

REACHING CONSENSUS IN SOCIAL NETWORKS

Manuel Mueller-Frank

The Public-Private Sector Research Center is a Research Center based at IESE Business School. Its mission is to develop research that analyzes the relationships between the private and public sectors primarily in the following areas: regulation and competition, innovation, regional economy and industrial politics and health economics.

Research results are disseminated through publications, conferences and colloquia. These activities are aimed to foster cooperation between the private sector and public administrations, as well as the exchange of ideas and initiatives.

The sponsors of the Public-Private Sector Research Center are the following:

- Ajuntament de Barcelona
- Departament d' Economia i Coneixement de la Generalitat de Catalunya
- Departament d' Empresa i Ocupació de la Generalitat de Catalunya
- Diputació de Barcelona
- EVERIS
- Fundació AGBAR
- Institut Català de les Empreses Culturals (ICEC)
- PricewaterhouseCoopers
- Sanofi

The contents of this publication reflect the conclusions and findings of the individual authors and not the opinions of the Center's sponsors.

Reaching Consensus in Social Networks

Manuel Mueller-Frank[‡]
IESE Business School

November 14, 2013

Abstract

This paper considers network based non-Bayesian opinion formation on a linearly ordered set of opinions. The general class of constricting and continuous Markov revision functions, that contains the standard weighted average revision functions, is analyzed. A revision function is constricting if the revised opinion is strictly higher (lower) ranked than the lowest (highest) ranked observed opinion. The main advantages of the general approach are that (i) it captures a wide range of applications, and (ii) the constricting property is easily testable.

It is shown that asymptotic consensus occurs in strongly connected networks whenever the revision functions of all agents are constricting and continuous. The revision function does not need to be the same across agents, or across time for a given agent. Additionally, asymptotic consensus is shown to hold almost surely if agents are subject to a natural class of probabilistic mistakes when forming their opinions.

Keywords: Networks, social learning, consensus, non-Bayesian learning, boundedly rational learning, convergence, partisans, cognitive dissonance.

JEL: D83, D85.

*I thank Itai Arieli, Eddie Dekel, Margaret Meyer, Stephen Morris, Alessandro Pavan, Marciano Siniscalchi, Dimitri Vayanos, Joergen Weibull and Peyton Young for comments and suggestions. I am grateful to seminar audiences at the University of Oxford, the conference of the Society for the Advancement of Economic Theory, and the Transatlantic Economic Theory workshop (PSE). This paper incorporates and extends upon some results appearing in an earlier working paper, Mueller-Frank [25].

[‡]E-mail: mmuellerfrank@iese.edu

1 Introduction

Social networks play an important role as communication platforms. Individuals interact with their social peers on an ongoing basis and use the information gained through the interaction when forming opinions or making decisions. In such interactive environments individuals adjust their opinions and actions in response to those of their social peers. This tendency has been documented empirically, as for example in Rogers and Shoemaker [28], and experimentally in Choi, Gale and Kariv [6] and Lorenz, Rauhut, Schweitzer and Helbing [16]. In the theoretical literature opinion processes under repeated interaction in social networks have received substantial attention recently.¹ The theoretical literature focuses on observational learning, where the state of the world is fixed over time and agents observe private information only once, prior to the onset of interaction.² A central question then is how learning of the agents should be modeled.

The literature takes two different approaches. In the Bayesian approach, all agents are assumed to be Bayesian, that is they update their opinions by making fully rational inferences regarding the private information of all agents based on the opinions they observe. Such models serve as a benchmark but fail to represent opinion formation in a realistic manner as the required inferences can be exceedingly complex. The non-Bayesian approach is founded on the assumption that agents are unwilling or unable to undertake complex inferences but instead use simple learning heuristics when updating their opinions. The standard model of non-Bayesian updating in social networks goes back to DeGroot [7]: the opinion of an agent in a given period is a weighted average of the last period opinions the agent observed. The strength and reason for the popularity of the DeGroot model is its tractability. However, to the best of my knowledge there are no empirical or experimental studies providing conclusive evidence for weighted average revision functions.³

The main objective of this paper is to provide an alternative model of non-Bayesian opinion formation in a general environment. Rather than focusing on a particular functional form I consider a general class of revision functions whose defining functional property is easily testable and, in fact, is satisfied in a recent experimental study of opinion formation by Lorenz, Rauhut, Schweitzer and Helbing [16]. The main question of the paper concerns

¹See for example DeMarzo, Vayanos and Zwiebel [8], Gale and Kariv [11], Golub and Jackson [12], Jadbabaie, Molavi, Sandroni and Tahbaz-Salehi [14], Mossel, Sly and Tamuz [21], and Mueller-Frank [22, 23].

