Naive learning through probability matching

Itai Arieli[*], Yakov Babichenko[†], Manuel Mueller-Frank[‡]

February 19, 2019

Abstract

We analyze boundedly rational updating in a repeated interaction network model with binary states and actions. We decompose the updating procedure into a deterministic stationary Markov belief updating component inspired by DeGroot updating and pair it with a random probability matching strategy that assigns probabilities to the actions given the underlying boundedly rational belief. This approach allows overcoming the impediments to consensus and naive learning inherent in deterministic updating functions in coarse action environments. We show that if a sequence of growing networks satisfies vanishing influence, then the eventual consensus action equals the realized state with a probability converging to one.

1 Introduction

Social networks play a very important role as a conduit of information. Individuals interact with their peers and adjust their behavior and opinions responding to the observed behavior and opinions. Observational learning serves as one explanation for such adjustment; here individuals draw inferences from observed behavior to

[*]Faculty of Industrial Engineering, Technion- Israel Institute of Technology, iarieli@tx.technion.ac.il

[†]Faculty of Industrial Engineering, Technion- Israel Institute of Technology, iarieli@tx.technion.ac.il

[‡]IESE Business School, University of Navarra, mmuellerfrank@iese.edu.

the underlying information giving rise to the behavior. The theoretical literature on observational learning takes two different approaches, the Bayesian approach and the boundedly rational approach. The Bayesian approach assumes that all agents are assumed to be fully rational in their updating procedure. Instead, in the boundedly rational approach agents apply simple heuristics when updating their behavior. The Bayesian approach[1] serves as a powerful benchmark but the indirect inferences of Bayesian agents on the information of unobserved agents via observed agents is computationally prohibitively complex, see Hazla, Jadbabaie, Mossel, and Rahimian [13].

This paper analyzes boundedly rational updating in a model of repeated interaction in a social network. Our focus is on the coarsest possible communication environment where the set of actions or opinions is binary. Possible examples are settings where agents provide a binary guess on whether or not an uncertain event will be realized (opinion environment), or where they face an adoption decision (action environment). We ask two key questions. First, whether and how boundedly rational agents can eventually achieve consensus. Second, we ask to which extent the consensus action aggregates information and how naive learning for large networks may be achieved. The crucial innovation relative to the literature that allows us to achieve positive results is that we disentangle the belief updating process from the process of taking actions. In particular, we introduce a novel approach to belief updating inspired by DeGroot updating and, more importantly, allow agents to randomize over the actions for a given subjective belief. This approach enables us to establish a binary action variant of Golub and Jackson's [9] naive learning result which they establish for the DeGroot model and hence the case of uncountable actions.

We consider the *canonical social learning setting* with a common prior, binary actions, binary states, and conditional iid signals. The underlying utility function assigns a utility of 1 if the agents period action matches the realized state, and a utility of 0 otherwise. Time is discrete. Agents are organized in a finite, strongly

[1] See for example, Gale and Kariv [8], Rosenberg et al. [21], Mueller-Frank [18], Mossel, Sly, and Tamuz. [16], and Mossel, Mueller-Frank, Sly, and Tamuz [17]

connected network and simultaneously select a binary action in every period.[2]

The behavior of each agent in every period can be broken down into two steps. First he forms a belief over the state space which is identified by a real number between zero and one, corresponding to the probability assigned to one state. In the first period this belief is Bayesian and determined by a private signal that is conditional i.i.d. across agents. Instead, in every later period, the belief is subjective and formed in a boundedly rational manner. We assume that the belief of agent i in period t depends only on i's last period belief and the last period observed actions of his neighbors in the network (including himself). This assumption is typically called *imperfect recall* or *Markov property* in the literature.[3] Additionally, we assume that the updating behavior is *stationary*, i.e., time-independent. Let the state space be denoted by $\{0,1\}$. The Markov and stationarity assumptions imply that the belief updating procedure of any agent i can be described by an *updating function*

$$f_i : \{0,1\}^m \times [0,1] \to [0,1] .$$

Here m denotes the number of agents in the network. To give the network a role in the updating process, we assume that the updating function of any agents is *local*, i.e., invariant in the last period actions of non-neighbors and thus a function of only actions of neighbors. We emphasize that each agent observes only the realized actions of his neighbors, and not their beliefs respectively strategies. An updating system $\mathbf{f} = (f_1, ..., f_m)$ denotes the tuple of updating functions across all agents.

After forming his belief, each agent takes a binary action. The underlying utility function assigns a utility of 1 if the action matches the realized state and 0 otherwise. By *strategy* we denote the mapping from subjective beliefs to mixed strategies over the two actions. As for the case of the updating process, we assume that the strategy is stationary. We focus on one particular strategy known in par-

[2]We will use the term action and guess interchangeably. For the action environment think of the standard utility function for binary states and actions, where an action matching the realized state achieves utility of one, and a utility of zero for a mismatch.

[3]See Molavi, Tahbaz-Salehi and Jadbabaie [15] for the former and Mueller-Frank and Neri [19] for the latter.

allel literature as *probability matching*.[4] For a given belief, such a strategy selects a given action with a probability exactly equal to the belief of the corresponding state.

We aim to understand the asymptotic properties of the process of actions when pairing different belief updating systems with probability matching strategies. In particular we focus on *consensus* and *naive learning.* For a given network, *consensus* holds if there exists a finite time period after which all agent play the same action. Under the probability matching assumption, this is possible only if *extreme consensus* holds with respect to the subjective beliefs. That is, the subjective belief of all agents simultaneously converges to either one or to zero.

To define naive learning, consider a sequence of networks and corresponding updating systems. We say that *naive learning* holds, if the probability of the consensus action equaling the state for a given network, converges to one as the size of the network goes to infinity.

We first consider *ϵ-DeGroot updating systems.* As in the standard DeGroot model, we assign a stochastic weight matrix W to a given network such that for each agent weights sum up to one and non-neighbors have weight zero. The belief of an agent in any period is then equal to a weighted average between his last period belief (weight $1-\epsilon$) and a DeGroot weighted average of the vector of last period realized actions (weight ϵ).

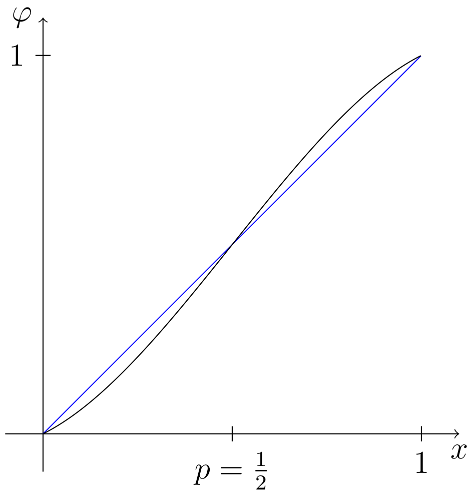
Proposition 1 shows that consensus holds for *ϵ-DeGroot* updating, for every $\epsilon \in (0,1]$. Moreover, the probability that the consensus is on action 1 equals what we call the *DeGroot consensus*, which equals the product of the centrality measure of the weight matrix W with the first period beliefs of the agents. This demonstrates in particular that under *ϵ-DeGroot* updating the probability with which the asymptotic consensus belief equals the state is bounded away from 1, for every ϵ and every network size. Thus naive learning fails for *ϵ-DeGroot* updating.

Our main result shows that incorporating *belief-extremization* into the *ϵ-DeGroot* updating allows agents not only to reach consensus but also to naively learn the

[4]See Vulkan [22] for a survey.

correct state. More concretely, we define an *extremization function* $\varphi : [0,1] \to [0,1]$ which maps every belief to a slightly more extreme belief with respect to the prior,[5] We consider (ϵ, δ)-*extremization DeGroot* updating where the belief of

Figure 1: The function $\varphi(x)$ for prior $p = \frac{1}{2}$.



agent i in any period is equal to a weighted average between his last period belief x_i (weight $1-\epsilon$), a DeGroot weighted average of the vector of last period realized actions (weight $\epsilon(1-\delta)$), and an extremization $\varphi(x_i)$ of his last period belief x_i (weight $\epsilon\delta$). Theorem 1 shows that for sufficiently small choice of the parameters ϵ and δ naive learning holds for (ϵ, δ)-*extremization DeGroot* updating systems, if the influence of the most influential agent, with respect to the centrality measure of the DeGroot matrix W, converges to zero as the size of the network goes to infinity.

