Observational Learning in Random Networks

Julian Lorenz, Martin Marciniszyn, and Angelika Steger

Institute of Theoretical Computer Science, ETH Zurich
8092 Zurich, Switzerland
{jlorenz,mmarcini,steger}@inf.ethz.ch

Abstract. In the standard model of observational learning, n agents sequentially decide between two alternatives a or b, one of which is objectively superior. Their choice is based on a stochastic private signal and the decisions of others. Assuming a rational behavior, it is known that informational cascades arise, which cause an overwhelming fraction of the population to make the same choice, either correct or false. Assuming that each agent is able to observe the actions of *all predecessors*, it was shown by Bikhchandani, Hirshleifer, and Welch [1, 2] that, independently of the population size, false informational cascades are quite likely.

In a more realistic setting, agents observe just a *subset of their predecessors*, modeled by a random network of acquaintanceships. We show that the probability of false informational cascades depends on the edge probability p of the underlying network. As in the standard model, the emergence of false cascades is quite likely if p does not depend on n. In contrast to that, false cascades are very unlikely if $p = p(n)$ is a sequence that decreases with n. Provided the decay of p is not too fast, correct cascades emerge almost surely, benefiting the entire population.

1 Introduction

In recent years, there has been growing interest in modeling and analyzing processes of observational learning, first introduced by Banerjee [3] and Bikhchandani, Hirshleifer, and Welch [1, 2]. In the model of [1, 2], individuals make a once-in-a-lifetime choice between two alternatives sequentially. Each individual has access to private information, which is hidden to other individuals, and also observes the choices made by his predecessors. Since each action taken provides an information externality, individuals may start to imitate their predecessors so as to maximize their objective. Although such *herding behavior* is a locally optimal strategy for each individual, it might not be beneficial for the population as a whole. In the models of [3] and [1, 2], imitation may cause an informational cascade such that all subsequent individuals make the same decision, regardless of their private information. One of the main results in [3] and [1, 2] states that the probability of a cascade that leads most members of the population into the false decision is constant, independently of the population size.

This result seems counterintuitive to our every day experience since at many occasions taking the choice of others into account is wise and beneficial for the entire society. In fact, imitation has been recognized as an important manifestation of intelligence and social learning. For instance, in his popular bestseller "The Wisdom of Crowds" [4], Surowiecki praises the superior judgment of large groups of people over an elite few. This became evident, for example, when Google launched their web search engine, at that time offering a superior service quality. Encouraged by their acquaintances, more and more users adopted Google as their primary index to the web. Moreover, the Google search engine itself leverages the wisdom of crowds by ranking their search results with the PageRank algorithm [5].

The reason that herding could be rather harmful in the model studied in [1, 2] is that each individual has unlimited observational power over the actions taken by *all predecessors*. In a more realistic model, information disseminates not perfectly so that individuals typically observe merely a *small subset of their predecessors*. In this paper, we propose a generalization of the sequential learning model of [1, 2]. Suppose the population has size n. For each individual $i \in \{1, \ldots, n\}$, a set of acquaintances $\Gamma(i)$ among all predecessors $j < i$ is selected, where each member of $\Gamma(i)$ is chosen with probability $p = p(n)$, $0 \leq p \leq 1$, independently of all other members. Only the actions taken by members of $\Gamma(i)$ are revealed to the individual i, all other actions remain unknown to i. Thus, the underlying social network is a random graph according to the model of Erdős and Rényi [6]. Setting $p = 1$ resembles the model of [1, 2].

Extending the result of [1, 2], we show that if p is a constant, the probability that a false informational cascade occurs during the decision process is constant, i.e., independent of the population size n. On the other hand, if $p = p(n)$ is a function that decays with n arbitrarily slowly, the probability of a false informational cascade tends to 0 as n tends to infinity. Informally speaking, almost all members of fairly large, moderately linked social networks make the correct choice with probability very close to 1, which is in accordance with our every day experience.

1.1 Model of Sequential Observational Learning in Networks

We consider the following framework of sequential learning in social networks that naturally generalizes the setting in [1, 2]. There are n individuals (or equivalently, *agents* or *decision-makers* in the following), $V = \{v_1, \ldots, v_n\}$, facing a once-in-a-lifetime decision between two alternatives a and b. Decisions are made sequentially in the order of the labeling of V. One of the two choices is objectively superior, but which one that is remains unknown to all individuals throughout. Let $\theta \in \{a, b\}$ denote that superior choice. The a-priori probabilities of being the superior choice are

$$\mathbb{P}[\theta = a] = \mathbb{P}[\theta = b] = \frac{1}{2} \ .$$

Each agent $v_i \in V$ makes his choice $\mathrm{ch}(v_i) \in \{a, b\}$ based on two sources of information: a private signal $s(v_i) \in \{a, b\}$ and public information. The private signal $s(v_i)$ is only observed by the individual v_i. All private signals are identically and independently distributed, satisfying $\mathbb{P}[s(v_i) = \theta] = \alpha$. That is, α is the probability that a private signal correctly recommends the superior choice. We assume that $1/2 < \alpha < 1$, excluding the trivial case $\alpha = 1$.

The actions $\{\mathrm{ch}(v_i) \mid 1 \leq i \leq n\}$ are public information, but an individual v_i can only observe the actions of a subset $\Gamma_i \subseteq V_{i-1} = \{v_1, \ldots, v_{i-1}\}$ of acquaintances. For all agents v_i, $2 \leq i \leq n$, each of the possible acquaintances $v_j \in V_{i-1}$ is included with probability $0 \leq p = p(n) \leq 1$ into Γ_i, independently of all other elements in V_{i-1}. Equivalently, the underlying social network can be represented as a labeled, undirected random graph $G = G_{n,p}$ on the vertex set V, where each possible edge is included with probability p, independently of all other edges. Then the set of acquaintances Γ_i of agent v_i is given by $\Gamma_G(v_i) \cap V_{i-1}$, where $\Gamma_G(v_i)$ denotes the neighborhood of v_i in G. It is easily seen that both representations are equivalent [7, 8] and yield a random graph in the classical model of Erdős and Rényi [6]. We shall assume throughout this paper that the social network is exogenously determined before all decisions take place and represented in form of a random graph $G = G_{n,p}$.

Various models of social networks were proposed in the literature (see e.g. [9]). The classical random graph model of Erdős and Rényi is analytically well understood and, despite its idealistic assumptions, powerful enough to explain essential features of sequential social learning well. Moreover, it naturally generalizes the model proposed in [1, 2], which is captured in the case $p = 1$.

1.2 Main Result

All agents employ the following deterministic rule for making decisions, which is a slight variation of the decision rule in [1, 2].

Definition 1 (Decision rule). *Suppose individual v_i has received the private signal $s(v_i)$, and, among his acquaintances $\Gamma(i)$, m_a chose option a and m_b chose option b. Then we have*

$$\mathrm{ch}(v_i) = \begin{cases} a & \text{if } m_a - m_b \geq 2 \ , \\ b & \text{if } m_b - m_a \geq 2 \ , \\ s(v_i) & \text{otherwise } . \end{cases}$$

One can show that on a complete graph this strategy is locally optimal for each individual assuming that the actions of acquaintances are given in an aggregated form, that is, agent v_i merely observes how many times either of the options a and b was chosen before (see Lemma 11 in the appendix).

For any two sequences a_n and b_n, $n \in \mathbb{N}$, we write $a_n \ll b_n$ if

$$\lim_{n \to \infty} \frac{a_n}{b_n} = 0 \ .$$

Then our result reads as follows.

Theorem 1. *Suppose a social network with n agents $V = \{v_1, \ldots, v_n\}$ is given as a random graph $G = G_{n,p}$ with vertex set V and edge probability $p = p(n)$. Assume that private signals are correct with probability $1/2 < \alpha < 1$ and each agent applies the decision rule in Definition 1. Let $c_{\alpha,p}(n)$ be a random variable counting the number of agents that make the correct choice.*

(i) If $n^{-1} \ll p \ll 1$, we have

$$\lim_{n\to\infty} \mathbb{P}\left[c_{\alpha,p}(n) = (1 - o(1))n\right] = 1 \ . \tag{1}$$

(ii) If $0 \le p \le 1$ is a constant, then there exists a constant $\varrho = \varrho(\alpha, p) > 0$ such that

$$\lim_{n\to\infty} \mathbb{P}\left[c_{\alpha,p}(n) = o(n)\right] \ge \varrho \ . \tag{2}$$

In moderately linked social networks as in (i), the entire society benefits from learning. Note that if agents ignored the actions of others completely, typically a $(1-\alpha)$-fraction of the population would make the false decision. On the other hand, if each individual has very many acquaintances on average as in (ii), incorrect informational cascades that lead almost the entire population into the false decision are quite likely.

In very sparse random networks with $p \le c/n$ for some constant $c > 0$, no significant herding will arise since those networks typically contain γn isolated vertices for some constant $\gamma = \gamma(c) > 0$ [7, 8]. These agents make their decision independently of all other agents and, hence, we expect that both groups of agents, choosing a and b respectively, contain a linear fraction of the population.

The crucial difference between the model of [1, 2], which assumes that the underlying graph of the social network is complete, and our model is that in the former the probability of a false informational cascade primarily depends on the decision of very few agents at the beginning of the process. For instance, with constant probability the first three agents make the false decision, no matter which decision rule they apply. Since in a complete graph each subsequent agent observes these actions, locally optimal imitation will trick the entire population into the false decision. In contrast to that, information accumulates locally in the beginning if the underlying network is sparse as in (i). During a relatively long phase of the process, individuals make an independent decision because none of their acquaintances has decided yet. Hence, after that phase typically a fraction very close to α of these agents made the correct choice, creating a bias towards it. In later phases of the process, agents observe this bias among their acquaintances and, trusting the majority, make the correct decision, thereby increasing the bias even more. In the end, almost all agents are on the correct side.

