Are Imperfect Reviews Helpful in Social Learning?

Tho Ngoc Le*, Vijay G. Subramanian**, Randall A. Berry*
*EECS Department, Northwestern University
**EECS Department, University of Michigan
Email: thole2012@u.northwestern.edu, vgsubram@umich.edu , rberry@eecs.northwestern.edu

***Abstract*—Social learning encompasses situations in which agents attempt to learn from observing the actions of other agents. It is well known that in some cases this can lead to *information cascades* in which agents blindly follow the actions of others, even though this may not be optimal. Having agents provide reviews in addition to their actions provides one possible way to avoid "bad cascades." In this paper, we study one such model where agents sequentially decide whether or not to purchase a product, whose true value is either good or bad. If they purchase the item, agents also leave a review, which may be imperfect. Conditioning on the underlying state of the item, we study the impact of such reviews on the asymptotic properties of cascades. For a good underlying state, using Markov analysis we show that depending on the review quality, reviews may in fact increase the probability of a wrong cascade. On the other hand, for a bad underlying state, we use martingale analysis to bound the tail-probability of the time until a correct cascade happens.**

I. INTRODUCTION

People often seek to learn from observing others when faced with new decisions. On-line platforms facilitate acquiring such information at a much greater scale than was previously possible. A basic question is then to understand how such information facilitates learning. One common approach for studying such questions is as a game among Bayesian agents. These agents sequentially make a binary decisions given their own private information as well as observations of the decisions of previous agents. A key result, first shown in [2] and [3], is that in such models *herding* or an *information cascade* can occur in which from some point onward all agents ignore their private information and follow the actions of the previous agents. Though individually optimal, this may result in the agents making a choice that is not socially optimal.

One reason for incorrect herding is that agents observe the actions of other agents *before* the other agents receive their pay-off, and so these actions reflect the agents' estimates of the true pay-off and not the true pay-off itself. Indeed, if agents instead were able to see the true pay-off obtained by others, then as shown in [9] there would never be an incorrect cascade in which agents buy a bad product. The use of reviews and on-line recommendation systems can be viewed as an attempt to provide other agents with this information. However, due for example to user errors, such reviews may only be an imperfect representation of this information (instead of the true pay-off as in [9]). Studying social learning in the presence of such imperfect reviews is the objective of this paper. More precisely, we consider a variation of the model in [2], [3], where agents have the option to either buy or not buy a given item, whose true value is one of two binary states (good or bad). In addition to the actions of the previous agents, agents also see a history of reviews before making their decisions. However, these reviews are not a perfect indication of the true state of the good due to two effects: first, as we have already mentioned, these reviews are imperfect and second, agents can only leave a review if they buy the good and so no additional information is given for agents that choose not to buy.[1]

The second effect in the previous paragraph introduces asymmetry in the resulting user behavior depending on the underlying state of the product. As an example of this, assume there is a new product in the market. If the product is "good," one would hope that by introducing reviews, agents will eventually learn the product's true value. However, this is true only if there are a sufficient number of reviews submitted by agents. If not, there is still the possibility of an incorrect cascade, where at some point all subsequent agents stop buying. This could happen, for example, if the first few agents have noisy private signals that tell them not to buy the product. As a result, those agents do not buy the product, and so generate no reviews. If sufficient agents do this, it can lead to a cascade in which no agents buy the product. In such a cascade no reviews are generated to stop this wrong cascade. On the other hand, when the product is "bad" an incorrect cascade cannot persist: if more and more agents buy the product, more reviews are collected. By the law of large numbers, eventually a sufficient number of negative reviews will stop the wrong cascade and agents will learn the true value of the product.

We analyze these two cases separately. Conditioned on the state of the product being good, we study the probability of an incorrect cascade. We show that even for reviews of nearly perfect strength, there exist scenarios where one would prefer having no review at all since the probability of wrong cascade is lower. Conditioned on the state being bad, we instead focus on the time until a correct cascade occurs and give a tail bound on this probability that illustrates the impact of review quality.