The key technical tool in our analysis is the *stochastic approximation method* (see, e.g., Benaim [4] and Roth and Sandholm [20]), which allows us to approximate the discrete stochastic learning process via a deterministic differential equation.

[5] The actual choice of the mapping φ does not play a role in our arguments. For technical reasons we set φ to be differentiable.

2 Discussion and Related Literature

Roughly the literature on social learning under repeated interaction in social networks can be partitioned according to a 2×2 matrix. The row dimension corresponds to the assumption on the updating process, be it Bayesian or boundedly rational. The column dimension corresponds to the coarseness of the action space, either binary or uncountable. The majority of the existing literature analyzes the boundedly rational, rich action case and in particular focuses on the DeGroot model [6][6]. In the standard DeGroot model, the opinion space is equal to a real numbered interval. The opinion of any agent in a given period is formed by taking a weighted average of the last period's opinions she observed, that is, those of her neighbors and herself. Under weak conditions on the network, a consensus emerges in DeGroot updating systems, where all agents engage in DeGroot updating. Golub and Jackson [9] consider the case where the initial opinions of agents are equal to the state of the world plus a mean zero noise term which is independent across agents. They show that as the size of the network goes to infinity, the consensus opinion converges to the state of the world in probability if the influence of the most influential agent vanishes to zero. We show that the same condition of vanishing influence enables naive learning in a binary action variant of the DeGroot model when paired with probability matching.

For a continuous action set, a recent work by Levy and Razin [14] show that under the *Bayesian peer influence* heuristic naive learning holds.[7] A follow up work by Arieli, Babichenko and Shlomov [1] builds upon Levy and Razin [14] and characterizes a condition for a class of information structures under which naive learning holds for the same heuristic in any information structure that belongs to the class.

In the boundedly rational binary action literature, the two most related papers are the following. Mueller-Frank and Neri [19] analyze the case of binary actions

[6]See, for example, DeMarzo, Vayanos, and Zwiebel [7], and Golub and Jackson [9].

[7]The Bayesian peer influence heuristic treats observed actions as if they were based only on the private signal realization of the agent. The same heuristic is also called *naive inference* by Eyster and Rabin [11] and Quasi-Bayesian updating by Mueller-Frank and Neri [19].

for general deterministic action dynamics on finite networks. They show that for any deterministic stationary Markov updating system to achieve consensus and information aggregation very strong conditions on the information structure and/or the network structure need to be satisfied. As we show, the results for random updating system differ substantially, as we show information aggregation to be possible without any conditions beyond conditional iid signals and vanishing influence. Chandrasekhar, Larreguy and Xandri [5] analyze what they call the *DeGroot action model.* Implicitly their model coincides with our uniform weighting 1-DeGroot model paired with a deterministic strategy that selects the more likely action. They show that naive learning fails within the DeGroot action model due to so called clans, which are groups of agents who each have more edges within than beyond the group. Our random strategy of probabilistic matching, however, prevents clans to remain stuck on an incorrect action thus overcoming the impediment to information aggregation identified by Chandrasekhar et al. [5].[8]

In the Bayesian literature on binary actions the most related papers are Mossel, Sly, and Tamuz [16] and Mossel, Mueller-Frank, Sly and Tamuz [17]. The former provides a sufficient condition on infinite networks such that asymptotically all agents select the optimal action with probability one. The latter shows that if signals are unbounded then asymptotic consensus implies that the consensus action aggregates all private information. In comparison to these Bayesian results on the canonical social learning setting, our naive learning result holds even for some networks that are not captured by the sufficient condition established in Mossel, Sly and Tamuz [16][9], and also holds for all informative signal structures, bounded as well as unbounded.

There is a vast amount of experimental evidence for probability matching.[10] This raises questions as probability matching violates expected utility maximization. In the words of Kenneth Arrow [2]:

[8]The conceptual role of clans in preventing information aggregation also appears in Mueller-Frank and Neri [19] in form of a more general formulation.

[9]One example is a $k-$regular graph where k grows with n unboundedly.

[10]For a survey see Vulkan [22].

"*We have here an experimental situation which is essentially of an economic nature in the sense of seeking to achieve a maximum of expected reward, and yet the individual does not in fact, at any point, even in a limit, reach the optimal behavior. I suggest that this result points out strongly the importance of learning theory, not only in the greater understanding of the dynamics of economic behavior, but even in suggesting that equilibria maybe be different from those that we have predicted in our usual theory.*"

Our contribution in context of the literature on probability matching is to merge the concept with a model of boundedly rational learning. We show that together they enable networks composed exclusively of boundedly rational agents to asymptotically aggregate information efficiently, even if in the coarsest possible communication environment with binary choices. This identifies a trade-off of probability matching. While in the short run it is not expected utility maximizing in the long run it enables agents to aggregate all private information and select the action that is optimal given the realized state.

3 The Model

A finite set of agents $V = \{1, ..., m\}$ is organized in strongly connected social network $G = (V, E)$, where E is the set of edges among agents.[11] We denote by N_i the set of neighbors of agent i, including agent i herself. Agents share a common prior over a binary state space $\Omega = \{0, 1\}$, where $p = \Pr[\omega = 1]$. The interaction among agents evolves in discrete time as follows. In period $t = 0$ the state of the world is drawn according to the common prior. At the beginning of period $t = 1$, each agent observes a private signal $s_i \in S$ which is drawn according to a state-dependent distribution $F_\omega \in \Delta(S)$ for the realized state, independently across agents. Thus the *information structure* of our model is described by the triple (p, F_0, F_1).

In each period $t \in \mathbb{N}$, $t \geq 1$, all agents simultaneously take a binary action

[11]A network is strongly connected if for every agent there exists a directed path connecting her to every other agent in the network.

where the action space equals the state space, $A = \Omega$. At the beginning of period t, each agent observes the actions of her neighbors of period $t-1$. Agents receive a period-payoff of 1 if their action matches the state, and a payoff of 0 otherwise,

$$u_i\left(a_i^t, \omega\right) = 1\left[a_i^t = \omega\right].^{12}$$

In period $t = 1$, each agent i forms a Bayesian posterior belief on the state space given his realized signal s_i. From period $t \geq 2$ and on, we assume that agents update their posterior belief in a boundedly rational fashion and allow them to play a mixed strategy. The latter is a crucial departure to the social learning literature where the focus has been on deterministic updating and action dynamics. This departue plays an important role in establishing our main result.

Thus when considering the behavior of an agent at a given time period t, he essentially performs two tasks. First, he forms his belief given his available information and, second he plays a mixed strategy given his subjective belief. In the following we denote the former by *updating* process and the latter by *strategy*. We now introduce a novel approach to both updating and strategies to the social learning literature.

Let x_i^t denote the subjective belief of agent i in period t, $x_i^t = \Pr_i^t[\omega = 1]$, $x_i^t \in [0,1]$. The *strategy* of an agent is a mapping from his subjective belief to a probability distribution over the binary action space which also can be denoted by $[0,1]$, i.e. $\sigma_i : [0,1] \to [0,1]$ where $\sigma_i(x_i^t) = \Pr_{\sigma_i}[a_i = 1 \,|x_i^t]$. Note that we assume the strategy of agents to be stationary, and thus invariant in time. This is a weak assumption as it is also satisfied by myopic Bayesian agents.[13] In the boundedly rational social learning literature on the binary action case, agents also apply a threshold belief rule that determines their pure action, see for example Chandrasekhar et al. [5], and Mueller-Frank and Neri [19]. We focus on one particular mixed strategy known as *probability matching* where the probability of action ω is equal to the subjective probability of state ω, i.e., $\sigma_i(x_i) = x_i =$

[12] $1[.]$ denotes the indicator function.

[13] A myopic Bayesian agent maximizes, in every period, his expected utility in that given period.

$\Pr_i[\omega = 1]$. Thus, but for the case of extreme beliefs $x_i = 0, 1$ and the case $x_i = \frac{1}{2}$, probability matching violates expected utility maximization.