Before presenting the proof of Theorem 1, let us make these ideas more precise. For any j, $1 \le j \le n$, let $V_j = \{v_1, \ldots, v_j\}$ denote the set of the first j agents. Recall that $\theta \in \{a, b\}$ denotes the objectively superior choice between a and b. For any set of agents $V' \subseteq V$, let

$$C(V') = \{v \in V' : \mathrm{ch}(v) = \theta\}$$

be the subset of agents in V' who made the correct decision. We denote the cardinality of $C(V')$ by $c(V')$. Suppose that in the first group of $j \geq 1$ agents approximately an α-fraction made the correct decision. The first important observation is that the subsequent agent v_{j+1} makes the correct choice with probability at least α if v_{j+1} obeys the decision rule in Definition 1.

Lemma 1. *Suppose the underlying social network is a random graph $G_{n,p}$ with edge probability $0 \leq p \leq 1$. Let $1/2 < \alpha < 1$ be fixed. Then there exists $\varepsilon > 0$ such that setting $\bar{\alpha} = (1-\varepsilon)\alpha$, for all $1 \leq j \leq n-1$, we have*

$$\mathbb{P}\left[\mathrm{ch}(v_{j+1}) = \theta \;\middle|\; c(V_j) \geq \bar{\alpha} j\right] \geq \alpha \;\; ,$$

provided agent v_{j+1} obeys the decision rule in Definition 1.

So basically, following the majority and using the private signal only to break ties does not decrease the chances of any agent even if his acquaintances are randomly selected, provided that there is a bias among all predecessors towards the right direction. This enables us to show that, throughout the first stage, a bias of $\bar{\alpha} > 1/2$ remains stable in the group of decided agents. Once this group has reached a critical mass, new agents adopt the correct choice with very high probability since the bias among their acquaintances is so evident. More specifically, we can show the following "herding" lemma .

Lemma 2. *Suppose the underlying social network is a random graph $G_{n,p}$ with edge probability $1 \ll p \leq 1$. Let $1/2 < \bar{\alpha} < 1$ be fixed. Then there exists a constant $\delta > 0$ and $j_0 \geq 1$ satisfying $j_0 = \mathcal{O}\big(p^{-1}\big)$ such that for all $j_0 \leq j \leq n-1$, we have*

$$\mathbb{P}\left[\mathrm{ch}(v_{j+1}) = \theta \;\middle|\; c(V_j) \geq \bar{\alpha} j\right] \geq 1 - \mathrm{e}^{-\delta p j} \;\; ,$$

provided agent v_{j+1} obeys the decision rule in Definition 1.

Thus, most agents opt for θ with probability very close to 1 in the second stage. What makes the crucial difference between parts (i) and (ii) of Theorem 1 is that if p is a constant, the assumption $c(V_j) \geq \bar{\alpha} j$ in Lemmas 1 and 2 is met in the process only with probability bounded away from 1. Then it is quite likely that agents experience a bias towards the false direction among their acquaintances, and the same herding behavior as before evokes a false informational cascade.

1.3 Related Results

As already mentioned, Bikhchandani, Hirshleifer, and Welch [1, 2] consider the case when the social network is a complete graph. Here informational cascades arise quickly, and it is quite likely that they are false. The authors of [1, 2] consider a decision rule that is slightly different from the one in Definition 1. As shown in Sect. A.1 of the appendix, although both rules are locally optimal, false informational cascades are more likely with the rule in [1, 2].

Models of observational learning processes were investigated in several papers. Banerjee [3] analyzes a model of sequential decision making that provokes herding behavior; as before, each decision-maker can observe the actions taken by *all* of his predecessors. In the model of Çelen and Kariv [10], decision-makers can only observe the action of their *immediate* predecessor. Banerjee and Fudenberg [11] consider the model in which each agent can observe the actions of a sample of his predecessors. This is comparable to our model with an underlying random network $G_{n,p}$. However, their model of making decisions is different; at each point in time, a proportion of the entire population leaves and is replaced by newcomers, who simultaneously make their decision. Similarly to our result, the authors of [11] show that, under certain assumptions, informational cascade are correct in the long run. In the learning process studied by Gale and Kariv [12], agents make decisions simultaneously rather than in a sequential order, but they may repeatedly revise their choice. Watts [13] studies random social networks, in which agents can either adopt or not. Starting with no adopters, in each round all agents update their state according to some rule depending on the state of their neighbors. In this model, the emergence of global informational cascades also depends on the density of the underlying random network.

1.4 Organization of the Paper

The paper is organized as follows. In Sect. 2 we present the proof of Theorem 1(i). An outline of this proof is contained in Sect. 2.1, where we also state a series of technical lemmas, which are proved in Sect. 2.2. The proof of Theorem 1(ii) is deferred to the appendix (see Sect. C). We conclude with experimental results in Sect. 3.

2 Proof of Theorem 1(i)

Suppose $n^{-1} \ll p \ll 1$ is given as in the theorem, and consider a random graph $G = G_{n,p}$ on the vertex set V with edge set E. For any set $V' \subseteq V$, let $E(V')$ denote the set of edges induced by V' in G. Recall that $C(V')$ denotes the subset of agents in V' who made the correct decision. Let $\bar{C}(V') = V' \setminus C(V')$ be its complement and set $c(V') = |C(V')|$ and $\bar{c}(V') = |\bar{C}(V')|$. The binomial distribution with mean np is denoted by $\mathrm{Bin}(n, p)$.

2.1 Outline of the Proof

The proof of of Theorem 1(i) is based on a series of auxiliary lemmas that we state here. The proofs of these lemmas are deferred to Sect. 2.2.

We subdivide the process of decision making into three phases as follows:

Phase I: Agents $V_I = \{v_1, \ldots, v_{k_0}\}$ with $k_0 = p^{-1}\omega^{-1/2}$.
Phase II: Agents $V_{II} = \{v_{k_0+1}, \ldots, v_{k_1}\}$ with $k_1 = p^{-1}\omega^{1/2}$.
Phase III: Agents $V_{III} = \{v_{k_1+1}, \ldots, v_n\}$.

In *Phase I*, the phase of the *early adopters*, most decision-makers have no more than one neighbor who has already decided. Such an agent always follows the private signal according to the decision rule in Definition 1, regardless of all others. Therefore, almost all agents make their decisions based solely on their private signal, which yields approximately an α-fraction of individuals who opted for θ. More specifically, we can establish the following lemma.

Lemma 3. *Let $\omega = \omega(n)$ be a sequence satisfying $1 \ll \omega \ll n$. Let $1/2 < \alpha < 1$, $0 < p \leq 1/\omega$ and $k_0 = p^{-1}\omega^{-1/2}$ be given. Then we have*

$$\mathbb{P}\left[c(V_{k_0}) \geq \left(1 - k_0^{-1/9}\right)\alpha k_0\right] = 1 - o(1) \quad .$$

Note that if $0 < p \leq 1$ is a constant independent of n, Phase I breaks down; there is no $k_0 \geq 1$ such that the number of correctly decided agents in V_{k_0} is roughly k_0 with probability $1 - o(1)$. That is exactly what makes the situation in part (ii) of Theorem 1 different.

In *Phase II*, more and more agents face decisions of their acquaintances. As stated in Lemma 1, everybody makes a correct choice with probability at least α assuming that roughly an α-fraction of the preceding agents took the right action. The following lemma asserts that approximately this fraction of correct decisions is maintained throughout the second phase.

Lemma 4. *Let $\omega = \omega(n)$ be a sequence satisfying $1 \ll \omega \ll n$. Let $1/2 < \alpha < 1$, $0 < p \leq 1/\omega$ and $k_0 = p^{-1}\omega^{-1/2}$ and $k_1 = p^{-1}\omega^{1/2}$ be given. Then we have*

$$\mathbb{P}\left[c(V_{k_1}) \geq \left(1 - k_0^{-1/18}\right)\alpha k_1 \;\middle|\; c(V_{k_0}) \geq \left(1 - k_0^{-1/9}\right)\alpha k_0\right] = 1 - o(1) \quad .$$

At the beginning of *Phase III*, every agent v_i has $\mathbb{E}\left[|\Gamma_i|\right] \geq pk_1 \gg 1$ decided neighbors on average. With high probability v_i disregards the private signal and follows the majority vote among its acquaintances, thereby making the correct choice.

Lemma 5. *Let $p > 0$, $\bar{\alpha} > 1/2$ and $k \geq 1$ be given. Then, for all $i > k$, we have*

$$\mathbb{P}\left[c(\Gamma_i \cap V_k) - \bar{c}(\Gamma_i \cap V_k) \geq \frac{2\bar{\alpha} - 1}{3} pk \;\middle|\; c(V_k) \geq \bar{\alpha}k\right] \geq 1 - 2\exp\left(-pkC\right) \quad .$$

where $C = (2\bar{\alpha} - 1)^2/(18\bar{\alpha})$. Furthermore, if $p \geq \omega/n$ and $k \geq k_1 = p^{-1}\omega^{1/2}$ hold for some sequence $\omega = \omega(n)$ with $1 \ll \omega \ll n$, then for all $i > k$ we have

$$\mathbb{P}\left[c(\Gamma_i \cap V_k) - \bar{c}(\Gamma_i \cap V_k) \geq \omega^{1/3} \;\middle|\; c(V_k) \geq \bar{\alpha}k\right] \geq 1 - \mathrm{e}^{-\omega^{1/3}} \quad . \tag{3}$$

Using this strong probability bound, we can prove that with high probability actually almost all agents make the correct choice in Phase III.

Lemma 6. *Let $\omega = \omega(n)$ be a sequence satisfying $1 \ll \omega \ll n$. Let $1/2 < \alpha < 1$, $\omega/n \leq p \leq 1/\omega$, $k_0 = p^{-1}\omega^{-1/2}$ and $k_1 = p^{-1}\omega^{1/2}$ be given. Then we have*

$$\mathbb{P}\left[c(V_n) \geq \left(1 - \omega^{-1/20}\right)n \;\middle|\; c(V_{k_1}) \geq \left(1 - k_0^{-1/18}\right)\alpha k_1\right] = 1 - o(1) \quad .$$

Combining Lemmas 3, 4 and 6, Theorem 1 follows immediately.