²See for example DeMarzo, Vayanos and Zwiebel [8], Gale and Kariv [11], Golub and Jackson [12], and Mueller-Frank [22, 23].

³See section 2 for details.

the long run properties of aggregate behavior. That is, under which conditions do all agents in a network converge to the same opinion?

This question is addressed in the following model. A finite set of agents are organized in a social network and engage in repeated interaction. In each of countable rounds of interaction every agent announces his opinion to all his neighbors in the network. The set of opinions is assumed to be linearly ordered. The first round opinion of each agent might arise from rational inference on private information while subsequent opinion revisions are assumed to be based upon a simple learning heuristic; the opinion of an agent in round t is a function of the last period opinions he observed. The revision functions therefore satisfy a Markov property as the revision depends only on the last period observed opinions. I consider the class of continuous and constricting Markov revision functions, that includes the set of weighted average revision functions as a subset. A revision function is constricting if the revised opinion is strictly higher ranked than the lowest ranked observed opinion and strictly lower ranked than the highest ranked observed opinion. The model intends to capture only observational learning. Beyond the opinions of other agents there are no further sources of information that might affect one's opinion.

The paper presents two formal results. Theorem 1 establishes that if the social network is strongly connected, the revision functions of all agents are continuous and constricting, and all agents satisfy finite types, then asymptotic consensus holds.⁴ The individual property of constricting updating hence leads to consensus in the aggregate in any strongly connected network. Despite the fact that all agents communicate only locally, their opinions converge in the long run. Moreover, asymptotic consensus does not require agents to be homogeneous and it is robust against time inconsistent behavior as long as the set of revision functions an agent applies over time is finite.⁵ In other words, Theorem 1 establishes that any kind of updating behavior that has a non-vanishing attraction to the observed opinions leads to eventual consensus in strongly connected networks. In an environment with a rich set of opinions long run disagreement requires some agents to stop paying attention to some of their neighbors opinions.⁶

For the second result, I restrict attention to real numbered opinions and consider the

⁴An agent satisfies finite types if the number of different revision functions the agent uses over time is finite.

⁵DeMarzo, Vayanos and Zwiebel [8] consider weighted average revision functions and allow agents to change their revision function over time. However, they require the relative weights assigned to neighbors to remain constant over time.

⁶The set of opinions is assumed to be a linear continuum and hence satisfies richness.

robustness of asymptotic consensus when agents are subject to probabilistic mistakes when forming their opinions. Probabilistic mistakes are introduced in the following way. The opinion of an agent in a given period consists of the sum of his revised opinion, based upon a Markov revision function, and a random error term. Theorem 2 characterizes a condition on the stochastic process of mistakes such that asymptotic consensus occurs with probability one in strongly connected networks, if the underlying revision functions of all agents are constricting and continuous and all agents satisfy finite types.

The existing economics literature on non-Bayesian learning mainly relies on the DeGroot [7] model.⁷ Instead, this paper considers a general class of Markov revision functions, that includes the set of weighted average revision functions as a subset, and further generalizes the DeGroot model by allowing the set of opinions to be linearly ordered as opposed to equal the set of real numbers. The main benefits of generalizing the class of revision functions are the following. First, as already mentioned, in contrast to weighted average revision functions, the constricting property is easily testable and there is some experimental evidence for it. The data generated in a recent experiment by Lorenz, Rauhut, Schweitzer and Helbing [16] indicates that individuals indeed satisfy the constricting property when revising their opinion.⁸ The second benefit is that the use of weighted averages for specific applications is limited while the general class of revision functions captures types of opinion formation that are of interest to applied theorists. In section 2, I provide an example of constricting and continuous revision functions that might serve to model opinion formation subject to cognitive dissonance. The third benefit of the more general approach is that weighted average revision functions impose restrictions on the set of opinions. For example, there exists no natural analog of a weighted average function in the case of linearly ordered opinion spaces. The constricting property on the other hand naturally extends from real numbered opinions to linearly ordered opinions. Moreover, the constricting property can be applied to opinion formation on topological vector spaces and the consensus result carries forward.⁹