We now turn to describe the process of belief updating. As is common in the literature, we assume that the subjective belief of agents is updated in a *Markovian* and *stationary* manner. That is, the belief in period $t+1$ is determined only by the belief and behavior in period t and the updating behavior is independent of time. Let $\mathbf{a}^t = (a_1^t, ..., a_m^t)$ denote the realized vector of period t actions. Our assumptions then imply that the updating procedure of any agent i in every period $t \geq 1$ can be described by a time-independent updating function $f_i(\mathbf{a}^t, x_i^t) = x_i^{t+1}$ where the domain equals $\{0,1\}^m \times [0,1]$ and the range $[0,1]$. Thus the belief x_i^t of agent i in period t, depends only on his last period belief x_i^{t-1} and the vector of last-period's observed actions. Note that f_i is invariant in the last period actions of non-neighbors, since they are not observed by i. A collection of Markov, stationary updating functions $\mathbf{f} = (f_1, ..., f_m)$ is denoted as Markov, stationary *updating system*. For the sake of brevity, in the following we denote such systems simply as updating systems without explicitly mentioning the Markov and stationary property.

We first analyze two particular updating functions which are binary action variants of the DeGroot model. To relate our analysis to the standard DeGroot model assume that each agent is endowed with a stochastic vector $\mathbf{w}_i$ such that all entries are between 0 and 1, sum up to 1 and where every non-neighbor of i has weight 0 and every neighbor (including i) has positive weight. A natural way to convert the approach of taking weighted averages of continuous opinions to beliefs over binary states, is to form the belief according to a weighted average of the last period observed actions. That is, a *DeGroot* updating function of agent i is defined as

$$f(\mathbf{a}, x) = \sum_{i \in V} w_{ij} a_i$$

Thus a *DeGroot updating system* $\mathbf{f} = (f_1, ..., f_m)$ is described by a stochastic matrix W where $W_{ij} = w_{ij}$.

Note that a DeGroot updating system is invariant in the last period random strategies and only a function of the realized actions. Instead, an ϵ-DeGroot updating function is a mixture between the last period belief and a DeGroot updating function. Formally, an *ϵ-DeGroot* updating function of agent i is defined as

$$f^{\epsilon}(\mathbf{a},x) = (1-\epsilon)x + \epsilon \sum_{i \in V} w_{ij} a_i.$$

Thus an *ϵ-DeGroot updating system* $\mathbf{f}$ is described by the pair (W, ϵ). The parameter ϵ can be intuitively understood as the *inertia* of the updating function. Clearly, a DeGroot updating system is equal to an ϵ-DeGroot updating system for $\epsilon = 1$. We can also write the ϵ-DeGroot processes recursively in a matrix form as follows:

$$\mathbf{x}^{t+1} = (1-\epsilon)\mathbf{x}^t + \epsilon W \mathbf{a}^t, \tag{1}$$

where $\mathbf{a}^t = (a_1^t, \dots, a_i^t)$ is the vector of actions taken by the agents at time t.

The literature on boundedly rational learning in networks mainly focuses on the asymptotic properties of the opinion process. In particular, the question is whether in the long run agents behavior converges to consensus and whether the consensus opinion aggregates the private information. We now turn to these questions in the following section.

4 Consensus

One of the prominent features of the standard DeGroot updating procedure is that in any strongly connected network agents asymptotically reach consensus. A DeGroot updating system is described by an aperiodic and irreducible weight matrix W and a continuum opinion space $\hat{A}$, say $\hat{A} = [0, 1]$. Let μ^W denote the centrality measure of W. For any first period opinion vector $\hat{\mathbf{a}}$, all agents converge to the following limit opinion

$$\langle \hat{\mathbf{a}}^1, \mu^W \rangle = \sum_{i \in V} \mu_i^W \times \hat{a}_i^1.$$

In our binary opinion environment, and for a belief updating system $\mathbf{f}$ paired with probability matching, we define consensus as follows.

Definition 1. Consider an updating system $\mathbf{f}$ on network G paired with the probability matching strategy. *Consensus* holds if with probability one there exists (random) time period t^* such that $a_i^t = a^* \in \{0,1\}$ for any $t \geq t^*$ and every agent $i \in V$.

Consider an aperiodic and irreducible weight matrix W and a corresponding ϵ-DeGroot updating system with $\epsilon \in (0,1]$. Again denote the centrality measure of W by μ^W. Our first result establishes the consensus properties of such stochastic updating systems.

Proposition 1. Let the strategy of all agents satisfy probability matching. Consensus holds for any ϵ-DeGroot updating system (W, ϵ) on network G, with $\epsilon \in (0,1)$. Moreover, for a first period belief vector $\mathbf{x}^1$, asymptotic agreement on action 1 occurs with probability

$$\langle \mathbf{x}^1, \mu^W \rangle = \sum_{i \in V} \mu_i^W \times x_i^1.$$

Thus other than the strong negative results that have been established for deterministic updating in binary action environments, our stochastic approach does achieve long run consensus with probability one. The necessity of the stochastic strategy for asymptotic consensus is immediately clear when considering the case of highly connected groups where each agent has more neighbors within the group rather than without. For belief threshold updating functions as employed in Chandrasekhar et al. [5], and Mueller-Frank and Neri [19] first period consensus in such groups cannot be broken over time. Thus with positive probability asymptotic disagreement holds for networks with highly connected groups. To overcome this random strategies are necessary. We show in Proposition 1 that ϵ-DeGroot updating paired with probability matching is sufficient for consensus. Additionally, similarly to the standard DeGroot model, the asymptotic consensus action can be characterized in terms of the first period belief vector and the centrality measure of the weight matrix. In the standard DeGroot model the limit action is simply a weighted average of the initial, real numbered actions. Instead, in the ϵ-DeGroot model the probability of the limit action being 1 is equal to the (centrality measure) weighted average of the first period beliefs.

Proposition 1 is proven using two lemmas. In Lemma 1 we show that the sequence of random variables

$$\left\{ f_t = \langle \mu^W, \mathbf{x}^t \rangle = \sum_{i \in [m]} x_i^t \mu_i^W \right\}_{t \in \mathbb{N}}$$

forms a martingale. Therefore, by the martingale convergence theorem, the sequence converges almost surely to a limit.

We then show in Lemma 2 that the martingale convergence implies that the sequence x_i^t converges to either 0 or 1 with probability one. This is possible only if, for any two agents i,j, the limit beliefs coincide with probability one, i.e., of $\lim_{t\to\infty} x_i^t = \lim_{t\to\infty} x_j^t$. Otherwise there would exist two neighbors i, j such that $\lim_{t\to\infty} x_i^t = 1 - \lim_{t\to\infty} x_j^t$, , with positive probability. We show this to be impossible. We further show that

$$\lim_{t\to\infty} x_i^t = \lim_{t\to\infty} x_j^t \in \{0, 1\}$$

implies that from some finite time onward any agent i plays the action which equals to $\lim_{t\to\infty} x_i^t \in \{0, 1\}$.

Furthermore, the fact that $E(\langle \mu^W, \mathbf{x}^t \rangle | \mathbf{x}^1) = \langle \mu^W, \mathbf{x}^1 \rangle$ holds for every $t \geq 1$, implies that the expectation of the limit $E(\langle \mu^W, \mathbf{x}^\infty \rangle | \mathbf{x}^1)$ equals $\langle \mu^W, \mathbf{x}^1 \rangle$. Since $\langle \mu^W, \mathbf{x}^\infty \rangle \in \{0, 1\}$ with probability one, we must have that $\langle \mu^W, \mathbf{x}^\infty \rangle = 1$ with probability $\langle \mu^W, \mathbf{x}^1 \rangle$ and $\langle \mu^W, \mathbf{x}^\infty \rangle = 0$ with probability $1 - \langle \mu^W, \mathbf{x}^1 \rangle$.

There are two properties of ϵ-DeGroot updating systems that are worthwhile to point out. The first property relates to the characterization of the asymptotic consensus action and its implication on information aggregation. To introduce the implication, note that for an information structure (p, F_0, F_1) the first period posterior beliefs of agents are conditional i.i.d. and distributed according to the state-dependent distribution $G_\omega \in \Delta[0, 1]$.