Proof (of Theorem 1 (i)). We consider the following three events

$$\begin{aligned} E_1: &\quad c(V_{k_0}) \ge \left(1 - k_0^{-1/9}\right) \alpha k_0 \ , \\ E_2: &\quad c(V_{k_1}) \ge \left(1 - k_0^{-1/18}\right) \alpha k_1 \ , \\ E_3: &\quad c(V_n) \ge \left(1 - \omega^{-1/20}\right) n \ . \end{aligned}$$

By Lemmas 3, 4 and 6, we have

$$\mathbb{P}\left[\overline{E}_3\right] \le \mathbb{P}\left[\overline{E}_3 \mid E_2\right] + \mathbb{P}\left[\overline{E}_2 \mid E_1\right] + \mathbb{P}\left[\overline{E}_1\right] = o(1) \ .$$

□

2.2 Proofs of Auxiliary Lemmas

Here we present the proofs of Lemmas 3, 4, 5, and 6 that were stated in the previous section. We will frequently make use of the following Chernoff tail bounds. The reader is referred to standard textbooks, e.g. [7, 8], for proofs.

Lemma 7. *Let $X_1, \ldots, X_n$ be independent Bernoulli trials with $\mathbb{P}[X_i = 1] = p_i$. Let $X = \sum_{i=1}^n X_i$ and $\mu = \mathbb{E}[X] = \sum_{i=1}^n p_i$. Then we have*

(a) $\mathbb{P}[X \ge (1+\delta)\mu] \le \mathrm{e}^{-\mu\delta^2/3}$ *for all* $0 < \delta \le 1$,
(b) $\mathbb{P}[X \le (1-\delta)\mu] \le \mathrm{e}^{-\mu\delta^2/2}$ *for all* $0 < \delta \le 1$,
(c) $\mathbb{P}[X \ge t] \le \mathrm{e}^{-t}$ *for all* $t \ge 7\mu$ *and*
(d) $\mathbb{P}[X \ge \mu + t] \le \mathrm{e}^{-\frac{t^2}{2(\mu+t/3)}}$ *for all* $t \ge 0$.

We first give the proof of Lemma 3, which makes an assertion on the number of correct decision-makers in Phase I.

Proof (of Lemma 3). For all $2 \le i < k_0$, we have

$$\begin{aligned} \mathbb{P}\left[|\Gamma_{i+1}| \ge 2\right] &= \sum_{j=2}^{i} \binom{i}{j} p^j (1-p)^j \le \sum_{j=2}^{i} (ip)^j \\ &\le k_0 p^2 \sum_{j=0}^{\infty} (k_0 p)^j \le \frac{k_0^2 p^2}{1 - k_0 p} \ . \end{aligned} \tag{4}$$

Let $A = \{v_i : |\Gamma_i| \le 1, \ 2 \le i \le k_0\}$, and $B = V_{k_0} \setminus A$ its complement. Note that all individuals in the set A make their decision solely based on their private signals. For individuals in B we don't know whether they have observed an imbalance $|\Delta| \ge 2$ in the actions of their neighbors and chosen to follow the majority, disregarding their private signals. But because of (4) and the definition of k_0 we have

$$\mathbb{E}\left[|B|\right] = \sum_{i=1}^{k_0} \mathbb{P}\left[|\Gamma_{i+1}| \ge 2\right] \le \frac{k_0^3 p^2}{1 - k_0 p} = {k_0}^3 p^2 \cdot (1 + o(1)) \ .$$

Let $\mathcal{E}$ denote the event that $|B| < {k_0}^3 p^2 \omega^{2/3} = k_0 \omega^{-1/3}$. As $\omega \to \infty$ we can apply Lemma 7 (c) and deduce that

$$\mathbb{P}\left[\ \overline{\mathcal{E}}\ \right] = \mathbb{P}\left[|B| \geq {k_0}^3 p^2 \omega^{2/3}\right] \leq \mathrm{e}^{-{k_0}^3 p^2 \omega^{2/3}} = \mathrm{e}^{-k_0 \omega^{-1/3}} = o(1) \tag{5}$$

by definition of k_0. Since by the decision rule in Definition 1 all individuals $v_i \in A_{k_0}$ follow their private signals, we have $\mathbb{E}\left[\bar{c}(A)\right] = (1-\alpha)\,|A|$. Clearly, we have $|A| \leq k_0$, and conditional on $\mathcal{E}$, we have $|A| \geq k_0\left(1 - \omega^{-1/3}\right)$. Therefore,

$$(1-\alpha)\, k_0 \left(1 - \omega^{-1/3}\right) \leq \mathbb{E}\left[\bar{c}(A)\ \middle|\ \mathcal{E}\right] \leq (1-\alpha)\, k_0 \ \ .$$

Using $k_0 \geq \omega^{1/2}$, Chernoff bounds imply

$$\begin{aligned}
\mathbb{P}\left[c(A) \leq \left(1 - k_0^{-1/9}\right)\alpha k_0 \ \middle|\ \mathcal{E}\right] &= \mathbb{P}\left[\bar{c}(A) \geq |A| - \left(1 - k_0^{-1/9}\right)\alpha k_0 \ \middle|\ \mathcal{E}\right] \\
&\leq \mathbb{P}\left[\bar{c}(A) \geq \left(1 + \frac{\alpha\, k_0^{-1/9}}{1-\alpha}\right)\mathbb{E}\left[\bar{c}(A)|\mathcal{E}\right] \ \middle|\ \mathcal{E}\right] \\
&\leq \mathrm{e}^{-\frac{\alpha^2 k_0^{-2/9}}{3(1-\alpha)^2}\mathbb{E}[\bar{c}(A)|\mathcal{E}]} = \mathrm{e}^{-\Theta\left(k_0^{7/9}\right)} = o(1) \ \ .
\end{aligned}$$

Thus, we have

$$\mathbb{P}\left[c(V_{k_0}) \geq \left(1 - k_0^{-1/9}\right)\alpha k_0 \ \middle|\ \mathcal{E}\right] \geq \mathbb{P}\left[c(A) \geq \left(1 - k_0^{-1/9}\right)\alpha k_0 \ \middle|\ \mathcal{E}\right] = 1 - o(1) \ \ .$$

Since

$$\mathbb{P}\left[c(V_{k_0}) \geq \left(1 - k_0^{-1/9}\right)\alpha k_0\right] \geq \mathbb{P}\left[c(V_{k_0}) \geq \left(1 - k_0^{-1/9}\right)\alpha k_0 \ \middle|\ \mathcal{E}\right] \cdot \mathbb{P}\left[\mathcal{E}\right] \ \ ,$$

we conclude with (5) $\mathbb{P}\left[c(V_{k_0}) \geq \left(1 - k_0^{-1/9}\right)\alpha k_0\right] = 1 - o(1)$. □

Before we proceed with the proof of Lemma 4, we need to state and prove a slightly stronger version of Lemma 1 in Sect. 1.2.

Lemma 8. *For every $1/2 < \alpha < 1$ there exists an $\varepsilon > 0$ such that if we have $c(V_k) \geq (1-\varepsilon)\alpha k$ for $k \geq 1$, then for all $i > k$ with $\Gamma_i \subseteq V_k$ we have*

$$\mathbb{P}\left[\mathrm{ch}(v_i) = \theta\right] \geq \alpha \ \ .$$

Proof (of Lemma 8). Let $c(V_k) = \bar{\alpha} k$ for some constant $\bar{\alpha} > 0$. Furthermore, let

$$\Delta \ = \ c(V_k \cap \Gamma_i) \ - \ \bar{c}(V_k \cap \Gamma_i)$$

be the difference in the number of neighbors of agent i in $C(V_k)$ and in $\overline{C}(V_k)$, and let $p_j = \mathbb{P}\left[\Delta = j\right]$ denote the probability that this difference is exactly j. Let $\ell_1 = \min\{\bar{\alpha}k, (1-\bar{\alpha})k + j\}$ and $\ell_2 = (1-\bar{\alpha})k \leq \ell_1$. Then for all $j \geq 2$, we have

$$p_j = \sum_{s=j}^{\ell_1} \binom{\bar{\alpha}k}{s}\binom{(1-\bar{\alpha})k}{s-j} p^{2s-j}(1-p)^{k-(2s-j)}$$

and

$$p_{-j} = \sum_{s=j}^{\ell_2} \binom{(1-\bar{\alpha})k}{s} \binom{\bar{\alpha}k}{s-j} p^{2s-j} (1-p)^{k-(2s-j)} \ .$$

For $r \geq s \geq 1$, let $r^{\underline{s}} = r\,(r-1)\ldots(r-s+1)$ be the falling factorial. For all $j \geq 1$ and $j \leq s \leq \ell_2$, we have

$$\begin{aligned}
\binom{\bar{\alpha}k}{s}\binom{(1-\bar{\alpha})k}{s-j} &= \frac{(\bar{\alpha}k)^{\underline{s}}((1-\bar{\alpha})k)^{\underline{s-j}}}{s!(s-j)!} \\
&= \frac{(\bar{\alpha}k)^{\underline{s-j}}((1-\bar{\alpha})k)^{\underline{s}}}{s!(s-j)!} \cdot \prod_{t=0}^{j-1} \frac{\bar{\alpha}k - s + j - t}{(1-\bar{\alpha})k - s + j - t} \\
&\geq \binom{(1-\bar{\alpha})k}{s}\binom{\bar{\alpha}k}{s-j} \cdot \left(\frac{\bar{\alpha}}{1-\bar{\alpha}}\right)^j \ .
\end{aligned}$$

Therefore we have

$$p_j \geq \left(\frac{\bar{\alpha}}{1-\bar{\alpha}}\right)^2 p_{-j} \qquad \forall\, j \geq 2 \ ,$$

and

$$\begin{aligned}
\mathbb{P}\left[\Delta \geq 2\right] &\geq \left(\frac{\bar{\alpha}}{1-\bar{\alpha}}\right)^2 \mathbb{P}\left[\Delta \leq -2\right] \\
&= \left(\frac{\bar{\alpha}}{1-\bar{\alpha}}\right)^2 \Big(1 - \mathbb{P}\left[-1 \leq \Delta \leq 1\right] - \mathbb{P}\left[\Delta \geq 2\right]\Big) \ .
\end{aligned}$$

Thus, we have

$$\mathbb{P}\left[\Delta \geq 2\right] \geq \frac{1}{1 + \left(\frac{1-\bar{\alpha}}{\bar{\alpha}}\right)^2} \Big(1 - \mathbb{P}\left[-1 \leq \Delta \leq 1\right]\Big) \ . \tag{6}$$

Let $\varepsilon < \frac{2(\alpha-1)+\sqrt{\alpha^{-1}-1}}{2\alpha-1}$. A straightforward calculation shows that

$$\frac{1}{1 + \left(\frac{1-\bar{\alpha}}{\bar{\alpha}}\right)^2} \geq \alpha \qquad \forall\, \bar{\alpha} \geq (1-\varepsilon)\alpha \ . \tag{7}$$

Because of the decision rule given in Definition 1, using (6) and (7) we have

$$\mathbb{P}\left[\mathrm{ch}(v_i) = \theta\right] = \alpha \mathbb{P}\left[-1 \leq \Delta \leq 1\right] + \mathbb{P}\left[\Delta \geq 2\right] \geq \alpha$$

for all $\bar{\alpha} \geq (1-\varepsilon)\alpha$. □

Note that Lemma 1 follows immediately from Lemma 8. Using Lemma 8, we now present the proof of Lemma 4, which asserts that roughly an α-fraction of correct decision-makers is maintained throughout Phase II.