Beyond the added value of mathematical generalization, the main benefit of considering a linear continuum as opinion space is that opinions are no longer required to be real numbered but instead can be of a more abstract form.¹⁰

⁷See for example DeMarzo, Vayanos and Zwiebel [8], Golub and Jackson [12], and Mueller-Frank [23]. There also exists a large literature on non-Bayesian learning in computer science and electrical engineering. See section 5 for a discussion.

⁸See section 2 for details.

⁹See the conclusion for a brief discussion.

¹⁰As long as the opinion space is endowed with a natural order and satisfies the additional properties of a linear continuum.

The rest of the paper is organized as follows. In section 2 I present the model, discuss the benefits of considering the general class of functions and review recent experimental results that provide evidence for the constricting property. Section 3 provides Theorem 1 and its proof. Section 4 considers the extension of probabilistic mistakes and provides Theorem 2. Section 5 discusses the main innovations in context of the existing literature. Section 6 concludes. The proof of Theorem 2 and the proofs of all lemmas are presented in appendix A. In appendix B I provide an example of a constricting and continuous optimal action dynamic, and counter-examples for failure of consensus when relaxing finite types and continuity.

2 The Model

A finite set of agents $V = \{1, \dots, v\}$ is organized in a social network represented by a directed graph, $G = (V, E)$. The nodes in the graph correspond to the agents and the edges, $E \subset V^2$, correspond to the connections between agents. Let N_i denote the neighborhood of agent i ,

$$N_i = \{j \in V : ij \in E\}.$$

A directed path between nodes k_1 and k_l is a sequence of nodes k_1, \dots, k_l such that $k_{f+1} \in N_{k_f}$ for $f = 1, \dots, l-1$. A graph G is strongly connected if there exists a directed path from every node $i \in V$ to every other node $j \in V$.

The focus of the paper lies on the effect of repeated interaction in social networks on the process of opinions chosen by individuals. Interaction takes place in countable rounds. In the first round each agent i announces an opinion $\mathbf{a}_i^1 \in A$. The set of opinions A is assumed to be uncountable and endowed with a strict total order such that (i) A has the least upper bound property, and (ii) for all $a' \prec a'''$ there exists a'' such that

$$a' \prec a'' \prec a'''.$$

The two conditions on A make it a linear continuum. The order topology on A is denoted by τ_A and the corresponding product topology on $A^v = \mathbf{A}$ by $\tau_{\mathbf{A}}$. From the second round onward the opinion \mathbf{a}_i^t of agent i in a given round is a function $f_i^t(\mathbf{a}^{t-1})$ of the opinions of her neighbors and herself in the previous round,

$$f_i^t : \mathbf{A} \rightarrow A.$$

The revision functions are assumed to be local, i.e. if $\mathbf{a}_j = \mathbf{a}'_j$ for all $j \in N_i \cup \{i\}$, then $f_i^t(\mathbf{a}) = f_i^t(\mathbf{a}')$. Consider a mapping $\mathbf{f}^t : \mathbf{A} \rightarrow \mathbf{A}$ such that

$$\mathbf{f}^t(\mathbf{a}) = (f_1^t(\mathbf{a}), \dots, f_v^t(\mathbf{a}))$$

where the revision functions f_i^t are the components of \mathbf{f}^t . The sequence of opinions can be defined recursively in the following manner. For all $t = 2, 3, \dots$

$$\mathbf{a}^t = \mathbf{f}^t(\mathbf{a}^{t-1}) = (f_1^t(\mathbf{a}^{t-1}), \dots, f_v^t(\mathbf{a}^{t-1})).$$

I make three assumptions on the revision functions of agents. While I do not impose a time consistency requirement, in the sense of agents applying the same revision function f_i in all periods, I allow for only finitely many revision functions for each agent. Each agent has a finite type, i.e. for each agent i there exists a finite set of revision functions $\{f_i^\delta\}_{\delta \in \Delta}$ such that $f_i^t \in \{f_i^\delta\}_{\delta \in \Delta}$ for all $t \in \mathbb{N}$. The second assumption is continuity of the revision function f_i^t for all agents i and time periods t . The third assumption is that the revision functions f_i^t are constricting for all agents i and rounds t .