Corollary 1. Consider a network G, an ϵ-DeGroot updating systems (W, ϵ), an information structure (p, F_0, F_1) and let all agents engage in probability matching. The probability that the consensus action equals the realized state of the world is bounded away from 1 and equal to

$$pE_{G_1}[x_i^1] + (1-p)E_{G_0}[1 - x_i^1]. \tag{2}$$

The bound in Corollary 1 depends only on the information structure and is independent of ϵ and the network. Moreover, for any non-trivial information structure, i.e., which does not reveal the state with probability one, the term in equation (2) is bounded away from one. This directly implies that naive learning as in Golub and Jackson [9] fails for ϵ-DeGroot updating systems on binary environments.

One feature of the asymptotic behavior established in Proposition 1 that might be surprising on first glance is that the inertia parameter ϵ does not impact asymptotic behavior. However, the inertia parameter ϵ has an important implication for the path of opinions as we will show next. The following proposition links the path of subjective beliefs and hence mixed strategies to the centrality measure, depending on the inertia parameter ϵ.

Proposition 2. Let the strategy of all agents satisfy probability matching. For *any* $\eta > 0$ there exists a time T such that for every $M > T$ the probability $\mathbf{P}(\sup_{\frac{T}{\epsilon} \leq t \leq \frac{M}{\epsilon}} |\mathbf{x}^t - \mathbf{1}\langle \mathbf{x}^1, \mu^W \rangle| > \eta)$ decreases to 0 when ϵ approaches zero.

Thus the inertia parameter increases the time frame within which agents randomize according to the centrality measure of the weight matrix W. This feature will be used in the next section where we analyze the question of naive learning.

To see the logic behind Proposition 2, consider a continuous time process $\mathbf{y}(t)$ that is derived from our discrete processes $\mathbf{x}^t$ as follows:

$$\mathbf{y}(t) = \mathbf{x}^{\lfloor \frac{t}{\epsilon} \rfloor}.$$

Thus we view the process $\mathbf{y}(t)$ as a continuous-time piecewise-constant process where "jumps" occur according to the $\mathbf{x}^t$ process only at times $t = k\epsilon$, for some k. We then apply stochastic approximations techniques to $\mathbf{y}(t)$ (see Benaim [4] and Roth and Sandholm [20]) that allow us, for small values of ϵ, to approximate the behavior of the processes $\mathbf{y}(t)$ using a deterministic differential equation. We show that the solution to the differential equation is absorbed at the consensus point $\mathbf{1}\langle \mu^W, \mathbf{x}^1 \rangle$ which is the DeGroot consensus point. That is, the deterministic process which approximates the behavior of $\mathbf{y}(t)$ converges to the belief vector where the subjective belief of *any* agent is $\langle \mu^W, \mathbf{x}^1 \rangle$.

We conclude that $\mathbf{y}(t)$ spends arbitrarily large time near the consensus as ϵ approaches zero.

5 Naive Learning

In their classical paper, Golub and Jackson [9] have shown that under certain conditions DeGroot updating systems induce long run opinions to perfectly aggregate the dispersed private information in large networks. To relate our result to theirs, it is instructive to formally introduce the naive learning result established in Golub and Jackson [9].

In their model, there is a true state of the world $p \in [0,1]$ and the first period opinion a_i^1 of each agent is drawn independently from a distribution with mean μ and variance $\sigma^2 > 0$. This approach is equivalent to assuming that in the first period agents communicate their subjective expected value of the state given a private signal with mean equalling to the state plus a zero mean noise term. *Naive learning* holds if for a sequence of networks with DeGroot updating systems described by their weight matrix $\langle W_n \rangle_{n\in\mathbb{N}}$ the limit opinion $\hat{\mathbf{a}}^*(W_n)$ converges to the true state in probability, as the size of the network goes to infinity. Golub and Jackson [9] show that naive learning is satisfied if the maximal centrality weight $\mu_+(W_n) = \max_{i\in V} \mu_i^{W_n}$ converges to zero as the size of the network goes to infinity.

We first formulate a notion of naive learning for a given network. Recall that x_i^t doubles as agent i's subjective probability of state 1 in period t, and the probability with which he selects action 1 in that same period.

Definition 2. An updating system $\mathbf{f}$ induces *η-naive learning* if, with probability at least $1-\eta$ there exists a time t^* such that $a_t^i = \omega$ for every agent $i \in V$ and any time $t \geq t^*$.

This notion of naive learning provides a lower bound on the probability with which limit beliefs and hence also actions equal the realized state of the world. By Corollary 1, *η-learning* fails under ϵ-stochastic DeGroot updating systems, for sufficiently small η and every network size m. We now introduce our notion of naive learning as a limit property for a sequence of networks which increase in size, following the approach chosen by Golub and Jackson [9].

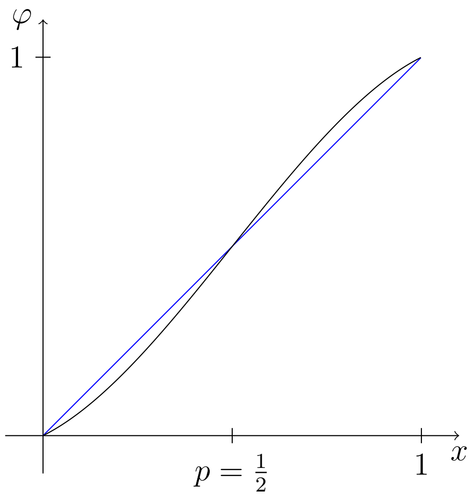
Definition 3. For a given sequence of networks $\langle G_m \rangle_{m\in\mathbb{N}}$, a sequence of corresponding updating systems $\langle \mathbf{f}^m \rangle_{m\in\mathbb{N}}$ satisfies *naive learning* if there exists a sequence $\langle \eta_m \rangle_{m\in\mathbb{N}} \in \mathbb{R}$ with $\lim\limits_{m\to\infty} \eta_m = 0$ such that $\mathbf{f}^m$ satisfies η_m-learning for every $m \in \mathbb{N}$.

Clearly, naive learning fails for any sequence of ϵ-DeGroot updating systems by Corollary 1. Thus, to achieve learning we require a different updating system. We next consider a variation of ϵ-DeGroot updating systems that induce slightly more extreme updated beliefs. We define an *extremization function* $\varphi : [0,1] \to [0,1]$ which maps every belief to a slightly more extreme belief (see Figure 1). Formally, let $\varphi(x) : [0,1] \to [0,1]$ be defined as follows:

$$\varphi(x) = x + x(1-x)(x-p),$$

and recall that p denotes the prior probability of state $\omega = 1$.

Figure 2: The function $\varphi(x)$ for $p = \frac{1}{2}$.



The important features of the extremization function $\varphi(x)$ that we rely on are: (1) $\varphi(x) > x$ for all $x \in (p,1)$, and (2) $\varphi(x) < x$ for all $x \in (0,p)$. Thus we can replace $\varphi(x)$ with any other Lipschitz continuous function from $[0,1]$ to $[0,1]$ with the above two properties.

We consider (ϵ,δ)-*extremization DeGroot* updating systems where the belief of agent i in any period is equal to a weighted average between his last period belief x_i (weight of $1-\epsilon$), a DeGroot weighted average of the vector of last period realized actions (weight of $\epsilon(1-\delta)$), and an extremization $\phi(x_i)$ of his last period belief x_i (weight of $\epsilon\delta$). Formally,

$$g(\mathbf{a},x) = (1-\epsilon)x_i + \epsilon\big((1-\delta)\sum_{j\in V} w_{ij}a_j + \delta\varphi(x_i)\big).$$

Thus, an (ϵ,δ)-*extremization DeGroot* updating system can be described by the triplet (W,ϵ,δ). Given g_i and a pair $(x_i^t, \mathbf{a}^t)$, the (subjective) period-$t+1$ belief of agent i is given by

$$x_i^{t+1} = g_i(\mathbf{a}^t, x_i^t).$$

Equivalently, we can write the processes in matrix form as follows

$$\mathbf{x}^{t+1} = (1-\epsilon)\mathbf{x}^t + \epsilon\big((1-\delta)W\mathbf{a}^t + \delta\vec{\varphi}(\mathbf{x}^t)\big), \tag{3}$$

where $\vec{\varphi}(\mathbf{x}) = (\varphi(x_1),\ldots,\varphi(x_n))$.

