1-\epsilon_2-\epsilon_3)},$$
$$\mathbb{P}(\mathcal{E}_n) \geq e^{-n(u+\epsilon_1+\epsilon_2+\epsilon_3)}.$$

Taking logarithms, dividing by n, and taking the limit, we see that

$$\left| \lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}(\mathcal{E}_n) - u \right| < \epsilon_1 + \epsilon_2 + \epsilon_3.$$

Since $\epsilon_1, \epsilon_2, \epsilon_3$ are arbitrary constants, we conclude that

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}(\mathcal{E}_n) = u.$$

■

The above result may be simplified further. Notice that when $\bar{\theta}\bar{q}+\theta q \in [0, 1-\delta]$, the value of $\hat{I}(\theta) = 0$, since the infimum in its definition may be calculated to be at $\bar{\theta}\bar{q}+\theta q$. Since we are minimizing $\theta D(1/2||p) + \hat{I}(\theta)$, we would like to consider the smallest possible value of θ such that $\bar{\theta}\bar{q}+\theta q \leq 1-\delta$. It follows that there is no need to consider $\theta > \frac{\delta-q}{\bar{q}-q}$. Furthermore, for $\theta \in \left[0, \frac{\delta-q}{\bar{q}-q}\right]$, we have the equality $\hat{I}(\theta) = I_\theta(1-\delta)$. This means that we may write exponential decay of the error probability as

$$\inf_{\theta \in \left[0, \frac{\delta-q}{\bar{q}-q}\right]} \theta D(1/2||p) + I_\theta(1-\delta).$$

This expression does not simplify further, but since the function being minimized is known in closed form, it is convenient to simulate this using MATLAB or similar softwares.

V. Discussion

In Figure 1, we have compared the learning rates for the student for the low-effort and the high-effort strategies. The black region indicates the region where the low-effort strategy dominates. This shows that if the teacher to student channel does not have a lot of noise, then it is better for the teacher to not teach her best estimate to the student; i.e., send uncoded information. A rough intuition for this is that the teacher may receive many incorrect observations initially by chance. In this case, the high-effort strategy (which relies on the teacher's majority opinion) has a significant delay in correcting the teacher's opinion. However, if the teacher is following the low-effort strategy, the flipped observations in the beginning have no effect on the teacher's future communications. Furthermore, since the student has a relatively clean channel, she does need a high-effort strategy to learn quickly. A surprising threshold of $q \approx 0.15$ also emerges from the figure: if the teacher to student channel is more noisy that this threshold, then it is *always* beneficial to use the high-effort strategy, no matter how bad the teacher's observations.

There are various other strategies that the teacher and student may employ that we have not discussed here. For example, when the teacher is using the high-effort strategy, the student may "ignore" the first few observations and use the majority rule only on the latter observations, since they are more likely to be correct. Bayesian strategies are also worth analyzing. In general, it is an open problem to determine the optimal joint strategies they may employ to maximize the learning rate, or indeed establish non-trivial upper bounds on the best possible learning rates.

We also note that our result in Section III is closely related to the famous Ballot Theorem in combinatorics [9]. Theorem 1 is essentially a more refined analysis of the classical Ballot Theorem setting. Extending Theorem 1 to more general Markov chains, such as the Brownian motion, could potentially lead to newer versions of Ballot Theorems.

Acknowledgments

VJ gratefully acknowledges the support by NSF grant CCF-1841190.

References

[1] E. Mossel, O. Tamuz, et al. Opinion exchange dynamics. *Probability Surveys*, 14:155–204, 2017.
[2] C. Chamley. *Rational herds: Economic models of social learning*. Cambridge University Press, 2004.
[3] X. Vives. How fast do rational agents learn? *The Review of Economic Studies*, pages 329–347, 1993.
[4] A. Jadbabaie, P. Molavi, and A. Tahbaz-Salehi. Information heterogeneity and the speed of learning in social networks. *Columbia Business School Research Paper*, 2013.
[5] P. Molavi, A. Tahbaz-Salehi, and A. Jadbabaie. Foundations of non-bayesian social learning. *Columbia Business School Research Paper*, 2017.
[6] M. Harel, E. Mossel, P. Strack, and O. Tamuz. The speed of social learning. *arXiv preprint arXiv:1412.7172*, 2014.
[7] P. Dupuis and R. S. Ellis. *A weak convergence approach to the theory of large deviations*, volume 902. John Wiley & Sons, 2011.
[8] R. P. Stanley. *Catalan numbers*. Cambridge University Press, 2015.
[9] L. Addario-Berry and B. A. Reed. Ballot theorems, old and new. In *Horizons of combinatorics*, pages 9–35. Springer, 2008.