Definition 1 $f_i : \mathbf{A} \rightarrow A$ is *constricting* if $\mathbf{a}_j \neq \mathbf{a}_i$ for some $j \in N_i$ implies

$$\inf_{j \in N_i \cup \{i\}} \mathbf{a}_j \prec f_i(\mathbf{a}) \prec \sup_{j \in N_i \cup \{i\}} \mathbf{a}_j.$$

Intuitively, the constricting property seems reasonable. Agents position their revised opinion in the strict interior of the two most extreme opinions they observed in the last round. Since weighted averages, the type of revision function typically analyzed in the literature, are continuous and constricting they are a subclass of the class of revision functions considered here. The reader might wonder why it is worthwhile to consider the more general function class. The main reason for the wide adoption of weighted average revision functions in the literature is their tractability. By looking at a more general class of functions one loses tractability. The question is whether the gains from generalization offset its cost.

When departing from the full rationality assumption one should always be concerned with the verifiability of the assumptions one imposes on behavior in order to avoid arbitrariness. To the best of my knowledge there are no experimental or empirical studies that validate weighted average revision functions. More general properties of updating behavior might be easier to verify in practice. Lorenz, Rauhut, Schweitzer and Helbing [16] provide an

experimental analysis of whether observing the opinion of others can undermine the wisdom of crowd effect. In their study subjects were asked questions with a numerical answer, for example, what is the length of the border between Italy and Switzerland. They consider several different treatments but the one of interest to this paper is the full information treatment in which subjects observe the first round answers of 11 other subjects and are asked to revise their opinion. While the objective of Lorenz et al. was not to identify properties of individual updating behavior, their data is perfectly suited to test the constricting property. In fact, in 281 out of 288 observations the constricting property on individual opinion revision is satisfied. When considering only the revised opinion of those agents who gave the lowest or largest answer in the first period, there are 48 observations and in 45, or 93.75 percent, of them the constricting property is satisfied. Therefore, their data indicates that the constricting property might be a good approximation of opinion formation in certain environments.

The second benefit associated with the more general class of revision functions is that weighted averages are limited in terms of their applications. In particular, under a weighted average revision function the weights assigned to the opinions of neighbors are fixed and cannot vary with the opinion of the agents. The following example provides a type of continuous and constricting revision functions which might be interesting from an applied perspective and where weights depend on opinions. Assume a real numbered set of opinions, i.e. $A = \mathbb{R}$ and consider the following revision function

$$f_i(\mathbf{a}) = \sum_{j \in N_i \cup \{i\}} \mathbf{a}_j \times w_{ij}(\mathbf{a}).$$

where $w_{ij} : \mathbf{A} \rightarrow \mathbb{R}$ is defined as follows: let $g : \mathbb{R}_+ \rightarrow \mathbb{R}_{++}$ be a continuous, weakly decreasing function and

$$w_{ij}(\mathbf{a}) = \frac{g(|\mathbf{a}_i - \mathbf{a}_j|)}{\left(\sum_{k \in N_i} g(|\mathbf{a}_i - \mathbf{a}_k|) \right) + W}$$

$$w_{ii}(\mathbf{a}) = \frac{W}{\left(\sum_{k \in N_i} g(|\mathbf{a}_i - \mathbf{a}_k|) \right) + W}$$

where $W \in \mathbb{R}_{++}$. The weight an agent assigns to his neighbors opinion is then decreasing in the distance of their opinions. This revision function might be used to model cognitive dissonance which refers to the conflict or uncomfortable feeling caused by inconsistent attitudes and actions. One way of capturing the conflict resulting from inconsistency is by

applying it to intertemporal opinion change of a given agent. When forming his opinion the agent considers the last period announcements of his neighbors and himself. When faced with opinions of his neighbors that widely differ from his own opinion expressed in the last period, the agent might feel discomfort when weighing the distant opinions highly as it leads to a large shift in his opinion from one period to the next, which he might perceive as having had a faulty opinion in the last period. When on the contrary the observed opinions of his neighbors are close to his own, the agent might feel acknowledged and can assign higher weight to his neighbors opinions as it only leads to a small inter-temporal opinion change. Widely differing opinions are incorporated, but in a marginal way.