Proof (of Lemma 4). We consider groups W_i of $m = p^{-1/3}\omega^{-1/4} \geq \omega^{1/12}$ individuals, resulting in $\ell = (k_1 - k_0)/m \leq k_1/m \leq k_1 p^{1/3}\omega^{1/4}$ groups between individuals k_0 and k_1. Let $\mathcal{E}_i$ be the event that there is at most one individual in W_i that has a neighbor in W_i, i.e. $|E(W_i)| \leq 1$. Let $\mathcal{E} = \mathcal{E}_1 \wedge \cdots \wedge \mathcal{E}_\ell$. Since $m^2 p = o(1)$, for n sufficiently large, we have

$$\mathbb{P}\left[\,\overline{\mathcal{E}}_i\,\right] \leq \sum_{j=2}^{\binom{m}{2}} \binom{\binom{m}{2}}{j} p^j \leq \sum_{j=2}^{\binom{m}{2}} m^{2j} p^j \leq m^4 p^2 \sum_{j=0}^{\infty} m^{2j} p^j$$
$$\leq \frac{m^4 p^2}{1 - m^2 p} \leq 2m^4 p^2 \ , \tag{8}$$

and

$$\mathbb{P}\left[\,\overline{\mathcal{E}}\,\right] \leq \ell \cdot \mathbb{P}\left[\,\overline{\mathcal{E}}_i\,\right] \leq 2m^4 p^2 \ell \leq 2\,p k_1 \omega^{-3/4} = 2\,\omega^{-1/4} \ . \tag{9}$$

We have

$$\mathbb{P}\left[c(V_{k_1}) < \left(1 - k_0^{-1/18}\right)\alpha k_1\right] \leq \mathbb{P}\left[c(V_{k_1}) < \left(1 - k_0^{-1/18}\right)\alpha k_1 \,\middle|\, \mathcal{E}\right] + \mathbb{P}\left[\,\overline{\mathcal{E}}\,\right] \ ,$$

and defining $\mathcal{A}_i$ as the event that $c(W_i) \geq \alpha\left(1 - k_0^{-1/18}\right)m$,

$$\mathbb{P}\left[c(V_{k_1}) < \left(1 - k_0^{-1/18}\right)\alpha k_1 \,\middle|\, \mathcal{E}\right] \leq$$
$$\mathbb{P}\left[c(V_{k_1}) < \left(1 - k_0^{-1/18}\right)\alpha k_1 \,\middle|\, \mathcal{E} \wedge \mathcal{A}_1 \wedge \cdots \wedge \mathcal{A}_\ell\right] +$$
$$\sum_{j=0}^{\ell-1} \mathbb{P}\left[\,\overline{\mathcal{A}}_j \,|\, \mathcal{E} \wedge \mathcal{A}_1 \wedge \cdots \wedge \mathcal{A}_j\right] \ .$$

Since $\mathcal{E} \wedge \mathcal{A}_1 \wedge \cdots \wedge \mathcal{A}_\ell$ implies $c(V_{k_1}) \geq \left(1 - k_0^{-1/18}\right)\alpha k_1$, we conclude

$$\mathbb{P}\left[c(V_{k_1}) < \left(1 - k_0^{-1/18}\right)\alpha k_1\right] \leq \sum_{j=0}^{\ell-1} \mathbb{P}\left[\,\overline{\mathcal{A}}_j \,|\, \mathcal{E} \wedge \mathcal{A}_1 \wedge \cdots \wedge A_j\right] + \mathbb{P}\left[\,\overline{\mathcal{E}}\,\right] \ . \tag{10}$$

Let $\bar{\alpha} = \left(1 - k_0^{-1/18}\right)\alpha$. The event $\mathcal{E} \wedge \mathcal{A}_1 \wedge \cdots \wedge \mathcal{A}_{j-1}$ means that before the individuals in group W_j have to make a decision, we have

$$c(V_{k_0+(j-1)m}) \geq \bar{\alpha}(k_0 + (j-1)m) \ ,$$

and there is at most one individual $w_j \in W_j$ with a neighbor in W_j that made his decision before w_j. Let $\widehat{W}_j = W_j \setminus w_j$ and $\hat{m} = m - 1$. Lemma 8 asserts, that there is an $\varepsilon > 0$ and $\bar{k} \geq 1$ (which both depend only on α), such that for all $k \geq \bar{k}$ we have $\mathbb{P}[\mathrm{ch}(v) = \theta] \geq \alpha$ for all $v \in \widehat{W}_j$, if $1 - k_0^{-1/18} < \varepsilon$. But since $k_0 \geq \omega$, for n sufficiently large we certainly have $k_0 \geq \bar{k}$ and $\bar{\alpha} \geq (1-\varepsilon)\alpha$. Hence, we have $\mathbb{E}\left[c(\widehat{W}_j)\right] \geq \alpha\,\hat{m}$. Chernoff bounds imply

$$\mathbb{P}\left[c(\widehat{W}_j) \leq \left(1 - 2\,k_0^{-1/18}\right)\alpha\hat{m}\right] \leq \mathrm{e}^{-2\,\alpha\hat{m}k_0^{-1/9}} \leq \mathrm{e}^{-\alpha m k_0^{-1/9}} \ .$$

Since for n sufficiently large we have

$$\mathbb{P}\left[\,\overline{\mathcal{A}}_j \mid \mathcal{E} \wedge \mathcal{A}_1 \wedge \cdots \wedge \mathcal{A}_{j-1}\right] = \mathbb{P}\left[c(W_j) \le \left(1 - k_0^{-1/18}\right)\alpha m\right] \le \mathbb{P}\left[c(\widehat{W}_j) \le \left(1 - 2\,k_0^{-1/18}\right)\alpha \hat{m}\right] \ ,$$

we also have

$$\mathbb{P}\left[\,\overline{\mathcal{A}}_j \mid \mathcal{E} \wedge \mathcal{A}_1 \wedge \cdots \wedge \mathcal{A}_{j-1}\right] \le \mathrm{e}^{-\alpha m k_0^{-1/9}} = \mathrm{e}^{-\alpha p^{-2/9}\omega^{-7/36}} \ .$$

Furthermore, since $\ell \le p^{-2/3}\omega^{3/4}$, we have

$$\sum_{j=1}^{\ell} \mathbb{P}\left[\,\overline{\mathcal{A}}_j \mid G \wedge \mathcal{A}_1 \wedge \cdots \wedge \mathcal{A}_{j-1}\right] \le \ell\ \mathrm{e}^{-\alpha p^{-\frac{2}{9}}\omega^{-\frac{7}{36}}} = o(1) \ . \tag{11}$$

Thus, because of (9), (10) and (11) we can conclude

$$\mathbb{P}\left[c(V_{k_1}) \ge \left(1 - k_0^{-\frac{1}{18}}\right)\alpha k_1\right] = 1 - o(1) \ .$$

□

We continue with the proof of Lemma 5, which is a slightly stronger version than Lemma 2 in Sect. 1.2.

Proof (of Lemma 5). Let $N_g = C(\Gamma_i \cap V_k)$ and $N_b = \overline{C}(\Gamma_i \cap V_k)$ be the neighbors of i in V_k who made the correct (respectively false) decision, and $n_g = |N_g|$. Let $n_b = |N_b|$. We have $n_g \sim \mathrm{Bin}(c(V_k), p)$ and $n_b \sim \mathrm{Bin}(k - c(V_k), p)$. Let $\mu_g = p\,\bar{\alpha}\,k$ and $\mu_b = p\,(1-\bar{\alpha})\,k$. Then we have $\mathbb{E}\left[n_g\right] \ge \mu_g$ and $\mathbb{E}\left[n_b\right] \le \mu_b$. Define

$$\delta = \frac{1}{3} - \frac{\mu_b}{3\mu_g} = \frac{2\,\bar{\alpha} - 1}{3\,\bar{\alpha}} \ . \tag{12}$$

We have

$$\begin{aligned}\mathbb{P}\left[n_g - n_b < \delta\mu_g\right] &= \mathbb{P}\left[n_g - (1-\delta)\mu_g < n_b - (1-2\delta)\mu_g\right] \\ &\le \mathbb{P}\left[n_g - (1-\delta)\mu_g < n_b - (1-2\delta)\mu_g \;\middle|\; n_g \ge (1-\delta)\mu_g\right] + \\ &\quad\ \mathbb{P}\left[n_g < (1-\delta)\mu_g\right] \ ,\end{aligned}$$

and thus

$$\mathbb{P}\left[n_g - n_b < \delta\mu_g\right] \le \mathbb{P}\left[n_b > (1-2\delta)\mu_g\right] + \mathbb{P}\left[n_g < (1-\delta)\mu_g\right] \ . \tag{13}$$