Recall from Golub and Jackson [9] that a sequence of weight matrices $\langle W_m\rangle_{m\in\mathbb{N}}$ satisfies *vanishing influence* if the corresponding sequence of centrality measures μ_m has the property that its maximal coordinate, $\max\limits_{i\in V}[\mu_m]_i$, converges to zero as m goes to infinity. Thus we say that a sequence of updating systems $\langle \mathbf{g}^m\rangle_{m\in\mathbb{N}}$ described by $\langle(W_m,\epsilon_m,\delta_m)\rangle_{m\in\mathbb{N}}$ satisfies *vanishing influence* if the sequence of weight matrices $\langle W_m\rangle_{m\in\mathbb{N}}$ does. We now can state our main result.

Theorem 1. Let the strategy of all agents satisfy probability matching. For a sequence of networks $\langle G_m\rangle_{m\in\mathbb{N}}$, a sequence of (ϵ,δ)-*extremization DeGroot* updating systems $\langle \mathbf{g}^m\rangle_{m\in\mathbb{N}}$ described by $\langle(W_m,\epsilon_m,\delta_m)\rangle_{m\in\mathbb{N}}$ satisfies *naive learning* if it satisfies vanishing influence.

We next outline the proof of Theorem 1. Consider a network G,with weight matrix W and associated centrality measure μ^W. Let F be the information structure and define l_F^G as follows

$$l_F^G = p\mathbf{P}(\sum_{i\in[m]} \mu_i^W x_i^1 > p|\omega=1) + (1-p)\mathbf{P}(\sum_{i\in[m]} \mu_i^W x_i^1 < p|\omega=0).$$

l_F^G represents the probability that the weighted average according to the centrality measure of the first period posterior belief lies above the prior p in state $\omega = 1$ and below the prior p in state 0. Since $E[x_i^1|\omega = 1] > p$ and $E[x_i^1|\omega = 0] < p$ for every agent $i \in V$ and since signals are conditionally independent, it follows from the weak law of large numbers that for large networks (G, E, W) with vanishing influence the term l_F^G approaches one. At a high level, we will show that, for an appropriate choice of ϵ and δ, the consensus action equals the state of the world with a probability that is arbitrarily close to l_F^G. Note that to prove this, it is sufficient to show that if $\sum_{i\in[m]} \mu_i^W x_i^1 > p$ then, for small enough $\delta > 0$ and $\epsilon > 0$, the stochastic process $\langle \mathbf{x}^t \rangle_{t\in\mathbb{N}}$ induced by $\mathbf{g}$ converges to $\mathbf{1}$ with arbitrarily high probability. Similarly if $\sum_{i\in[m]} \mu_i^W x_i^1 < p$, then the process $\langle \mathbf{x}^t \rangle_{t\in\mathbb{N}}$ induced by $\mathbf{g}$ needs to converges to $\mathbf{0}$ with arbitrarily high probability. In the remainder of the outline we will focus on the former case.

Essentially, a careful selection of ϵ and δ assures that convergence of $\langle \mathbf{x}^t \rangle_{t\in\mathbb{N}}$ to $\mathbf{1}$ occurs in two phases. In the first phase, $\langle \mathbf{x}^t \rangle_{t\in\mathbb{N}}$ reaches the centrality measure internal consensus on $\mathbf{1}\big(\sum_{i\in[m]} \mu_i^W x_i^1\big)$. That is, there exists a time period t such that $\mathbf{x}^t$ lies arbitrarily close to the belief vector $\mathbf{1}\big(\sum_{i\in[m]} \mu_i^W x_i^1\big)$, with arbitrary large probability. This is shown using a stochastic approximation technique. Under the ϵ-DeGroot process with $\delta = 0$ and ϵ sufficiently small, the approximation of the intermediate centrality consensus is shown to hold in Proposition 2. By selecting a δ sufficiently close to 0, we assure that $\langle \mathbf{x}^t \rangle_{t\in\mathbb{N}}$ approximates the ϵ-DeGroot process sufficiently to satisfy the desired closeness to $\mathbf{1}\big(\sum_{i\in[m]} \mu_i^W x_i^1\big)$.

In the second phase of the stochastic process $\langle \mathbf{x}^t \rangle_{t\in\mathbb{N}}$ convergence towards a consensus of $\mathbf{1}$ occurs. This part of the proof is divided in several steps. As a first step we introduce a δ-dependent differential equation $\dot{\mathbf{z}}^\delta$ that has the following property: for any $\delta \in [0, 1]$ if $\mathbf{z}^\delta(0) = \mathbf{x}$ satisfies $x_i > p$ for every $i \in V$, then $\lim_{t\to\infty} \mathbf{z}^\delta(t) = \mathbf{1}$. In particular, we show that if $\mathbf{z}^\delta(t) \in [\alpha, \beta] \subset (p, 1)$ then the derivative $\dot{z}_i^\delta$ is positive and bounded away from zero. As a second step, we show that by choosing ϵ sufficiently small, the stochastic process $\langle \mathbf{x}^t \rangle_{t\in\mathbb{N}}$ can be approximated by the deterministic process $\mathbf{z}^\delta$ for an arbitrary duration of time.

That is, if for some $\mathbf{x}^{t'}$ we have $\sum_{i\in[m]} \mu_i^W x_i^{t'} > p$ then $\langle \mathbf{x}^t \rangle_{t\in\mathbb{N}}$ gets arbitrarily close to $\mathbf{1}$ with arbitrary high probability. The complication here is to assure that ϵ and δ are appropriately chosen such that the respective deterministic process is approximated in the two phases of the stochastic process. As a third step, we show that if $\mathbf{x}^{t'}$ reaches a point close to $\mathbf{1}$ then, for any $\delta > 0$, then the probability of absorbance to $\mathbf{1}$ goes to one as the distance to $\mathbf{1}$ decreases.

6 Discussion of results and conclusion

The existing literature on boundedly rational updating dynamics so far has focused attention on deterministic Markov updating functions. For the case of binary actions, it has been shown that generally, efficient information aggregation fails. Mueller-Frank and Neri [19] identify three main impediments to information aggregation. First and foremost, at the level of an individual agent, the coarse nature of the action space generally disallows to retain, aggregate and diffuse local relevant information. Second, even if one agent might aggregate all relevant information, the stationary nature of the updating functions might prevent the diffusion of the optimal action. For example, highly connected small subgroups that initially agree on one action cannot be induced to switch their action. Last but not least, royal families or highly influential small groups might induce consensus on the incorrect action.

Our approach differs from the existing analysis in allowing for random strategies. More precisely, we decompose the updating procedure into a deterministic stationary Markov belief updating component and pair it with a random strategy which assigns probabilities to the actions given the underlying boundedly rational belief. The belief updating procedures we consider are binary action variants of the DeGroot model. We focus on a probability matching strategy where agents select a given action with a probability equal to their subjective belief of the corresponding state. We show that this approach overcomes the impediments to consensus and information aggregation inherent in deterministic updating functions. At a high level, we overcome the first impediment to learning by allowing each agent to keep track of one additional variable, his belief. We believe that our approach is intuitive and naturally founded in Bayesian updating. Bayesian agents equally would form a belief in each period prior to selecting an action as a function of their belief, but, clearly, the Bayesian updating procedure is much more complex than our stationary, Markov belief updating function. The impediment to information diffusion can be naturally overcome through random strategies as thus initially agreement on the wrong action in a highly connected group can be broken. However, as our result shows, information aggregation can still be prevented by highly influential groups. Nevertheless, by adjusting the weight matrices, for any sequence of networks we can construct a sequence of corresponding updating systems such that naive learning holds.

In regards to the literature on probability matching, we identify a trade-off of probability matching that was previously unknown (to the best of our knowledge): while it violates expected utility maximization in the short run, it induces information aggregation asymptotically in social network settings.

There are several interesting avenues for future work. Our analysis focuses on the canonical social learning environment with binary states and binary actions. A natural question is how one might adapt our approach to a general finite environment. A further important question concerns the robustness of our results to small perturbations to the updating functions. We leave this for future work.

7 Appendix

The following lemma is a fundamental observation that we use throughout. We let $\mathcal{F}_t$ be the sigma field generated by $\{x^\tau\}_{\tau \leq t}$.