Finally, weighted averages are only applicable to a limited set of opinion spaces. For example, there is no natural analog of weighted average revision functions in more general linearly ordered spaces. The constricting property on the other hand, generalizes in a direct manner from real numbered opinions to linearly ordered opinions and (a little less directly) to topological vector spaces.

Some readers might be concerned with the somewhat mechanical nature of opinion formation that is considered here. Generally, it is the standard norm in the non-Bayesian opinion formation literature to consider revision functions that satisfy the Markov property and hence the mechanical nature. Moreover, the opinion formation process can be thought of as a learning dynamic in games and there are classes of games where standard learning dynamics satisfy the constricting property. Hence the learning dynamics of certain types of games fall into the general class of updating dynamics considered in this paper. Please see the appendix B for an example of a constricting and continuous learning dynamic.

3 The Robustness of Asymptotic Consensus

The main objective of the paper is to analyze the long run implications of non-Bayesian opinion formation that satisfies the constricting property. There exists a large body of theoretical literature that establishes the human tendency towards uniform behavior and opinions. The question is whether this tendency towards asymptotic consensus carries forward to the general non-Bayesian opinion formation model considered here. Theorem 1 answers this question affirmatively.

Theorem 1 *Let the network G be strongly connected. If the revision functions of all agents are continuous and constricting and all agents satisfy finite types, then for every initial opinion vector $\mathbf{a}^1 \in \mathbf{A}$ there exists an opinion $L(\mathbf{a}^1) \in A$ such that the opinions of each agent converge to $L(\mathbf{a}^1)$.*

The theorem states that asymptotic consensus is a robust outcome under the general class of constricting and continuous revision functions. If all agents in any strongly connected networks satisfy the constricting property and their revision functions are continuous, then all agents reach consensus asymptotically. Moreover, the result does not depend on time consistent behavior. Agents do not need to apply the same revision function in each communication round but can switch their revision function infinitely often over time, as long as the set of functions they are choosing from is finite. Hence asymptotic consensus is a quite robust outcome and does not rely on the assumption of weighted average revision functions made so far in the literature.

Prior to presenting the proof and the lemmas involved in it, let me provide some insight on what could prevent asymptotic consensus. For simplicity consider the special case in which the set of opinions is \mathbb{R} . It is quite intuitive that continuous and constricting revision functions imply that when considering only the largest and smallest opinion expressed in the network in a given round the corresponding sequences are weakly monotone and bounded. Hence the sequence of smallest and largest opinions converge. But they need not converge to the same limit. So one viable approach to proving the theorem would be to show that the limit of largest and smallest announcements is indeed the same. However, the proof presented here has a different structure. The reason for choosing the presented structure is that its broad idea can be maintained to prove both theorems.¹¹

Let $\mathbf{C} \subset \mathbf{A}$ be the set of all consensus opinion vectors,

$$\mathbf{C} = \{\mathbf{a} \in \mathbf{A} : \mathbf{a}_i = \mathbf{a}_j \text{ for all } i, j \in V\}.$$

The broad structure of the proof is to show that i) the sequence of opinion vectors $\{\mathbf{a}^t\}_{t \in \mathbb{N}}$ has a convergent subsequence, ii) this limit of the subsequence is a consensus vector, and iii) the sequence $\{\mathbf{a}^t\}_{t \in \mathbb{N}}$ converges to the limit of the converging subsequence and hence a consensus vector. The proof of Theorem 1 relies on the following five lemmas.

¹¹That is Theorem 1 and Theorem 2 in section 4. Moreover, the proof structure can be maintained when A is a topological vector space.