The Chernoff bound in Lemma 7 (b) implies that

$$\mathbb{P}\left[n_g < (1-\delta)\mu_g\right] \le \mathbb{P}\left[n_g < (1-\delta)\mathbb{E}\left[n_g\right]\right] \le \mathrm{e}^{-\mathbb{E}[n_g]\delta^2/2} \le \mathrm{e}^{-\mu_g\delta^2/2} \ , \tag{14}$$

and since $1 - 2\delta - \mu_b/\mu_g = \delta$ by (12), we have

$$\begin{aligned}\mathbb{P}\left[n_b > (1-2\delta)\mu_g\right] &= \mathbb{P}\left[n_b > \mathbb{E}\left[n_b\right] + (1 - 2\delta - \mathbb{E}\left[n_b\right]/\mu_g)\,\mu_g\right] \\ &\leq \mathbb{P}\left[n_b > \mathbb{E}\left[n_b\right] + (1 - 2\delta - \mu_b/\mu_g)\,\mu_g\right] \\ &= \mathbb{P}\left[n_b > \mathbb{E}\left[n_b\right] + \delta\mu_g\right] \enspace .\end{aligned}$$

Thus, using the Chernoff bound in Lemma 7 (d), we obtain

$$\mathbb{P}\left[n_b > (1-2\delta)\mu_g\right] \leq \exp\left(-\frac{\delta^2\mu_g^2}{2(\mathbb{E}\left[n_b\right] + \delta\mu_g/3)}\right) \leq \exp\left(-\frac{\delta^2\mu_g^2}{2(\mu_b + \delta\mu_g/3)}\right) \enspace .$$

Because of (12) we have $\mu_b + \delta\mu_g/3 \leq \mu_g$, and thus

$$\mathbb{P}\left[n_b > (1-2\delta)\mu_g\right] \leq \mathrm{e}^{-\mu_g\delta^2/2} \enspace . \tag{15}$$

Because of (13) - (15) and $\delta\mu_g = (2\bar{\alpha}-1)pk/3$, we conclude

$$\mathbb{P}\left[|C(\mathit{\Gamma}_i \cap V_k)| - |\overline{C}(\mathit{\Gamma}_i \cap V_k)| < \frac{2\bar{\alpha}-1}{3}pk\right] \leq 2\,\exp\left(-pk\frac{(2\bar{\alpha}-1)^2}{18\bar{\alpha}}\right) \enspace ,$$

and since $pk \geq \omega^{1/2}$, for n sufficiently large we have

$$\mathbb{P}\left[|C(\mathit{\Gamma}_i \cap V_k)| - |\overline{C}(\mathit{\Gamma}_i \cap V_k)| \geq \omega^{1/3}\right] \geq 1 - \mathrm{e}^{-\omega^{1/3}} \enspace .$$

□

Note that Lemma 2 is a straightforward corollary of Lemma 5; we omit the proof due to space restrictions. It remains to prove Lemma 6, which relies on the following lemma.

Lemma 9. *Let $\omega = \omega(n)$ be a sequence satisfying $1 \ll \omega \ll n$. Let $\frac{1}{2} < \bar{\alpha} < 1$, $\omega/n \leq p \leq 1/\omega$ and $k \geq k_1 = p^{-1}\omega^{1/2}$. Suppose we have*

$$c(V_k) \geq \bar{\alpha}k \enspace . \tag{16}$$

Then we have

$$\mathbb{P}\left[c(V_{2k} \setminus V_k) = \left(1 - \omega^{-1/19}\right)k\right] \geq 1 - \mathrm{e}^{-kp}\,.$$

The proof of Lemma 9 is deferred to Sect. B in the appendix due to space restrictions. Now we are ready to prove Lemma 6.

Proof (of Lemma 6). We consider subphases of increasing length. More precisely, the first subphase lasts to individual $2k_1$. The second subphase then lasts until individual $4k_1$. Thus, in general subphase j lasts from individual $k_1\,2^{j-1}$ until $k_1\,2^j$. We will have at most $\log(n-k_1) \leq \log n$ such subphases.

Inductively, for n sufficiently large we can employ Lemma 9 for each subphase, as assumption (16) iteratively holds. We obtain

$$\mathbb{P}\left[\frac{c(V_n \setminus V_{k_1})}{n-k_1} < 1-\omega^{-1/19}\right] \leq \sum_{j=0}^{\log n} \mathrm{e}^{-pk_1 2^j} = \mathrm{e}^{-pk_1} \sum_{j=0}^{\log n} \mathrm{e}^{-pk_1(2^j-1)}$$
$$\leq \mathrm{e}^{-pk_1} \sum_{j=0}^{\infty} \mathrm{e}^{-pk_1 j} = \frac{\mathrm{e}^{-pk_1}}{1-\mathrm{e}^{-pk_1}} = o(1) \quad .$$

Since $k_1 \leq n\omega^{-1/2}$, note that $c(V_n \setminus V_{k_1}) \geq (n-k_1)\big(1-\omega^{-1/19}\big)$ implies

$$c(V_n) \geq \big(1-k_1/n\big)\big(1-\omega^{-1/19}\big)n \geq \big(1-\omega^{-1/20}\big)n$$

for n sufficiently large. Thus, we conclude

$$\mathbb{P}\Big[c(V_n) \geq \big(1-\omega^{-1/20}\big)n\Big] = 1-o(1) \quad .$$

□

3 Numerical Experiments

The statements in Theorem 1 are asymptotic, asserting the emergence of informational cascades in the limit. As our numerical experiments show, these phenomena can be observed even with moderately small populations.

We conducted experiments with varying population size n and edge probability $p = p(n)$. For each value of n and p, we sampled $N = 2000$ instances of random graphs $G = G_{n,p}$ and of private signals $s(v_i)$, $v_i \in V(G)$. The sequential decision process was evaluated on each of those instances following the decision rule in Definition 1. We identified an informational cascade in such an experiment if at least 95% of all agents opted for the same choice. We computed the relative frequency of informational cascades among the N samples for each value of n and p.

We ran the simulation for $\alpha = 0.75$, $n \in \{100 \cdot i : 1 \leq i \leq 20\}$, and three distinct sequences p. The results are plotted in Fig. 1. The solid and the dotted line represent the relative frequencies of correct and false cascades, respectively, for constant $p = 0.5$. In accordance with Theorem 1(ii), both events occur with constant frequency independent of the population size. The dashed and the dash-dotted line represent the relative frequencies of correct cascades for $p = 1/\log n$ and $p = n^{-1/2}$, respectively. Confirming Theorem 1(i) those plots approach 1 as n grows.

Acknowledgment

Julian Lorenz was partially supported by UBS AG. We would like to thank Konstantinos Panagiotou and Florian Jug for helpful discussions.

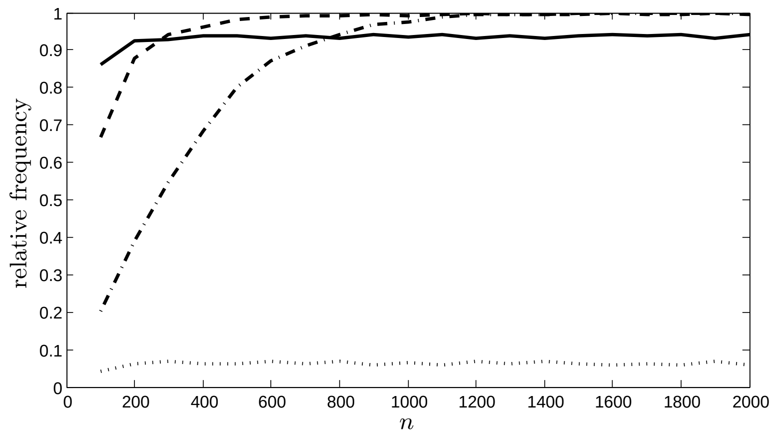

Fig. 1. Simulation results for $\alpha = 0.75$. The plot shows the relative frequencies of correct cascades for different values of the edge probability p as a function of n: $p = 0.5$ (solid line), $p = 1/\log n$ (dashed line), and $p = n^{-1/2}$ (dash-dotted line). The dotted line represents the relative frequency of incorrect cascades for $p = 0.5$.

References

1. Bikhchandani, S., Hirshleifer, D., Welch, I.: A theory of fads, fashion, custom, and cultural change in informational cascades. Journal of Political Economy **100**(5) (1992) 992–1026
2. Bikhchandani, S., Hirshleifer, D., Welch, I.: Learning from the behavior of others: Conformity, fads, and informational cascades. The Journal of Economic Perspectives **12**(3) (1998) 151–170
3. Banerjee, A.V.: A simple model of herd behavior. The Quarterly Journal of Economics **107**(3) (1992) 797–817
4. Surowiecki, J.: The Wisdom of Crowds. Anchor (2005)
5. Brin, S., Page, L.: The anatomy of a large-scale hypertextual Web search engine. Computer Networks and ISDN Systems **30**(1–7) (1998) 107–117
6. Erdős, P., Rényi, A.: On random graphs. I. Publ. Math. Debrecen **6** (1959) 290–297
7. Bollobás, B.: Random graphs. Second edn. Volume 73 of Cambridge Studies in Advanced Mathematics. Cambridge University Press, Cambridge (2001)
8. Janson, S., Łuczak, T., Rucinski, A.: Random graphs. Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience, New York (2000)
9. Barabási, A.L., Albert, R.: Emergence of scaling in random networks. Science **286**(5439) (1999) 509–512
10. Çelen, B., Kariv, S.: Observational learning under imperfect information. Games Econom. Behav. **47**(1) (2004) 72–86
11. Banerjee, A., Fudenberg, D.: Word-of-mouth learning. Games Econom. Behav. **46**(1) (2004) 1–22
12. Gale, D., Kariv, S.: Bayesian learning in social networks. Games Econom. Behav. **45**(2) (2003) 329–346 Special issue in honor of Robert W. Rosenthal.
13. Watts, D.J.: A simple model of global cascades on random networks. Proc. Natl. Acad. Sci. USA **99**(9) (2002) 5766–5771 (electronic)

A A Few Remarks on Decision Rules

A.1 The Decision rule of Bikhchandani, Hirshleifer, and Welch

Bikhchandani, Hirshleifer, and Welch [1, 2] use a slightly different version of the individuals' local decision rule.