Adding reviews is a way of changing the information structure in [2], [3]. There have been a variety of other papers that considered other changes in this structure such as changing the underlying network structure among the agents, e.g. [8], or changing the signal structure, e.g. [6]. In prior work ([10], [11]), we considered a variation of the information structure, where agent's observed noisy observations of the *actions* of others. This led to the following counter-intuitive

[1]For example, many on-line platform such as Amazon.com indicate *verified purchase reviews*; in our model only such reviews are considered.

result: the probability of incorrect herding is non-monotonic in the noise level. In other words, in some cases, more noise is actually beneficial. In this paper, agents perfectly observe the actions of previous agents and the only imperfection is in the reviews. Additionally, since only agents who buy the good can submit reviews, this leads to an asymmetry in the model that was not present in [10], [11]).

Another strand of related work is the literature on "word-of-mouth" learning (e.g. [4], [5], [7]) in which agents can communicate information about payoff of past actions. However, these models consider different settings (e.g. naive rule-of-thumb decision-based, random sampling of population); while our paper assumes that fully-rational agents can observe all past actions and reviews.

This remainder of the paper is organized as follows. In Section II we specify our model. The main results are presented in Sections III and IV for the case where the value of product is "good" and "bad," respectively. We conclude in Section V.

II. Model

We consider a model similar to [2] in which there is a countable population of agents, indexed $n = 1, 2, \ldots$ with the index reflecting the time and order of actions of the agents. Each agent n has an action choice A_n of saying either Yes (Y) or No (N) to a new item. The true value (V) of the item can be either good (G) or bad (B); both possibilities are assumed to be equally likely and the same for all agents. To reflect the agents' prior knowledge about the true value of the item, we assume that each agent n receives a private signal $S_n \in \{1 \text{ (high)}, 0 \text{ (low)}\}$. For each agent n who chooses $A_n = Y$, he submits a review $R_n \in \{G \text{ (Good)}, B \text{ (Bad)}\}$ representing his experience with the item after purchasing. However, if $A_n = N$ then he does not submit a review; which we denote by $R_n =^*$. Assume the probability that a private signal (resp. a review) aligns with V is $p \in (0.5, 1)$ (resp. $\delta \in [0.5, 1]$). That is:

$$\mathbb{P}[S_n = 1|V = B] = \mathbb{P}[S_n = 0|V = G] = 1 - p,$$
$$\mathbb{P}[S_n = 1|V = G] = \mathbb{P}[S_n = 0|V = B] = p, \text{ and if } A_n = Y,$$
$$\mathbb{P}[R_n = G|V = G] = \mathbb{P}[R_n = B|V = B] = \delta,$$
$$\mathbb{P}[R_n = G|V = B] = \mathbb{P}[R_n = B|V = G] = 1 - \delta.$$

We consider a homogeneous population where conditioned on V, the private signals and reviews are i.i.d. across all agents. Since $p \in (0.5, 1)$, the signals are informative, but not revealing; we call p the *signal quality*. In addition, let δ denote the *review's strength* which is independent of the signal.[2]

We assume that each agent takes his one-time action in exogenous order where the actions and reviews history is public information to subsequent agents. The agents are Bayes-rational and make decisions based on their own private signals and the public information. Each agent n updates his posterior belief about the true value V using his private signal S_n, the actions $A_1, \ldots, A_{n-1}$, and the reviews R_m for $m = 1, \ldots, n-1$.[3]