Lemma 1 *If the network is strongly connected and the revision functions f_i of all agents are continuous and constricting, then*

$$\mathbf{f}(\mathbf{a}) = \mathbf{a} \Leftrightarrow \mathbf{a} \in \mathbf{C}.$$

Lemma 1 characterizes the set of fixed points of the mapping \mathbf{f} as the set of consensus points. For $\mathbf{a} \in \mathbf{A}$ denote its supremum and infimum by $\mathbf{a}_+ = \sup_{i \in V} \mathbf{a}_i$ and $\mathbf{a}_- = \inf_{i \in V} \mathbf{a}_i$ respectively.¹²

Lemma 2 *If f_i^t is constricting and continuous for all $i \in V$ and $t \in \mathbb{N}$, then*

$$\mathbf{a}^{t+n} \in \mathbf{A}^t = \{\mathbf{a} \in \mathbf{A} : \mathbf{a}_-^t \preceq \mathbf{a}_i \preceq \mathbf{a}_+^t \ \forall i \in V\}$$

for all $n \in \mathbb{N}$.

Lemma 3 *The set $\mathbf{A}^1 \subset \mathbf{A}$ is compact, where*

$$\mathbf{A}^1 = \{\mathbf{a} \in \mathbf{A} : \mathbf{a}_-^1 \preceq \mathbf{a}_i \preceq \mathbf{a}_+^1 \ \forall i \in V\}.$$

Lemma 4 *Let G be strongly connected. If f_i^t is constricting and continuous for all $i \in V$ and $t \in \mathbb{N}$, then there exists a finite $k \in \mathbb{N}$ such that for all $\mathbf{a}^1 \in \mathbf{A} \setminus \mathbf{C}$ the corresponding sequence of opinions $\{\mathbf{a}^t\}_{t \in \mathbb{N}}$ satisfies $\mathbf{a}_-^t \succ \mathbf{a}_-^1$ and $\mathbf{a}_+^t \prec \mathbf{a}_+^1$ for all $t \geq k$.*

Lemma 5 *Let $\{\mathbf{a}^{t_\lambda}\}_{t_\lambda \in \Lambda}$ be a converging subsequence of $\{\mathbf{a}^t\}_{t \in \mathbb{N}}$ and suppose that $\{\mathbf{a}^{t_\lambda}\}_{t_\lambda \in \Lambda}$ converges to $\mathbf{a}^* \in \mathbf{C}$. Then $\{\mathbf{a}^t\}_{t \in \mathbb{N}}$ converges to \mathbf{a}^* .*

See below for the proof of Theorem 1.

Proof. The result is established in three steps.

1. The sequence $\{\mathbf{a}^t\}_{t \in \mathbb{N}}$ lies in \mathbf{A}^1 by Lemma 2. \mathbf{A}^1 is compact by Lemma 3. Hence $\{\mathbf{a}^t\}_{t \in \mathbb{N}}$ has a converging subsequence $\{\mathbf{a}^{t_\lambda}\}_{t_\lambda \in \Lambda}$ with limit \mathbf{a}^* , where $\Lambda \subset \mathbb{N}$ and $|\Lambda| = |\mathbb{N}|$

¹²Lemma 2 is an immediate consequence of Lemma 1 and the constricting property. Its proof is hence omitted.

2. Suppose that the limit of the converging subsequence $\{\mathbf{a}^{t_\lambda}\}_{t_\lambda \in \Lambda}$ is a consensus point $\mathbf{a}^* \in \mathbf{C}$. Then $\{\mathbf{a}^t\}$ converges to \mathbf{a}^* by Lemma 5.
3. Suppose on the contrary that $\mathbf{a}^* \notin \mathbf{C}$. For a given sequence of revision functions $\{f_i^t\}_{t \in \mathbb{N}}$ of agent i , denote by F_i the set of revision functions agent i uses

$$F_i = \{f : f = f_i^t \text{ for some } t\}.$$

Consider the following set of mappings $\mathbf{f} : \mathbf{A} \rightarrow \mathbf{A}$

$$\mathcal{F} = \{\mathbf{f} = (f_1, \dots, f_v) : f_i \in F_i \text{ for all } i \in V\}.$$