Definition 2 (Local decision rule in [1, 2]). *Suppose individual v_i has received the private signal $s(v_i)$, and, among his acquaintances $\Gamma(i)$, m_a chose option a and m_b chose option b. Let z be drawn uniformly at random from $\{a, b\}$. Then*

$$\mathrm{ch}(v_i) = \begin{cases} a & \text{if } m_a - m_b \geq 2 \quad \text{or} \quad (m_a - m_b = 1) \wedge (s(v_i) = a) \ , \\ b & \text{if } m_b - m_a \geq 2 \quad \text{or} \quad (m_b - m_a = 1) \wedge (s(v_i) = b) \ , \\ z & \text{if } (m_b - m_a = 1) \wedge (s(v_i) = a) \\ & \text{or } (m_a - m_b = 1) \wedge (s(v_i) = b) \ , \\ s(v_i) & \text{if } m_b - m_a = 0 \ . \end{cases}$$

As the following lemma shows, this decision rule yields inferior global behavior on the complete network $G = K_n$ compared to the decision rule given in Definition 1.

Lemma 10. *Suppose the network of acquaintanceships is $G = K_n$. Then the probabilities of ever entering a correct cascade are given by*

$$f_0 = \frac{\alpha^2}{1 - 2\alpha + 2\alpha^2} \quad \text{and} \quad g_0 = \frac{\alpha(1+\alpha)}{2(1-\alpha+\alpha^2)} \ ,$$

if all individuals employ the decision rule in Definition 1 and Definition 2, respectively. For all $1/2 < \alpha < 1$, we have $f_0 > g_0$.

Proof. We consider a Markov chain where the state variable Δ is the difference between correct and incorrect decision-makers. The decision rules in Definitions 1 and 2 yield transition probabilities as given in Fig. 2.

The probability of entering a correct informational cascade is the probability of being absorbed in state $\Delta \geq 2$ when starting in state $\Delta = 0$. Under the decision rule in Definition 1, for $-1 \leq i \leq 1$ let f_i be the probability of eventually being absorbed in state $\Delta \geq 2$ when starting in state $\Delta = i$. Analogously, let g_i be the corresponding probabilities under the decision rule in Definition 2. We obtain the systems of linear equations

$$f_1 = \alpha + (1-\alpha) f_0 \ , \quad f_0 = \alpha f_1 + (1-\alpha) f_{-1} \ , \quad f_{-1} = \alpha f_0 \ ,$$

and

$$g_1 = (1+\alpha)/2 + (1-\alpha) g_0/2 \ , \quad g_0 = \alpha g_1 + (1-\alpha) g_{-1} \ , \quad g_{-1} = \alpha g_0/2 \ ,$$

which yield

$$f_0 = \frac{\alpha^2}{1 - 2\alpha + 2\alpha^2} \quad \text{and} \quad g_0 = \frac{\alpha(1+\alpha)}{2(1-\alpha+\alpha^2)} \ ,$$

and it is straightforward to check that $f_0 > g_0$ for all $1/2 < \alpha < 1$. □

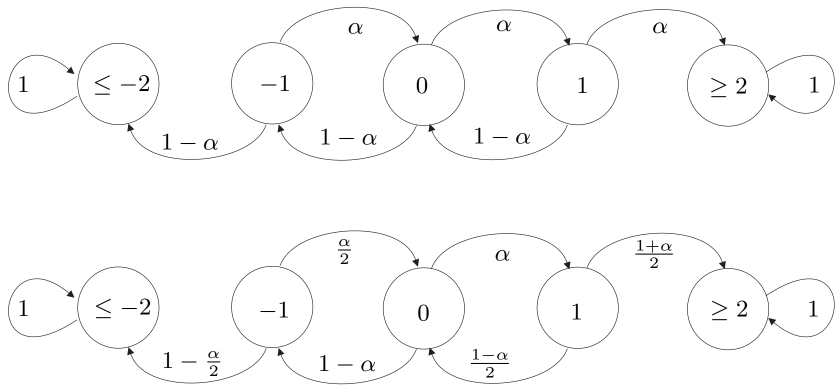

Fig. 2. Markov chains for the difference Δ of correct and incorrect decision-makers under the decision rules in Definition 1 (upper part) and in Definition 2 (lower part).

A.2 The Decision Rule in Definition 1

As one can show, the decision rule in Definition 1 is locally optimal under the following assumptions.

Lemma 11. *Let the social network be given as the complete graph on n vertices. Suppose that previous actions are observable in an aggregated form, and all individuals behave Bayes rational. Then by acting according to the rule in Definition 1 each agent maximizes the a-posteriori probability of making the correct decision.*

Proof. We consider a Markov chain with state variable

$$\Delta_j = \left|\left\{v \mid \operatorname{ch}(v) = \theta, v \in V_j\right\}\right| - \left|\left\{v \mid \operatorname{ch}(v) \neq \theta, v \in V_j\right\}\right| \ , \tag{17}$$

the difference after j individuals between correct and incorrect decision-makers, assuming that all individuals follow the decision rule in Definition 1. From the decision rule, for all $j \geq 0$ we have the transition probabilities

$$\mathbb{P}\Big[\Delta_{j+1} = \Delta_j + 1 \mid \Delta_j \geq 2\Big] = \mathbb{P}\Big[\Delta_{j+1} = \Delta_j - 1 \mid \Delta_j \leq -2\Big] = 1 \ , \tag{18}$$

and

$$\mathbb{P}\Big[\Delta_{j+1} = \Delta_j + 1 \,\Big|\, |\Delta_j| \leq 1\Big] = \alpha \ , \tag{19}$$

$$\mathbb{P}\Big[\Delta_{j+1} = \Delta_j - 1 \,\Big|\, |\Delta_j| \leq 1\Big] = 1 - \alpha \ . \tag{20}$$

For all $k \geq 1$ and $-k \leq i \leq k$, define

$$f_{i,k} = \mathbb{P}\Big[\Delta_k = i\Big] \ .$$

The probabilities $f_{i,j}$ will be useful later to prove the local optimality of the decision rule. From the transition probabilities (18)-(20), we will first compute $f_{i,j}$ explicitly. Since the first two individuals always decide independently, we have

$$f_{1,1} = \alpha,\ f_{1,-1} = 1-\alpha,\ f_{2,2} = \alpha^2,\ f_{2,0} = 2\alpha(1-\alpha) \text{ and } f_{2,-2} = (1-\alpha)^2 \ . \quad (21)$$

In order to have $\Delta_{2k+2} = 0$, we must have $\Delta_{2k} = 0$ and that the actions of the two individuals in $V_{j+2} \setminus V_j$ are a and b in any order. Thus, we have

$$f_{2k+2,0} = 2\alpha(1-\alpha)\, f_{2k,0} \ ,$$

and because of $f_{2,0} = 2\alpha(1-\alpha)$ we conclude

$$f_{2k,0} = 2^k \alpha^k (1-\alpha)^k \ . \quad (22)$$

By similar inductive reasoning with (21) as the base case and using (18)-(20) for the inductive step, we obtain

$$\begin{aligned} f_{2k-1,1} &= 2^{k-1}\alpha^k(1-\alpha)^{k-1} \ , \quad & f_{2k-1,-1} &= 2^{k-1}\alpha^{k-1}(1-\alpha)^k \ , \\ f_{2k,2} &= 2^{k-1}\alpha^{k+1}(1-\alpha)^{k-1} \ , \quad & f_{2k,-2} &= 2^{k-1}\alpha^{k-1}(1-\alpha)^{k+1} \ . \end{aligned}$$

Therefore, we have

$$\frac{f_{2k-1,1}}{f_{2k-1,1} + f_{2k-1,-1}} = \alpha \quad \forall\, k \geq 1 \ , \quad (23)$$

and

$$\frac{f_{2k,2}}{f_{2k,2} + f_{2k,-2}} = \frac{\alpha^2}{\alpha^2 + (1-\alpha)^2} \quad \forall\, k \geq 1 \ . \quad (24)$$

Because of (18), we have for $j \geq 2$ and $3 \leq i \leq j+1$

$$\frac{f_{j+1,i}}{f_{j+1,i} + f_{j+1,-i}} = \frac{f_{j,i-1}}{f_{j,i-1} + f_{j,-(i-1)}} \ , \quad (25)$$

and thus, inductively by (24) and (25), for all $j \geq 2$ and $2 \leq i \leq j$ we have

$$\frac{f_{j,i}}{f_{j,i} + f_{j,-i}} = \begin{cases} \frac{\alpha^2}{\alpha^2+(1-\alpha)^2} & \text{if } j \equiv i \mod 2 \ , \\ 0 & \text{otherwise} \ . \end{cases} \quad (26)$$

We will now prove that the decision rule in Definition 1 yields the locally optimal decision for each individual v_j, $1 \leq j \leq n$. For v_1, having no observations, the optimal decision is to follow his private signal, since $\alpha > 1/2$. Suppose now that the individual v_{j+1} has to make his decision, observing m_a individuals that made the choice a and m_b that made the choice b. By the induction hypothesis, we can assume that all his j predecessors followed the rule in Definition 1. We distinguish the cases j even and j odd.

j even: Note that $|m_a - m_b|$ is always even. If $m_a = m_b$, then v_{j+1} cannot learn anything about the correct decision by the observation of his predecessors. Since his signal $s(v_{j+1})$ is correct with $\alpha > 1/2$, his optimal choice is $s(v_{j+1})$ in that case. On the other hand, if $|m_a - m_b| \geq 2$, we have

$$\mathbb{P}\left[\theta = a \mid m_a - m_b = 2i\right] = \frac{\mathbb{P}\left[\Delta_j = 2i\right]}{\mathbb{P}\left[\Delta_j = 2i\right] + \mathbb{P}\left[\Delta_j = -2i\right]} = \frac{f_{j,2i}}{f_{j,2i} + f_{j,-2i}} \ ,$$

and also

$$\mathbb{P}\left[\theta = b \mid m_b - m_a = 2i\right] = \frac{\mathbb{P}\left[\Delta_j = 2i\right]}{\mathbb{P}\left[\Delta_j = 2i\right] + \mathbb{P}\left[\Delta_j = -2i\right]} = \frac{f_{j,2i}}{f_{j,2i} + f_{j,-2i}} \ .$$

Because of (26), decision rule in Definition 1 gives individual v_{j+1} a probability of making the correct choice of

$$\mathbb{P}\left[\mathrm{ch}(v_{j+1}) = \theta\right] = \frac{\alpha^2}{\alpha^2 + (1-\alpha)^2} > \alpha \qquad \text{for all } \frac{1}{2} < \alpha < 1 \ . \tag{27}$$

Since this yields a confidence strictly larger than the confidence α of his private signal, the decision rule in Definition 1 is indeed locally optimal.

j odd: $|m_a - m_b|$ is always odd, and $m_a = m_b$ can never occur. By analogous reasoning as in the case of j even, from (23) and (26) we obtain that the decision rule in Definition 1 is also locally optimal in the case of j odd, which completes the proof of the inductive step.