A. Public likelihood ratio as a Markov process

Let $q = 1 - p$. Denote the public history *after* agent n decides as $\mathcal{H}_n = \{A_1, R_1, \ldots, A_n, R_n\}$. Agents' decisions are based on calculations of the posterior probability of $V = B$ versus $V = G$ given the observed history $\mathcal{H}_n$. However, due to the independence of signals from history, agent $n+1$ can instead compare the public likelihood ratio ℓ_n of $V = B$ versus $V = G$, and his private belief which is a function of only his private signal. The private belief is calculated as q/p (resp. p/q) if $S_{n+1} = 1$ (resp. $S_{n+1} = 0$). On the other hand, using Bayes' rule and V being equally likely B or G, we can rewrite ℓ_n in its alternate form $\ell_n = \mathbb{P}[\mathcal{H}_n|V = B]/\mathbb{P}[\mathcal{H}_n|V = G]$. Since V is equally likely G or B, $\ell_0 = 1$. The higher ℓ_n is, the more likely $\mathcal{H}_n$ is indicating $V = B$. Moreover, since $\mathcal{H}_n$ is public information, for both V values ℓ_n can be updated as:

- If agent n follows his own signal then:

$$\ell_n = \begin{cases} \frac{p}{q}\ell_{n-1}, & \text{if } A_n = N \\ \frac{q}{p}\frac{1-\delta}{\delta}\ell_{n-1}, & \text{if } A_n = Y, R_n = G \\ \frac{q}{p}\frac{\delta}{1-\delta}\ell_{n-1}, & \text{if } A_n = Y, R_n = B. \end{cases} \tag{1}$$

- Otherwise, if agent n cascades then:

$$\ell_n = \begin{cases} \ell_{n-1}, & \text{if } A_n = N \\ \frac{1-\delta}{\delta}\ell_{n-1}, & \text{if } A_n = Y, R_n = G \\ \frac{\delta}{1-\delta}\ell_{n-1}, & \text{if } A_n = Y, R_n = B. \end{cases} \tag{2}$$

B. Agents' decision rule, cascades' condition and Markov property

Define $x = \log_{\frac{\delta}{1-\delta}} \frac{p}{q} \in [0, \infty]$ as the indicator of how strong the reviews are with respect to signals. In other words, the lower x is, the stronger the reviews are relative to the signals. Also, let a_n and r_n be two integer random variables denoting the two differences in actions ($\#Y - \#N$) and reviews ($\#G - \#B$), respectively. Note that while a_n excludes the actions caused by both types of cascades (since then the cascading actions provide no information), r_n is unchanged whenever a review is not made due to an agent not buying.

By (1) and (2), we can rewrite $\ell_n = (q/p)^{a_n}((1-\delta)/\delta)^{r_n} = (q/p)^{h_n}$ where the exponent $h_n = a_n + \frac{1}{x}r_n$. Since agent $n+1$ makes his decision by comparing ℓ_n to his private belief, agent $n+1$ cascades Y if $h_n > 1$, cascades N if $h_n < -1$, and follows his signal if $h_n \in [-1, 1]$.

By the above recursive relation, $\{\ell_n\}$ is a Markov process. Moreover, this is also true if in addition we condition on each value of V[5]. Equivalently, the dynamics of the process $\{\ell_n\}$ can be studied by investigating the 2-D Markov chain (a_n, r_n). Moreover, for special values of x, this investigation can be further simplified to a 1-D Markov chain where the state is denoted as h_n. We will study a few of such scenarios in Section III.

[2] The motivation being, while signal quality reflects product's marketing efficiency, the review strength is a function of the individual agent's behavior.

[3] For simplicity, we assume indifferent agents follow their own signals.

[5] This is an extension of results from [6].

C. Asymmetry under different types of cascades and product quality

This model exhibits asymmetric behaviors with respect both to the types of cascades (Y or N), and to the true value V of the item. In particular, the arrival of new information (reviews) depends on the action chosen by each agent.