Lemma 1. Consider the ϵ-DeGroot learning processes described in equation (1). The sequence of random variable $\{\langle \mu^W, \mathbf{x}^t \rangle = \sum_{i \in V} x_i^t \mu_i^W\}_t$ forms martingale with respect to $\{\mathcal{F}_t\}_t$.

Proof. To see this note that (1) implies that

$$E[\sum_{i\in V} x_i^{t+1}\mu_i^W|\mathcal{F}_t] = (1-\epsilon)\sum_{i\in V} x_i^t\mu_i^W + \epsilon\sum_{i\in V}\mu_i^W\sum_{j\in} w_{ij}x_j^t =$$
$$(1-\epsilon)\sum_{i\in V} x_i^t\mu_i^W + \epsilon(\mu^W)'Wx^t =$$
$$(1-\epsilon)\sum_{i\in V} x_i^t\mu_i^W + \epsilon\sum_{i\in V} x_i^t\mu_i^W. \quad (4)$$

Equation (4) follows from the fact that μ^W is a left eigenvector of W by definition.

□

Recall that $\mathbf{0}$ is the vector with m zeros and $\mathbf{1}$ is the vector with m ones. The following lemma shows that the ϵ-DeGroot model is absorbed to one of these vectors.

Lemma 2. The ϵ-DeGroot processes $\mathbf{x}^t$ of converges almost surely to either $\mathbf{1}$ or to $\mathbf{0}$.

Proof. Let $f_t = \langle\mu^W, x^t\rangle$. We can rewrite equation (1) as follows:

$$\mathbf{x}^{t+1} - \mathbf{x}^t = \epsilon(W\mathbf{a}^t - \mathbf{a}^t).$$

Therefore, since μ^W is an eigenvector of W we get

$$\langle\mathbf{x}^{t+1} - \mathbf{x}^t, \mu^W\rangle = f_{t+1} - f_t = \epsilon(\langle\mathbf{a}^t - \mathbf{x}^t\rangle.$$

Or equivalently that

$$f_{t+1} - f_t = \epsilon\sum_{i\in V}(a_i^t - x_i^t)\mu_i^W.$$

We note that conditional on the sigma field $\mathcal{F}_t$ the random variables $(x_i^t - a_i^t)_{i\in V}$ are independent and $Var(x_i^t - a_i^t) = x_i^t(1-x_i^t)$.

Therefore it follows that

$$E[Var(f_{t+1} - f_t)|\mathcal{F}_t] = \epsilon^2\sum_{i\in V}\mu_i^W x_i^t(1-x_i^t).$$

Let $\mu_m = \min\mu_i^W$, it follows that if $x_i^t \in [\eta, 1-\eta]$ for some $\eta > 0$, then $E[Var(f_{t+1} - f_t)|\mathcal{F}_t] \geq \mu_m\eta(1-\eta)\epsilon^2$. It follows from the martingale convergence theorem that $(f_t)_t$ is a convergent sequence and therefore we must have that

$$\lim_{t\to\infty} E[Var(f_{t+1} - f_t)|\mathcal{F}_t] = 0, \ a.s.$$

By the above, this entails that $\lim_{t\to\infty} x_i^t \in \{0,1\}$ for every $i \in V$.

It only left to show that $\lim_{t\to\infty} x_i^t = \lim_{t\to\infty} x_j^t$ with a probability one for any $i, j \in V$. To see this, note that otherwise we must have i, j such that $w_{ij} > 0$ and it holds with positive probability that $\lim_{t\to\infty} x_i^t = 1$ and $\lim_{t\to\infty} x_j^t = 0$. Therefore, in particular, $\lim_{t\to\infty} a_j^t = 1$. But since $x_i^{t+1} = (1-\epsilon)x_i^t + \epsilon \sum_{j\in V} w_{ij} a_j^t$ it must hold that $\liminf x_i^t \geq w_{ij}\epsilon$, a contradiction.

□

We next turn to prove Proposition 1.

Proof of Proposition 1. Proposition 1 follows directly from Lemma 1 and Lemma 2. We first show that $\mathbf{x}^t$ converges to $\mathbf{1}$ with probability $\langle \mu^W, \mathbf{x}^1 \rangle$ (and hence converges to $\mathbf{0}$ with probability $1 - \langle \mu^W, \mathbf{x}^1 \rangle$). To see this note that by the martingale property:

$$E[\lim_{t\to\infty} \langle \mu^W, \mathbf{x}^t \rangle | \mathcal{F}_1] = \langle \mu^W, \mathbf{x}^1 \rangle.$$

Therefore, since by Lemma 2 $\lim_{t\to\infty} \langle \mu^W, \mathbf{x}^t \rangle \in \{0,1\}$ we must have that this limit is 1 with probability $\langle \mu^W, \mathbf{x}^1 \rangle$.

We note that conditional on the event that $\lim_{t\to\infty} \mathbf{x}^t = \mathbf{1}$ we must have that the vector of actions $\mathbf{a}^t = \mathbf{1}$ for some time period t^* and on. To see this we note that at every time t where $a_i^t = 0$ we must have that $x_j^{t+1} \leq (1 - w_{ji})$. Since the network is strongly connected there exists an agent j for which $w_{ji} > 0$ for some agent j. Therefore, if $\lim_{t\to\infty} \mathbf{a}^t = \mathbf{1}$, then $a_i^t = 1$ for every i and all $t \geq t^*$ for some t^*. We note that conditional on the event $\lim_{t\to\infty} \mathbf{x}^t = \mathbf{1}$ we must have that the sequence of actions $\lim_{t\to\infty} \mathbf{a}^t = \mathbf{1}$. Similarly, conditional on the event that $\lim_{t\to\infty} \mathbf{x}^t = \mathbf{0}$ we must have that the vector of actions $\mathbf{a}^t = \mathbf{0}$ for some time period t^* and on. □

7.1 Stochastic Approximation

In order to prove Proposition 2 Theorem 1 use a stochastic approximation technique due to Benaim [4] and Roth and Sandholm [20]. Let $\mathbf{x}^t$ be an ϵ-dependent Markov process on a compact convex set $X = [0,1]^m$. Assume further that for any $\epsilon > 0$

$$x_i^{t+1} - x_i^t - \epsilon U_i^{t+1} = \epsilon g_i(x^t), \tag{5}$$

where U_i^{t+1} satisfies $E[U_i^{t+1}|\mathcal{F}_t] = 0$ and $g_i : X \to \mathbb{R}$ is a Lipschitz function. Assume further that $\mathbf{x}^t \in X$ for every $t \geq 0$ with probability one.

Consider the solution $z : \mathbb{R}_+ \to [0,1]^m$ to the following differential equation

$$\forall i \in V \ \dot{z}_i(t) = g_i(z(t)) \text{ and } z(0) = x^0. \tag{6}$$

Let $y(t) = x^{\lfloor \frac{t}{\epsilon} \rfloor}$. Proposition 3.1 in Benaim [4] and Theorem 3.1 in Roth and Sandholm [20] imply the following:

Proposition 3. For every $T > 0$ and $\eta > 0$

$$\lim_{\epsilon \to 0} \mathbf{P}(\sup_{0 \leq t \leq T} \|z(t) - y(t)\| > \eta \big| y(0) = z(0) = \mathbf{x}^1) = 0$$

uniformly in $x^0 \in X$

We note first that the ϵ-DeGroot learning process can be written as follows.

$$x_i^{t+1} - x_i^{t-1} - \epsilon U_i^{t+1} = \epsilon \sum_{j \in V} w_{ij}(x_j^t - x_i^t), \tag{7}$$

where $U_i^{t+1} = \sum_{j \in V} w_{ij}(a_j^t - x_j^t)$. Furthermore, since the expectation of a_j^t is x_j^t for any $j \in V$ by definition, we have that $E[U_i^{t+1}|\mathcal{F}_t] = 0$.