□

Remark 1. The alternative decision rule in Definition 2 is a locally optimal strategy for the individuals as well. The only difference between Definition 1 and Definition 2 is the coin flipping if an individual v_i observes a thin majority $|m_a - m_b| = 1$ in conjunction with $s(v_i)$ contradicting this majority vote. From (23) we see that such a thin majority vote has the same significance α to be correct, exactly the same as v_i's private signal. If the majority vote and $s(v_i)$ do not coincide, both choices a and b have therefore the same a-posteriori probability to be correct. Thus, flipping a coin is a locally optimal decision, as well as following the private signal as suggested by the rule in Definition 1. The reason that Definition 1 yields better global behavior on $G = K_n$ (as shown in Lemma 10) is that if individuals follow their private signals when $|m_a - m_b| = 1$ instead of flipping a coin, a greater information externality is provided, benefiting subsequent agents.

B Proof of Lemma 9

Let $m = p^{-1} \geq \omega$ and $\ell = k/m = kp \geq \omega^{1/2}$. For all $i = 0, \ldots, \ell - 1$, let $W_i = \{v_j \mid k + iw \leq j \leq k + (i+1)m\}$. That is, we consider ℓ groups until $2k$ agents have decided. Let A_i be the indicator variable defined by

$$A_i = \begin{cases} 1 & \text{if} \quad c(W_i) < \left(1 - \omega^{-1/18}\right) m \ , \\ 0 & \text{otherwise} \ . \end{cases}$$

Let $\mathcal{C}_i$ be the event that $\sum_{j=1}^{i} A_j \leq \ell\omega^{-1/4}$. We will first show that

$$\mathbb{P}\left[A_i = 1 \;\middle|\; \mathcal{C}_{i-1}\right] \leq 2\,\mathrm{e}^{-\omega^{1/3}} \quad . \tag{28}$$

To prove (28), let $B_i = \left\{w \in W_i : |\Gamma_G(w) \cap W_i| \geq \omega^{2/9}\right\}$ and let $\mathcal{E}_i$ be the event that $|B_i| \leq m^{8/9}$. Since $\overline{\mathcal{E}}_i$ implies $|E(W_i)| \geq \frac{1}{2}|B_i|\,\omega^{2/9} \geq \frac{1}{2}m^{8/9}\,\omega^{2/9}$, we have

$$\mathbb{P}\left[\,\overline{\mathcal{E}}_i\,\right] \leq \mathbb{P}\left[|E(W_i)| \geq \frac{1}{2}m^{8/9}\,\omega^{2/9}\right] \quad ,$$

and clearly $\mathbb{E}\left[E(W_i)\right] = pm(m-1)/2 = (p^{-1}-1)/2$. For n sufficiently large, we have $m^{\frac{8}{9}}\omega^{\frac{2}{9}}/2 \geq 7\,\mathbb{E}\left[E(W_i)\right]$, and hence the Chernoff bound in Lemma 7 (c) yields

$$\mathbb{P}\left[\,\overline{\mathcal{E}}_i\,\right] \leq \mathrm{e}^{-m^{8/9}\omega^{2/9}/2} \leq \mathrm{e}^{-\omega^{10/9}/2} \quad . \tag{29}$$

Note that $\mathcal{C}_{i-1}$ (i.e. $\sum_{j=1}^{i-1} A_j \leq \ell\omega^{-1/4}$) implies that for n sufficiently large there exists $1/2 < \tilde{\alpha} < \bar{\alpha}$ such that $c(V_{k+(i-1)m}) \geq \tilde{\alpha}\,(k+(i-1)m)$. That is, before the first individual of group W_i decides, we have at least a fraction $\tilde{\alpha}$ of correct decision-makers. By definition of B_i, an agent $v \in W_i \setminus B_i$ has at most $\omega^{2/9}$ neighbors within W_i, and because of (3) in the statement of Lemma 5,

$$\mathbb{P}\left[\mathrm{ch}(v) = \theta \;\middle|\; \mathcal{C}_{i-1}\right] \geq 1 - \mathrm{e}^{-\omega^{1/3}} \qquad \forall\, v \in W_i \setminus B_i \ .$$

Let

$$\mu = \mathbb{E}\left[c(W_i \setminus B_i) \;\middle|\; \mathcal{E}_i \wedge \mathcal{C}_{i-1}\right] \;\geq \left(1 - \mathrm{e}^{-\omega^{1/3}}\right)\left(1 - m^{-1/9}\right) m \ .$$

Then, for n sufficiently large, we have $\mu \geq m/2 \geq \omega/2$. Lemma 7 (b) implies

$$\mathbb{P}\left[c(W_i \setminus B_i) \leq \left(1 - 2\omega^{-1/3}\right)\mu \;\middle|\; \mathcal{E}_i \wedge \mathcal{C}_{i-1}\right] \leq \mathrm{e}^{-2\omega^{-2/3}\mu} \leq \mathrm{e}^{-\omega^{1/3}} \quad .$$

For n sufficiently large, we have $\left(1 - \omega^{-1/18}\right) m \leq \left(1 - 2\omega^{-1/3}\right)\mu$. Together with $c(W_i) \geq c(W_i \setminus B_i)$ and $2\,\omega^{-2/3}\mu \geq \omega^{1/3}$, we therefore have

$$\mathbb{P}\left[c(W_i) \leq \left(1 - \omega^{-1/18}\right) m \;\middle|\; \mathcal{E}_i \wedge \mathcal{C}_{i-1}\right] \leq \mathrm{e}^{-\omega^{1/3}} \quad .$$

Thus, using (29), we obtain

$$\mathbb{P}\left[A_i = 1 \;\middle|\; \mathcal{C}_{i-1}\right] \leq \mathrm{e}^{-\omega^{1/3}} + \mathbb{P}\left[\,\overline{\mathcal{E}}_i\right] \leq 2\,\mathrm{e}^{-\omega^{1/3}} \quad ,$$

which completes the proof of (28).

Suppose $\sum_{j=1}^{\ell} A_j \geq \ell\omega^{-1/4}$. Consider the first $\ell\omega^{-1/4}$ groups W_i for which $A_i = 1$. For each of them, we clearly have $\sum_{j=1}^{i-1} A_j \leq \ell\omega^{-1/4}$ and hence (28) holds for each of those groups. Therefore, for n sufficiently large

$$\begin{aligned}
\mathbb{P}\left[\sum_{j=1}^{\ell} A_j \geq \ell\omega^{-1/4}\right] &\leq \binom{\ell}{\ell\omega^{-1/4}} \left(2\,\mathrm{e}^{-\omega^{1/3}}\right)^{\ell\omega^{-1/4}} \\
&\leq \left(\mathrm{e}\;\omega^{1/4}\right)^{\ell\omega^{-1/4}} \left(2\,\mathrm{e}^{-\omega^{1/3}}\right)^{\ell\omega^{-1/4}} \\
&= \left(2\,\mathrm{e}^{1+(\log\omega)/4-\omega^{1/3}}\right)^{\ell\omega^{-1/4}} \leq \mathrm{e}^{-\ell} = \mathrm{e}^{-kp} \quad .
\end{aligned}$$

Since $\sum_{j=1}^{\ell} A_j < \ell\omega^{-1/4}$ implies

$$c(V_{2k} \setminus V_k) \geq \left(1-\omega^{-1/18}\right)\left(1-\omega^{-1/4}\right)k \geq \left(1-\omega^{-1/19}\right)k \ ,$$

we have

$$\mathbb{P}\left[c(V_{2k} \setminus V_k) \geq \left(1-\omega^{-1/19}\right)k\right] \geq 1-\mathrm{e}^{-kp} \ .$$

This concludes the proof of Lemma 9. □

C Proof of Theorem 1(ii)

First, we need to prove the following technical lemma, which is very similar to Lemma 5.