1) Y versus N cascades: If agent n faces $h_{n-1} > 1$, he chooses $A_n = Y$ regardless of his signal and thus initiates a Y cascade. A Y cascade does not last forever, unless the reviews are of perfect quality ($\delta = 1$). For example, if $R_n = B$, then $h_n = h_{n-1} - \frac{1}{x}$ could be below 1, which induces agent $n+1$ to use his own signal. Furthermore, if x is sufficiently small then $h_n < -1$, and agent $n+1$ initiates a N cascade. The dynamics of a Y cascade, once it gets started, are determined solely by the reviews process (and it does not depend on the signals). Regardless of the time a Y cascade was initiated, it can be broken by a sufficiently long sequence of bad reviews. Thus, the history process $\{\mathcal{H}_n\}$ could include sample paths where Y cascades start and stop infinitely often.

On the other hand, once $h_n < -1$, a N cascade starts and lasts forever. This is due to agents who choose N not generating reviews; thus the likelihood ratio stays constant as soon as any agent cascades to N. Subsequent agents who have the same signal strength are left in the same state as the one who initiated the cascade; thus they make the same action choice.

2) Good versus bad product: When $V = G$, a wrong cascade happens with positive probability. For example, if the first two agents have low signals, they both choose N; therefore no review is collected. As a result, all subsequent agents are drawn into a N cascade, which is irreversible. This possibility cannot be avoided by adjusting the reviews strength, δ, even to perfect quality. In case the reviews are perfect, we would still need a non-cascading agent who has a high signal for his review to take effect. In addition, for $V = G$, it is highly likely that there is an abundance of new information. If an agent n chooses Y, one review R_n is added to the common database. Since reviews are independent of signals, when $V = G$ more agents choose Y and new information begets further new information.

In other words, when $V = G$ the underlying Markov process have a drift toward the correct cascade, but there is no absorbing state on that side since h_n is unbounded above. However, multiple absorbing states for wrong cascades might exist. For $V = G$, the quantity of interest is the probability of wrong (N) cascade which is a function of both p and δ. One would expect the time until correct cascade to be infinite by considering the drift of the underlying Markov process. We will discuss this scenario in Section III.

On the other hand, when $V = B$, this model exhibits a different set of behaviors. Wrong cascades can never happen. The reason is as more agents purchase the item, more and more reviews are collected. Since reviews are informative, subsequent agents can track the difference in the number of reviews to learn the true value of V eventually. In other words, while there are only trapping states for correct cascade, the drift also leans toward this side. Thus, correct cascade happens with probability 1. In this scenario, we are interested in the distribution of the time (i.e. the number of agents) until a correct cascade happens. This will be studied in Section VI.

III. Probability of wrong cascade for $V = G$

In previous section, we discussed that wrong (N) cascades could happen if the product is good. In this section, we determine the probability of this happening. For a fixed p, as x varies the conditions on a_n and r_n when cascades happen also changes. As a result, the underlying 2-D Markov chains (a_n, r_n) have different structures (both in terms of their states space and transition probabilities). Despite the complexity of these dynamics for a generic value of x, interesting and non-intuitive insights can be drawn by looking at special values of x. In Proposition 1, we consider two cases: $x = 1$ and $x = 1/2$; in both cases we can simplify each state (a_n, r_n) to a state h_n of the corresponding 1-D Markov chains, as shown in Fig. 1 and 2, respectively.

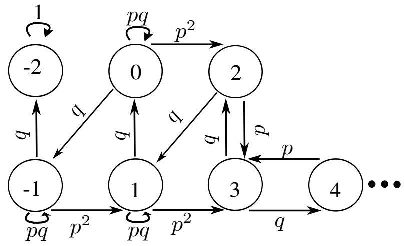

Figure 1: States transitions for $V = G$, and $x = 1$.

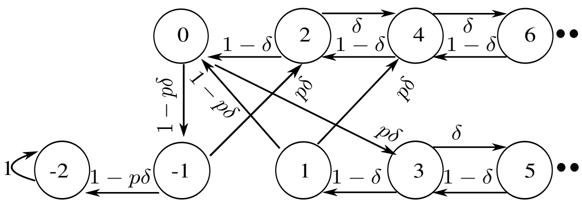

Figure 2: States transitions for $V = G$, and $x = 1/2$.