Therefore in our case the deterministic process $z : \mathbb{R}_+ \to [0,1]^m$ is the solution to the following differential equation:

$$\forall i \in V \ \dot{z}_i(t) = \sum_{j \in V} w_{ij}(z_j(t) - z_i(t)) \text{ and } z(0) = \mathbf{x}^1. \tag{8}$$

Alternatively we can write (8) as follows

$$\dot{z}(t) = (W - I)z \text{ and } z(0) = \mathbf{x}^1. \tag{9}$$

Proposition 3 implies that as ϵ decreases the stochastic processes $\{y(t)\}_{t \geq 0}$ lies arbitrarily close to the solution of the differential equation (8) starting at $z(0) = x(0)$. We let $\Phi : X \times \mathbb{R}_+ \to X$ be the semi-flow of the differential equation (8). That is, $\Phi(x,t)$ is the value $z(t)$ for the solution of (8) with $z(0) = x$.

We turn to prove Proposition 2.

Proof of Proposition 2. For any $\eta > 0$ we define recursively the following difference equation $r : [0, 1] \to X$:

$$r(0) = x \text{ and } r((k+1)\eta) - r(k\eta) = \eta(W - I)r(k) \tag{10}$$

where for any $t \neq k\delta$ we let $r(t) = r(\lfloor t/\eta \rfloor \eta)$. By a standard result in differential equations (see e.g., Hartman [12] Exercise 2.1), as η approaches zero r converges in the sup-norm topology to the solution of (9) z on any interval $[0, T]$. A simple induction shows that $r((k+1)\eta) = ((1-\eta)I + \eta W)^k x$ for any $k \geq 0$. We further note that $\lim_{k\to\infty}((1-\eta)I + \eta W)^k = \mathbf{1}(\mu_W)'$ which is a matrix with all rows equal $(\mu^W)'$. Therefore, $\lim_{t\to\infty} r(t) = \lim_{t\to\infty} z(t) = \mathbf{1}\langle \mu^W, x\rangle$. That is, for $z(0) = x$ the limit point of z is a consensus vector where all m coordinates equal $\langle \mu^W, x\rangle$.

It follows from the above and the compactness of X that for every $\eta > 0$ there exists a time T such that $|\Phi(x, t) - \langle \mu^W, x\rangle| < \eta$ for every $x \in X$. Therefore, Proposition 3 implies that for every M and η there exists ϵ_0 and a time T such that:

$$\mathbf{P}(\max_{T \leq t \leq M} |y(t) - \langle \mu^W, \mathbf{x}^1\rangle| > \eta) < \eta.$$

By definition of $y(t)$, this concludes the proof of Proposition 3.

□

7.2 Proof of Theorem 1

Similarly to the previous section we can write the (ϵ, δ)-*extremization DeGroot* as follows:

$$x_i^{t+1} - x_i^{t-1} - \epsilon U_i^{t+1} = \epsilon[(1-\delta)\sum_{j\in V} w_{ij}(x_j^t - x_i^t) + \delta(\varphi(x_i^t) - x_i^t)]. \tag{11}$$

where $U_i^{t+1} = (1-\delta)\sum_{j\in V} w_{ij}(a_j^t - x_j^t)$. Again note that since the expectation of a_j^t is x_j^t for any $j \in V$ we have that $E[U_i^{t+1}|\mathcal{F}_t] = 0$. Let $y(t) = x^{\lfloor \frac{t}{\epsilon} \rfloor}$. It now follows from Proposition 3 that the processes $y(t)$ can be approximated by the following differential equation:

$$\dot{z}^\delta = (1-\delta)Wz + \delta(\vec{\varphi}(z^\delta) - z^\delta),\ z^\delta(0) = \mathbf{x}^1 \tag{12}$$

where $\vec{\varphi}(z) = (\varphi(z_1), \ldots, \varphi(z_n))$. We start with the following lemma.

Lemma 3. For every $\eta > 0$ there exists $\delta > 0$ such that if $\langle \mu^W, \mathbf{x}^1 \rangle > p + \eta$ then the solution to (12) with initial condition $z^\delta(0) = \mathbf{x}^1$ satisfies $\lim_{t\to\infty} z^\delta(t) = \mathbf{1}$. Similarly, if $\langle \mu^W, \mathbf{x}^1 \rangle < p - \eta$, then the solution to (12) with initial condition $z^\delta(0) = \mathbf{x}^1$ satisfies $\lim_{t\to\infty} z^\delta(t) = \mathbf{0}$.

Proof. We show the first part of the lemma. The second part follows using similar considerations. We note first that the derivative $\dot{z}^\delta$ of equation (12) approaches to that of (9) as δ goes to zero. Therefore a standard application of Gronwall's inequality show that for any initial condition x, any a time T, and any $\beta > 0$ there exists a small enough $\delta > 0$ such that the two solutions z, z^δ satisfy $\sup_{0\le t\le T} \|z(t) - z^\delta(t)\| < \beta$ for any $\delta < \delta_0$.

We have shown in the proof of Proposition 2 that $\lim_{t\to\infty} z(t) = \mathbf{1}\langle \mu^W, x^0 \rangle$. Therefore, for any $\beta > 0$ and x^0 there exists a δ_0 such that for all $\delta < \delta_0$ there exists a time T such that the solution z^δ satisfy:

$$\|z^\delta(T) - \mathbf{1}\langle \mu^W, \mathbf{x}^1 \rangle\|_\infty < \beta.$$

From compactness of X, for any $\beta > 0$ we can find $\delta_0 > 0$ and a time T' such that for any $\delta < \delta_0$ and any $x \in X$:

$$\inf_{0\le t\le T'} \|z^\delta(t) - \mathbf{1}\langle \mu^W, \mathbf{x}^1 \rangle\|_\infty < \beta.$$

Hence, for any $\eta > 0$ we can find $\delta_0 > 0$ such that if $\langle \mu^W, \mathbf{x}^1 \rangle > p + \eta$, then $\|z^\delta(T) - \mathbf{1}\langle \mu^W, \mathbf{x}^1 \rangle\|_\infty < \frac{\eta}{2}$ for some $T \le T'$ and for any $\delta < \delta_0$. In particular, $z_i^\delta(T) > p + \frac{\eta}{2}$ for any $i \in V$.

In order to complete the proof we will next show that if $z_i^\delta(0) > p + \frac{\eta}{2}$ for every $i \in V$, then $\lim_{t\to\infty} z(t) = \mathbf{1}$. To see this note first that since $z_i^\delta(0) > p$ it holds that $\varphi(z_i^\delta) > z_i^\delta$ for every i. Recall that Let $z_{min}^\delta(t) = \min_{i\in V} z_i^\delta(t)$. Since

$$\dot{z}_i^\delta(t) = (1-\delta)\sum_{j\in V} w_{ij}(z_j^\delta(t) - z_i^\delta(t)) + \delta(\varphi(z^\delta) - z_i^\delta),$$

it follows that $\dot{z}_{min}^\delta > 0$. We note that since $(\varphi(x) - x)$ is bounded away from zero on every interval $[\alpha, \beta] \subseteq (p, 1)$. Therefore, it follows that $\lim_{t\to\infty} z_{min}^\delta(t) = 1$. Hence in particular $\lim_{t\to\infty} z^\delta(t) = \mathbf{1}$. This concludes the proof of the lemma. □

We next show the following.

Lemma 4. Let $\eta > 0$. Consider $\mathbf{x}^1 \in X$ such that $x_i^1 > 1 - \eta$ for every $i \in inV$ and let A_η be the event that $\lim_{t\to\infty} \mathbf{x}^t = \mathbf{1}$ and $x_i^t > p$ for every $i \in V$ and every t. Then the probability of A_η approaches one as η approaches zero. Similarly, if $\mathbf{x}^1 \in X$ satisfies $x_i^1 > 1 - \eta$ $i \in inV$ and B_η is the event that $\lim_{t\to\infty} \mathbf{x}^t = \mathbf{0}$ and $x_i^t < p$ for every $i \in V$ and every t, then the probability of B_η approaches one as η approaches zero.

Proof. We show the first part. We note that since $\mathbf{x}^t$ converges almost surely to either $\mathbf{0}$ or $\mathbf{1}$ and since $\langle \mu^W, x^t \rangle$ is a martingale we must have that the probability it converges to $\mathbf{1}$ is $\langle \mu^W, x^0 \rangle$. If $x_i^t \leq p$ for some i, then $\mathbf{x}^t$ converges to $\mathbf{0}$ with probability at least $\mu_*(1-p)$ where μ_* is the minimal coordinate in μ^W. Therefore the probability of A_η is at least $1 - \frac{\eta}{\mu_*(1-p)}$. □

Finally we have the following lemma.