Lemma 12. *Let $0 < p \leq 1$ and $0 < \delta \leq p/(3+2p)$ be fixed. Then there exists a constant $j_0 = j_0(p,\delta) \geq 1$ such that, for all $j_0 \leq j \leq n-1$ and all $i > j$, we have*

$$\mathbb{P}\left[c(\mathit{\Gamma}_i \cap V_j) + \delta j - \bar{c}(\mathit{\Gamma}_i \cap V_j) \geq 2 \;\middle|\; c(V_j) \leq \delta j\right] \leq \delta/2 \ .$$

Proof. Note that by symmetry Lemma 5 can also be stated as a lower bound for the probability of an imbalance towards the false decision among the neighbors of agent v_i. Because of $\bar{c}(V_j) \geq \bar{\alpha} j$ with $\bar{\alpha} = 1-\delta > 1/2$, this yields for all $i > j$,

$$\mathbb{P}\left[\bar{c}(\mathit{\Gamma}_i \cap V_j) - c(\mathit{\Gamma}_i \cap V_j) \geq \frac{2\bar{\alpha}-1}{3}pj \;\middle|\; c(V_j) \leq \delta j\right] \geq 1 - 2\exp\left(-Cpj\right) \ ,$$

where $C = (2\bar{\alpha}-1)^2/(18\bar{\alpha})$. Note that $\delta \leq p/(3+2p)$ implies $p \geq 3\delta/(1-2\delta)$, and hence

$$\frac{2\bar{\alpha}-1}{3}pj = \frac{1-2\delta}{3}pj \geq \delta j \ .$$

Since p and δ are constants, there exists a constant $j_0 \geq 1$ such that we have $2\exp\left(-Cpj\right) \leq \delta/2$ for all $j \geq j_0$. Thus, we have

$$\begin{aligned}
&\mathbb{P}\left[c(\mathit{\Gamma}_i \cap V_j) + \delta j - \bar{c}(\mathit{\Gamma}_i \cap V_j) \geq 2 \;\middle|\; c(V_j) \leq \delta j\right] \\
&\quad = 1 - \mathbb{P}\left[\bar{c}(\mathit{\Gamma}_i \cap V_j) - c(\mathit{\Gamma}_i \cap V_j) > \delta j - 2 \;\middle|\; c(V_j) \leq \delta j\right] \\
&\quad \leq 1 - \mathbb{P}\left[\bar{c}(\mathit{\Gamma}_i \cap V_j) - c(\mathit{\Gamma}_i \cap V_j) \geq \delta j \;\middle|\; c(V_j) \leq \delta j\right] \\
&\quad \leq 2\exp\left(-Cpj\right) \leq \delta/2 \ .
\end{aligned}$$

This completes the proof of Lemma 12. □

Now, we can show Part (ii) of Theorem 1.

Proof (of (ii) in Theorem 1). Suppose p is a constant. We shall show that there exist positive constants γ and ϱ such that

$$\mathbb{P}\left[c(V_n) \leq \gamma(\log n)^2\right] \geq \varrho \ .$$

Applying Lemma 12 with p and $\delta = p/(3+2p)$ yields an integer j_0, independent of n. Let

$$\varepsilon = \frac{\delta^2}{6(1+\delta)} \ ,$$

and choose $k \geq j_0$ such that

$$\mathrm{e}^{-\varepsilon\delta k} < 1/2 \ .$$

For n sufficiently large, we can subdivide the set of agents into classes of increasing size as follows. Let $W_0 = \{1, \ldots, k\}$. For any $1 \leq i \leq (\log n)/2$, let $W_i = \{(1+\delta)^{i-1}k+1, \ldots, (1+\delta)^i k\}$. For any fixed $i \geq 1$ and $U_i = \bigcup_{j=0}^{i} W_j$, we claim that

$$\mathbb{P}\left[c(U_i) \geq \delta|U_i| \;\middle|\; c(U_{i-1}) \leq \delta|U_{i-1}|\right] \leq \mathrm{e}^{-\varepsilon|U_i|} \ . \tag{30}$$

In order to prove (30), we need to show that typically at most $\delta|W_i|$ agents in W_i make the correct choice. Recall that θ denotes the superior choice, and let $\bar{\theta}$ be the inferior alternative. Unfortunately, we expect many edges between the agents in W_i causing dependencies. We introduce the following random variables so as to overcome this issue. For any agent $w_i \in W_i$, let $m_\theta(w_i)$ be the random variable counting how many acquaintances of w_i in U_{i-1} opted for θ plus the number of agents in W_i, i.e.,

$$m_\theta(w_i) = c(\Gamma(w_i) \cap U_{i-1}) + |W_i| = c(\Gamma(w_i) \cap U_{i-1}) + \delta|U_{i-1}| \ .$$

Furthermore, let $m_{\bar{\theta}}(w_i)$ be the number of w_i's acquaintances in U_{i-1} who decided for $\bar{\theta}$, i.e.,

$$m_{\bar{\theta}}(w_i) = \bar{c}(\Gamma(w_i) \cap U_{i-1}) \ .$$

Define the following random variable

$$\widetilde{\mathrm{ch}}(w_i) = \begin{cases} \theta & \text{if } m_\theta(w_i) - m_{\bar{\theta}}(w_i) \geq 2 \ , \\ \bar{\theta} & \text{if } m_{\bar{\theta}}(w_i) - m_\theta(w_i) \geq 2 \ , \\ s(v_i) & \text{otherwise} \ . \end{cases}$$

Let $X(w_i) = \mathbf{1}[\widetilde{\mathrm{ch}}(w_i) = \theta]$ be an indicator random variable. Note that $X(w_i)$ and $X(u_i)$ are independent for all $u_i \neq w_i \in W_i$. Moreover, we have

$$X(W_i) = \sum_{w_i \in W_i} X(w_i) \geq c(W_i)$$

since $\mathrm{ch}(w_i) = \theta$ implies $\widetilde{\mathrm{ch}}(w_i) = \theta$ for all $w_i \in W_i$. Then, owing to Lemma 12, we have

$$\mathbb{P}\left[\widetilde{\mathrm{ch}}(w_i) = \theta \;\middle|\; c(U_{i-1}) \leq \delta|U_{i-1}|\right] \leq \delta/2 \ .$$

Hence, we conclude that

$$\mathbb{E}\left[X(W_i) \;\middle|\; c(U_{i-1}) \leq \delta|U_{i-1}|\right] \leq (\delta/2)|W_i| \ .$$

It follows from Chernoff bounds (see Lemma 7 (a)) that

$$\mathbb{P}\left[c(W_i) \geq \delta|W_i| \;\middle|\; c(U_{i-1}) \leq \delta|U_{i-1}|\right]$$
$$\leq \mathbb{P}\left[X(W_i) \geq \delta|W_i| \;\middle|\; c(U_{i-1}) \leq \delta|U_{i-1}|\right] \leq \mathrm{e}^{-\delta|W_i|/6} \;\; .$$

Therefore, with probability

$$1 - \mathrm{e}^{-\delta|W_i|/6} = 1 - \mathrm{e}^{-(\delta/6)\left(1-(1+\delta)^{-1}\right)|U_i|} = 1 - \mathrm{e}^{-\varepsilon|U_i|}$$

we have

$$c(U_i) = c(U_{i-1}) + c(W_i) \leq \delta|U_{i-1}| + \delta|W_i| = \delta|U_i| \;\; ,$$

where $\varepsilon = (\delta/3)\left(1-(1+\delta)^{-1}\right) = \delta^2/3(1+\delta) > 0$, and (30) is proved.

Applying (30) inductively, yields that, for any $r \geq 1$, we have

$$\begin{aligned}
&\mathbb{P}\left[c(U_r) \geq \delta|U_r| \;\middle|\; c(U_0) \leq \delta|U_0|\right] \\
&\quad \leq \sum_{i=1}^{r} \mathrm{e}^{-\varepsilon|U_i|} = \sum_{i=1}^{r} \mathrm{e}^{-\varepsilon(1+\delta)^i k} \leq \sum_{i=1}^{r} \mathrm{e}^{-\varepsilon(1+i\delta)k} \\
&\quad = \sum_{i=0}^{r-1} \mathrm{e}^{-\varepsilon(1+(i+1)\delta)k} = \mathrm{e}^{-\varepsilon(1+\delta)k} \sum_{i=0}^{r-1} \mathrm{e}^{-\varepsilon i\delta k} \\
&\quad \leq \frac{\mathrm{e}^{-\varepsilon(1+\delta)k}}{1-\mathrm{e}^{-\varepsilon\delta k}} < 1
\end{aligned}$$

by the choice of k. Since k is fixed, we have $c(U_0) \leq \delta|U_0|$ with positive probability. Hence, there exists a constant $\widetilde{\varrho} > 0$ such that we have

$$\mathbb{P}\left[c(U_r) \leq \delta|U_r|\right] \geq \mathbb{P}\left[c(U_r) \leq \delta|U_r| \;\middle|\; c(U_0) \leq \delta|U_0|\right] \cdot \mathbb{P}\left[c(U_0) \leq \delta|U_0|\right] \geq \widetilde{\varrho}$$

for all $r \geq 1$. In particular, there exists a constant $\bar{\alpha} = \bar{\alpha}(\delta)$, $1/2 < \bar{\alpha} < 1$, such that, setting $\ell = \lceil (\log n)^2 \rceil$, we have

$$\mathbb{P}\left[\bar{c}(V_\ell) \geq \bar{\alpha}\ell\right] \geq \widetilde{\varrho} \;\; .$$

By symmetry we can apply Lemma 2 with parameters p and $\bar{\alpha}$ and obtain constants $\tau \leftarrow \delta$ and $j_1 \leftarrow j_0$, such that, for all j, $j_1 \leq j < n$, we have

$$\mathbb{P}\left[\mathrm{ch}(v_{j+1}) = \bar{\theta} \;\middle|\; \bar{c}(V_j) \geq \bar{\alpha}j\right] \geq 1 - \mathrm{e}^{-\tau p j} \;\; .$$

For n sufficiently large, we have $\ell \geq j_1$ and, therefore,

$$\begin{aligned}
\mathbb{P}\left[c(V_n) \leq (1-\bar{\alpha})\ell\right] &\geq \mathbb{P}\left[c(V_n) \leq (1-\bar{\alpha})\ell \;\middle|\; \bar{c}(V_\ell) \geq \bar{\alpha}\ell\right] \cdot \mathbb{P}\left[\bar{c}(V_\ell) \geq \bar{\alpha}\ell\right] \\
&\geq \left(\prod_{\ell}^{n-1} \mathbb{P}\left[\mathrm{ch}(v_{j+1}) = \bar{\theta} \;\middle|\; \bar{c}(V_j) \geq \bar{\alpha}j\right]\right) \cdot \widetilde{\varrho} \\
&\geq \left(\prod_{\ell}^{n-1} \left(1 - \mathrm{e}^{-\tau p j}\right)\right) \cdot \widetilde{\varrho} \\
&\geq \left(1 - \mathrm{e}^{-\tau p (\log n)^2}\right)^n \cdot \widetilde{\varrho} \geq \left(1 - \frac{1}{n}\right)^n \cdot \widetilde{\varrho} \geq \widetilde{\varrho}/4 = \varrho
\end{aligned}$$

for $\varrho = \widetilde{\varrho}/4$. This completes the proof of the theorem. □