Proposition 1. *1) Having reviews twice as strong as signals (i.e. $x = 1/2$) gives the same probability of wrong cascade as having reviews with the same strength as signals (i.e. $x = 1$), and*

2) Both cases give a higher probability of wrong cascade as compared to when having no review.

Proof. 1) The probabilities of wrong cascade can be calculated using the transition diagrams and the first-step technique in Markov chain analysis. For both values $x = 1$ and $x = 1/2$, we obtain a wrong cascade probability of $(q/p)^2$.

2) When there are no reviews, result from [2] gives a wrong cascade probability of a $\frac{(q/p)^2}{(q/p)^2+1}$, which is less than $(q/p)^2$. $\square$

Proposition 1 suggests that one should look at regions where reviews are even stronger. Unfortunately, Proposition 2 shows that except for the reviews having perfect strength, one cannot guarantee a better performance with reviews for all values of $p \in (0.5, 1)$. To illustrate, we consider all values $x < 1/3$, i.e. reviews are more than triple the signals' quality. In Fig. 3, the underlying Markov chain is 2-D where the first and second coordinates denote r_n and a_n, respectively.

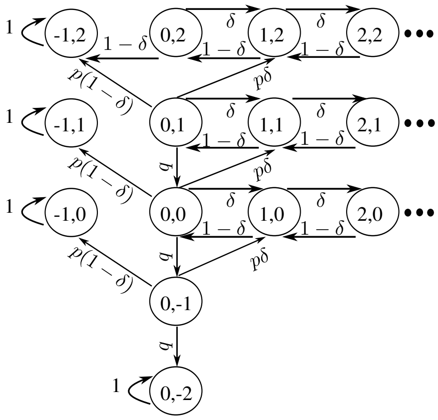

Figure 3: States transitions for $V = G$, and $x < \frac{1}{3}$.

Proposition 2. *Assume $0 < x < 1/3$ (i.e. reviews are more than triple the signals' quality):*

1) *$\mathbb{P}[wrong]$ decreases in the review quality, δ, and*
2) *At $x = 0$ (i.e. perfect reviews), $\mathbb{P}[wrong] = q^2$, which is lower than that for having no review.*
3) *For x bounded away from 0, there exists a threshold $p_0 \in (0.5, 0.75)$ such that for signal quality with $p < p_0$, we are better off having no reviews.*

Proof. 1) Again, using the first-step technique for Markov chains, we solve a system of linear equations for the probability of wrong cascade starting from state $(0, 0)$. This gives a result of:

$$\mathbb{P}[wrong] = [1 - p(2\delta - 2p\delta + 2p - p/\delta)] \, / \, [1 - 2pq(1 - \delta)]$$

which is decreasing in δ.

2) For perfect reviews, a wrong cascade happens if and only if the first two agents have low signals, which happens with probability $q^2 < \left[(q/p)^2\right] / \left[(q/p)^2 + 1\right]$.

3) p_0 is the solution to:

$$\mathbb{P}[wrong] = \left[(q/p)^2\right] / \left[(q/p)^2 + 1\right]. \tag{3}$$

First, we will show the existence of p_0 in $(0.5, 0.75)$. In fact, (3) is equivalent to:

$$f(p) \stackrel{\triangle}{=} p^3 \left[6\delta + (2/\delta) - 6\right] + p^2 \left[-14\delta - (2/\delta) + 10\right] + p \left[12\delta + (1/\delta) - 7\right] + (2 - 4\delta) = 0. \tag{4}$$

Note that $f(p)$ is continuous in p. It can be easily shown that for any $\delta \in [0.5, 1]$, $f(0.5) > 0$ and $f(0.75) < 0$. Thus, by the Mean Value Theorem there exists a root $p_0 \in (0.5, 0.75)$.