Lemma 5. Consider the (ϵ, δ)-*extremization DeGroot* . For every $\eta > 0$ there exists $\delta > 0$ such that (i) if $\langle \mu^W, \mathbf{x}^1 \rangle > p + \eta$, then $\lim_{t\to\infty} \mathbf{x}^t = \mathbf{1}$ with probability at least $1 - \eta$ for all sufficiently small ϵ and analogously (ii) if $\langle \mu^W, \mathbf{x}^1 \rangle < p - \eta$ then $\lim_{t\to\infty} \mathbf{x}^t = \mathbf{0}$ with probability at least $1 - \eta$ for all sufficiently small ϵ.

Proof. Assume that $\langle \mu^W, \mathbf{x}^1 \rangle > p + \eta$. Lemma 3 implies that the solution to (12) with initial condition $z^\delta(0) = \mathbf{x}^1$ satisfies $\lim_{t\to\infty} z^\delta(t) = \mathbf{1}$ for all sufficiently small $\delta > 0$. Fix such a $\delta > 0$. Proposition 3 implies that for any $\gamma > 0$ there exists an $\epsilon_0 > 0$ such that for any $\epsilon < \epsilon_0$ there exists a $t > 0$ such that $\|\mathbf{x}^t - \mathbf{1}\| < \gamma$ with probability at least $1 - \frac{\eta}{2}$.

We further note that if $x \in X$ satisfies $x_i \geq p$, then

$$E_{(\epsilon,\delta)}[x_i^{t+1} | \mathbf{x}^t = x] \geq E_\epsilon[x_i^{t+1} | x^t = x],$$

where the right-hand side conditional expectation is taken with respect to the ϵ-DeGroot processes and the left-hand side with respect to the (ϵ, δ)-extremization DeGroot process. Hence, in particular, for any γ the event A_γ defined in Lemma 4 has a higher probability under the (ϵ, δ)-extremization DeGroot process comparing with the ϵ-DeGroot processess.

Therefore, in particular, for small enough γ if $\mathbf{x}^t = x$ such that $\|\mathbf{x}^t - \mathbf{1}\|_\infty < \gamma$, then $\lim_{t\to\infty} \mathbf{x}^t = \mathbf{1}$ with probability at least $1 - \frac{\eta}{2}$. Overall we have that if $\langle v, \mathbf{x}^1 \rangle > p + \eta$, then $\lim_{t\to\infty} \mathbf{x}^t = \mathbf{1}$ with probability at least $1 - \eta$. □

Consider a network (G, E, W) with μ^W as the centrality measure. Let F be the information structure and let l_F^G be the following number

$$l_F^G = p\mathbf{P}(\sum_{i\in V} \mu_i^W x_i^0 > p|\omega = 1) + (1-p)\mathbf{P}(\sum_{i\in V} \mu_i^W x_i^1 < p|\omega = 0).$$

For large network (G, E, W) with vanishing μ^W the term l_F^G approaches one.

Corollary 2. For any network (G, E, W) and $r > 0$ there exists δ and ϵ such that for the (ϵ, δ)-extremization process $l_F^G - r$ naive learning holds.

Proof. To see this note that for any $r > 0$ we can find η such that:

$$p\mathbf{P}(\sum_{i\in V} \mu_i^W x_i^1 > p + \eta|\omega = 1) + (1-p)\mathbf{P}(\sum_{i\in V} \mu_i^W x_i^1 < p - \eta|\omega = 0) > l_F^G - \frac{r}{2}.$$

The corollary now follows from Lemma 5. □

We next prove Theorem 1.

Proof of Theorem. The proof of Theorem 1 easily follows from the above corollary. To see this take a sequence of networks $\{(G_n, E_n, W_n)\}_n$ with a sequence vanishing centrality measures $\{\mu^n\}$. We note that $E[x_i^1|\omega = 1] > p$ and $E[x_i^1|\omega = 0] < p$ (see e.g., Acemoglu et al.). Furthermore the variables $\{x_i^1\}_{i\in V}$ are independent conditional on the realized state of the world ω. Therefore, we conclude, as in Golub and Jackson, that since the sequence $\{\mu^n\}$ is of vanishing measures it holds that

$$\lim_{n\to\infty} \mathbf{P}(\sum_{i\in V} \mu_i^n x_i^1 > p|\omega = 1) = 1 \text{ and } \lim_{n\to\infty} \mathbf{P}(\sum_{i\in V} \mu_i^W x_i^1 < p|\omega = 0).$$

Therefore by definition $\lim_{n\to\infty} l_F^{G_n} = 1$ and the theorem follows from Corollary 2. □

References

[1] I. Arieli, Y. Babichenko, S. Shlomov, (2018) "Robust non-Bayesian Social Learning", *working paper.*

[2] K. Arrow (1962), "Utilities, attitudes, choices: A review note", *Econometrica* 26, 1-23.

[3] A. Banerjee, E. Breza, A.G. Chandrasekhar, M. Mobius (2016), "Naive learning with uninformed agents", *working paper.*

[4] M. Benaim, (1998) "Recursive algorithms, urn processes and chaining number of chain recurrent sets," *Ergodic Theory and Dynamical Systems,* 18(1), 53-87.

[5] A.G Chandrasekhar, H. Larreguy, J.P Xandri (2016), "Testing models of social learning on networks: evidence from a lab experiment in the field", *working paper*.

[6] M. H. DeGroot (1974), "Reaching a Consensus", *Journal of the American Statistical Association* 69, 118-121.

[7] P. DeMarzo, D. Vayanos, J. Zwiebel (2003), "Persuasion Bias, Social Influence and Uni-Dimensional Opinions", *Quarterly Journal of Economics* 118, 909-968.

[8] D. Gale, S. Kariv (2003), "Bayesian learning in social networks", *Games and Economic Behavior*, 45, 329-346.

[9] B. Golub, M. Jackson (2010), "Naive Learning in Social Networks and the Wisdom of Crowds", *American Economic Journal: Microeconomics* 2:1, 112-149.

[10] B. Golub, E. Sadler (2016), "Learning in Social Networks", *The Oxford Handbook of the Economics of Networks.*

[11] Eyster, E., and Rabin, M. (2010), "Naïve Herding in Rich-Information Settings", *American Economic Journal: Microeconomics*, 2, 221-243.

[12] Hartman, P. (2002). "Ordinary Differential Equations," Classics in Applied Mathematics, Vol. 38. Society for Industrial and Applied Mathematics (SIAM).

[13] J. Hazla, A. Jadbabaie, E. Mossel, M.A. Rahimian (2018), "Reasoning in Bayesian Opinion Exchange Networks is PSPACE-Hard", *working paper.*

[14] Levy G., R. Razin (2018). Information diffusion in networks with the Bayesian Peer Influence heuristic, *Games and Economic Behavior*, 109, 262-270.

[15] P. Molavi, A. Tahbaz-Salehi, A. Jadbabaie (2017), "A Theory of Non-Bayesian Social Learning", forthcoming in *Econometrica.*

[16] E. Mossel, A. Sly, O. Tamuz (2015), "Strategic learning and the topology of social networks", *Econometrica* 83, 1755-1794.

[17] E. Mossel, M. Mueller-Frank, A. Sly, O. Tamuz (2017), "Social Learning Equilibria", *working paper.*

[18] M. Mueller-Frank (2013), "A General Framework for Rational Learning in Social Networks", *Theoretical Economics* 8, 1-40.

[19] M. Mueller-Frank, C. Neri (2016), "Quasi-Bayesian Updating in Social Networks", *working paper*, IESE Business School..

[20] R. Gregory and W.H. Sandholm (2013), "Stochastic approximations with constant step size and differential inclusions." *SIAM Journal on Control and Optimization,* 51, 525-555.

[21] Rosenberg, D., Solan, E., and Vieille, N. (2009). Information externalities and emergence of consensus,' *Games and Economic Behavior*, 66, 979-994.

[22] N. Vulkan (2000) "An economist's perspective on probability matching", *Journal of Economic Surveys* 14, 101-118.