Finally, denote by \mathcal{F}^k the following set of composite mappings

$$\mathcal{F}^k = \{\mathbf{f}^1 \circ \mathbf{f}^2 \circ \dots \circ \mathbf{f}^k : \mathbf{f}^i \in \mathcal{F}, i = 1, \dots, k\}.$$

\mathcal{F}^k is the set of all mappings that are equal to a total of k compositions of mappings in \mathcal{F} . Note that all mappings in \mathcal{F}^k are continuous due to continuity of the revision functions of all agents. Furthermore, as all agents satisfy finite types, the set \mathcal{F} is finite which implies finiteness of \mathcal{F}^k . Denote by $\mathbf{F}^{t+1, t+k}$ the composition of mappings $\{\mathbf{f}^n\}_{n=t+1}^{t+k}$,

$$\mathbf{F}^{t+1, t+k} = \mathbf{f}^{t+k} \circ \mathbf{f}^{t+k-1} \circ \dots \circ \mathbf{f}^{t+1}.$$

We have $\mathbf{F}^{t+1, t+k} \in \mathcal{F}^k$ for all $t \in \mathbb{N}$. Consider the converging subsequence $\{\mathbf{a}^{t_\lambda}\}_{t_\lambda \in \Lambda}$ and the corresponding sequence of composite mappings $\{\mathbf{F}^{t_\lambda+1, t_\lambda+k}\}_{t_\lambda \in \Lambda}$. Note that

$$\{\mathbf{F}^{t_\lambda+1, t_\lambda+k}(\mathbf{a}^{t_\lambda})\}_{t_\lambda \in \Lambda} = \{\mathbf{a}^{t_\lambda+k}\}_{t_\lambda \in \Lambda}$$

is a subsequence of $\{\mathbf{a}^t\}_{t \in \mathbb{N}}$. As $\mathbf{F}^{t+1, t+k} \in \mathcal{F}^k$ for all $t \in \mathbb{N}$ and \mathcal{F}^k is finite, there exists a $\mathbf{F} \in \mathcal{F}^k$ and an infinite subsequence $\{\mathbf{a}^{t_\gamma}\}_{t_\gamma \in \Gamma}$ of $\{\mathbf{a}^{t_\lambda}\}_{t_\lambda \in \Lambda}$ such that $\mathbf{F}^{t_\gamma+1, t_\gamma+k} = \mathbf{F}$ for all $t_\gamma \in \Gamma$. Next suppose that $\{\mathbf{a}^{t_\lambda}\}_{t_\lambda \in \Lambda}$ converges to a point $\mathbf{a}^* \notin \mathbf{C}$ which directly implies that the sequence $\{\mathbf{a}^{t_\gamma}\}_{t_\gamma \in \Gamma}$ converges to $\mathbf{a}^* \notin \mathbf{C}$. Convergence of $\{\mathbf{a}^{t_\gamma}\}_{t_\gamma \in \Gamma}$ to \mathbf{a}^* and continuity of \mathbf{F} implies the convergence of $\{\mathbf{a}^{t_\gamma+k}\}_{t_\gamma \in \Gamma}$ to $\mathbf{F}(\mathbf{a}^*) = \mathbf{a}^{**}$. By Lemma 4 we have $\mathbf{a}_-^* \prec \mathbf{a}_-^{**}$ and $\mathbf{a}_+^{**} \prec \mathbf{a}_+^*$. As A is a linear continuum there exists a_- such that $\mathbf{a}_-^* \prec a_- \prec \mathbf{a}_-^{**}$ and a_+ such that $\mathbf{a}_+^{**} \prec a_+ \prec \mathbf{a}_+^*$. Consider the following open set in the product topology $\tau_{\mathbf{A}}$

$$\mathbf{A}_{\succ a_-} = \{\mathbf{a} \in \mathbf{A} : \mathbf{a}_i \succ a_- \text{ for all } i\}.$$