Now we show that p_0 is the only root of $f(p)$ in $(0.5, 1)$. Since $f(0) \leq 0$, $f(p)$ has another root $p_1 \in [0, 0.5)$. Moreover, $p = 1$ is another root of $f(p)$ (note that at $p = 1$, $\delta = 1$). In addition, since $f(p)$ is a cubic polynomial in p with positive highest order coefficient, we conclude that for $0.5 < p < p_0$, $f(p) > 0 \Rightarrow LHS(8) > RHS(8)$; and for $p_0 < p < 1$, $f(p) < 0 \Rightarrow LHS(8) < RHS(8)$. □

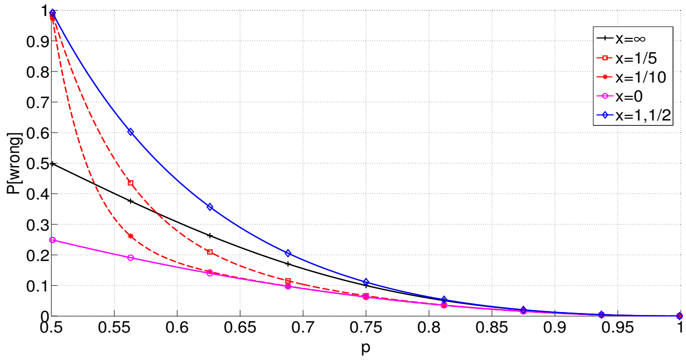

Figure 4: Wrong herding probability, $V = G$.

Fig. 4 illustrates both Propositions 1 and 2. For all cases, the probability of wrong cascade decreases in the signal quality p. Moreover, except for reviews with perfect accuracy, one would prefer having no reviews for low enough signal quality.

IV. Time until correct cascade for $V = B$

In Section II, we argue that for a bad product, only correct (N) cascades can happen, which also last forever. In this section, we examine the distribution of the time until a correct cascade by determining its tail exponent. In the following let $n \geq 0$. Conditioned on $V = G$ (resp. B), let $\{\mathscr{F}_n^G\}$ (resp. $\{\mathscr{F}_n^B\}$) be the sequence of σ-algebras generated by $\{\mathcal{H}_n\}$. Similar to the models in [6] and [8] where reviews do not exist, in our model the Markov process $\{\ell_n\}$ also exhibits the martingale property as presented in the following lemma:

Lemma 1. *$\{1/\ell_n\}$ (resp. $\{\ell_n\}$) is a martingale process conditioned on $V = B$ (resp. $V = G$) adapted to the filtration $\{\mathscr{F}_n^B\}$ (resp. $\{\mathscr{F}_n^G\}$)*

Proof. Given $\ell_1, \ldots, \ell_n$ and p, δ as common knowledge, subsequent agents know agent $n+1$'s decision rule. If agent $n+1$ follows an N cascade, $\ell_{n+1} = \ell_n$ thus the martingale property follows naturally. Otherwise, for $V = B$, if agent $n+1$ follows his own signal then:

$$\begin{aligned} &\mathbb{E}[1/\ell_{n+1} | \mathscr{F}_n^B] = \mathbb{P}[A_{n+1} = N | V = B] \left[(1-p)/p\right] / \ell_n \\ &+ \mathbb{P}[A_{n+1} = Y, R_{n+1} = G | V = B] \left[p/(1-p)\right] \left[\delta/(1-\delta)\right] / \ell_n \\ &+ \mathbb{P}[A_{n+1} = Y, R_{n+1} = B | V = B] \left[p/(1-p)\right] \left[(1-\delta)/\delta\right] / \ell_n \\ &= 1/\ell_n. \end{aligned} \tag{5}$$

Similarly if agent $n+1$ cascades to Y when $V = B$ then:

$$\mathbb{E}[1/\ell_{n+1} | \mathscr{F}_n^B] = \delta/\ell_n + (1-\delta)/\ell_n = 1/\ell_n. \tag{6}$$

From (5), (6), it follows that $\{\frac{1}{\ell_n}\}$ is a martingale for $V = B$. For $V = G$, similar method shows that $\{\ell_n\}$ is a martingale. □

Using Lemma 1 and techniques from [1], we use the martingale property to bound the tail probability of the time until correct cascade. This can also be used to give a bound on the expected time until correct cascade. Let X and Y be two random variables representing the increments $\Delta h_n = h_{n+1} - h_n$ for h_n in $[-1, 1]$ and $h_n > 1$, respectively. Let $f_1(\lambda)$ and $f_2(\lambda)$ be their corresponding moment generating functions (MGFs), where λ is a real variable. Let $\rho = \max(f_1(\lambda), f_2(\lambda))$ and define the random process $\{M_n\} = \left\{\frac{e^{\lambda h_n}}{\rho^n}\right\}$. We have:

Lemma 2. *$\{M_n\}$ is a super-martingale adapted to $\{\mathscr{F}_n^B\}$.*

Proof.

$$\frac{\mathbb{E}[M_{n+1}|\mathscr{F}_n^B]}{M_n} = \frac{\mathbb{E}[e^{\lambda h_{n+1}}/\rho^{n+1}|\mathscr{F}_n^B]}{e^{\lambda h_n}/\rho^n} = \frac{\mathbb{E}[e^{\lambda \Delta h_n}]}{\rho}$$
$$\leq \frac{max\left(\mathbb{E}[e^{\lambda X}], \mathbb{E}[e^{\lambda Y}]\right)}{\rho} = 1.$$

□

Let $\tau = \min\{n \geq 0 : h_n < -1\}$ be the stopping time when N cascade happens. Now we use Lemma 2 to give an upper-bound on the tail-probability of τ in the following proposition.

Proposition 3. *$\mathbb{P}[\tau > n] \leq e^{\lambda}\rho^n$, where $0 < \rho < 1$.*

Proof. For feasibility, we require $0 < \rho < 1$, which implies $\lambda \in (0, \ln(\frac{p}{1-p}))$. Since τ is a stopping time, so is $n \wedge \tau$. Thus $\{M_{n\wedge\tau}\}$ is also a super-martingale. Thus by $h_0 = 0$ we have:

$$\begin{aligned}
1 &= e^{\lambda h_0} = M_0 \geq \mathbb{E}[M_{n\wedge\tau}|\mathscr{F}_0^B] \\
&\geq \mathbb{E}[M_{n\wedge\tau}; \tau > n|\mathscr{F}_0^B]\mathbb{P}[\tau > n] \\
&= \mathbb{E}[e^{\lambda h_{n\wedge\tau}}/\rho^{n\wedge\tau}; \tau > n|\mathscr{F}_0^B]\mathbb{P}[\tau > n] \\
&\geq \mathbb{E}[e^{\lambda h_n}/\rho^n; \tau > n|\mathscr{F}_0^B]\mathbb{P}[\tau > n] \\
&\geq e^{\lambda(-1)}\rho^{-n}\mathbb{P}[\tau > n], \text{ since } h_n > -1 \text{ when } \tau > n \\
&\Rightarrow \mathbb{P}[\tau > n] \leq e^{\lambda}\rho^n.
\end{aligned}$$

□

The above bound is a function of n, the agent index, the dummy variable λ, and the two MGFs f_1, f_2. Our objective is to choose λ and ρ which minimize this bound. We solve this numerically and compare the minimum bound with the tail-probability obtained using Monte-Carlo simulation for different values of p and δ.

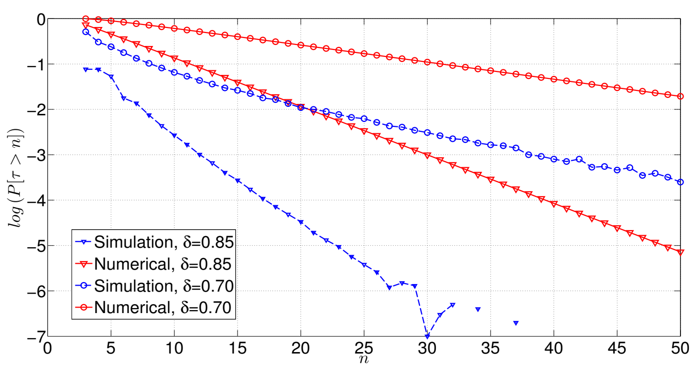

Figure 5: Tail-probability of time until N cascade, $p = 0.70$.

Fig. 5 shows that both simulation results and numerical bounds are decreasing as δ increases. Moreover, the higher value of δ, the faster the rate at which both the simulation and the numerical results decay to zero.

V. Conclusions and Future Work

This paper studied a simple observational Bayesian learning model. We assumed that subsequent agents can perfectly observe the previous actions and, in addition, feedback in the form of reviews, which depend on the actions. We showed that the reviews could increase the probability that agents misinterpret the true value of a good product. In practice, in online platforms like Yelp, Amazon, etc. customers reviews come with a variability of strengths. Even though this scenario was not considered in this paper, our results indirectly implied that a platform planner should opt to cut out the reviews of bad qualities and release only the "good" ones. In fact, this strategy is already adopted by many platforms, e.g. Amazon with verified purchase reviews, or Yelp with filtered reviews. Moreover, our results suggested that no matter how strong the reviews quality, agents might not perform better if their prior knowledge are limited. This implies that a platform planner should consider spending their budget on improving both the product's marketing efficiency and the reviews' reliability.

In the future work, we plan to study the possibility of having reviews with strengths non-homogeneously distributed across the population. Another possible direction is by considering having reviews when both type of actions are taken, where agents have the option to leave the reviews and assuming that not all agents would exercise this option.

References

[1] B. Hajek, *Hitting and occupation time bounds implied by drift analysis with applications*, Advances in Applied Probability, vol. 14, pp. 502-525, 1982

[2] S. Bikhchandani, D. Hirshleifer, and I. Welch. *A Theory of Fads, Fashion, Custom and Cultural Change as Informational Cascades*, J. Polit. Econ., vol. 100, No. 5, pp. 992-1026, 1992.

[3] A. Banerjee, *A Simple Model of Herd Behavior*, The Quarterly Journal of Economics, vol. 107, pp. 797-817, 1992.

[4] G. Ellison, D. Fudenberg, *Rules of thumb for social learning*, Journal of Political Economy, vol. 101, pp. 612-643, 1993.

[5] G. Ellison, D. Fudenberg, *Word-of-mouth communication and social learning*, The Quarterly Journal of Economics, vol. 110, pp. 93-125, 1995.

[6] L. Smith, P. Sorensen, *Pathological Outcomes of Observational Learning*, Econometrica, vol. 68, pp. 371-398, 2000.

[7] A. Banerjee, D. Fudenberg, *Word-of-mouth learning*, Games and Economic Behavior, vol. 46, pp. 1-22, 2004.

[8] D. Acemoglu, M. Dahleh, I. Lobel, and A. Ozdaglar, *Bayesian learning in social networks*, Review of Economic Studies, vol. 78, pp. 1201-1236, 2011.

[9] H. Cao, B. Han, and D. Hirshleifer, *Taking the road less traveled by: Does conversation eradicate pernicious cascades?*, Journal of Economic Theory, vol. 146, pp. 1418-1436, 2011.

[10] T. Le, V. Subramanian, R. Berry, *The Value of Noise for Informational Cascades*, IEEE ISIT Proceeding, 2014.

[11] T. Le, V. Subramanian, R. Berry, *The Impact of Observation and Action Errors on Informational Cascades*, IEEE CDC Proceeding, 2014.