As $\mathbf{A}_{\succ a'}$ is an open set containing \mathbf{a}^{**} and $\{\mathbf{a}^{t_\gamma+k}\}_{t_\gamma \in \Gamma}$ converges to \mathbf{a}^{**} there exists a round t'_γ such that for all $t_\gamma > t'_\gamma$ we have $\mathbf{a}^{t_\gamma+k} \in \mathbf{A}_{\succ a'}$. Consider an element of the sequence $\{\mathbf{a}^{t_\gamma+k}\}_{t_\gamma \in \Gamma}$ that lies in $\mathbf{A}_{\succ a'}$ and the corresponding set $\mathbf{A}^{t_\gamma+k}$. Note that such $\mathbf{A}^{t_\gamma+k}$ is a subset of $\mathbf{A}_{\succ a'}$. By Lemma 2 for all $t > t_\gamma + k$ the sequence $\{\mathbf{a}^t\}_{t \in \mathbb{N}}$ lies in $\mathbf{A}^{t_\gamma+k}$. Let i^- denote the agent that satisfies $\mathbf{a}_{i^-}^* = \mathbf{a}_-^*$. Consider the following open set $\mathbf{V}_{\mathbf{a}^*}$ containing \mathbf{a}^* ,

$$\mathbf{V}_{\mathbf{a}^*} = \{\mathbf{a} \in \mathbf{A} : \mathbf{a}_{i^-} \prec a_-, \mathbf{a}_j \prec \mathbf{a}_+^1 \text{ for } j \neq i^-\}.$$

As $\{\mathbf{a}^{t_\gamma}\}_{t_\gamma \in \Gamma}$ converges to \mathbf{a}^* and $\mathbf{V}_{\mathbf{a}^*}$ is an open set containing \mathbf{a}^* there exists a round t' such that for all $t_\gamma > t'$ all elements of the subsequence $\{\mathbf{a}^{t_\gamma}\}_{t_\gamma \in \Gamma}$ lie in $\mathbf{V}_{\mathbf{a}^*}$. The fact that $\mathbf{A}_{\succ a_-}$ and $\mathbf{V}_{\mathbf{a}^*}$ have an empty intersection leads to a contradiction as $\{\mathbf{a}^{t_\gamma}\}_{t_\gamma \in \Gamma}$ is a subsequence of $\{\mathbf{a}^t\}_{t \in \mathbb{N}}$. Therefore the limit of the converging subsequence has to be a consensus point. ■

Theorem 1 establishes sufficient conditions for asymptotic consensus. No mention is made of necessary conditions. But what can be said about necessary conditions for asymptotic consensus? Let me address the question in the following way: can asymptotic consensus be achieved by relaxing, for all agents, one of the three conditions on the revision functions while keeping the other two conditions as is? By providing a counter-example for each of the three cases I establish a certain tightness of the sufficient conditions. All counter-examples consider the special case in which the set of opinions equals \mathbb{R} .

A natural way to relax the constricting requirement is to consider weakly constricting revision functions for all agents, where the announcement has to be at least as large as the smallest last period announcement and at most as large as the largest last period announcement. One immediate counter-example to convergence then is given by revision functions where each agent announces his last period opinion. Continuity and finite types are satisfied but asymptotic consensus fails for all but consensus points. For examples establishing failure of consensus when relaxing continuity and finite types see appendix B.

4 Asymptotic Consensus under Probabilistic Mistakes

Theorem 1 has shown asymptotic consensus to be a robust outcome of non-Bayesian opinion formation on linearly ordered opinion spaces. The objective of this section is to further

test the robustness of consensus through the introduction of probabilistic mistakes. Suppose that the revision functions of agents are continuous and constricting, and agents satisfy finite types but are subject to probabilistic mistakes when forming their opinion. For this extension suppose that $A = \mathbb{R}$. Let \mathbf{X}_i^t be a real valued random variable. The opinion of agent i in period t is given by

$$\mathbf{a}_i^t = f_i^t(\mathbf{a}^{t-1}) + \mathbf{X}_i^t$$

which leads to the following vector of opinions in period t

$$\mathbf{a}^t = \mathbf{f}^t(\mathbf{a}^{t-1}) + \mathbf{X}^t$$

where \mathbf{X}^t is a random vector composed of the individual error terms. I make the following assumption on the stochastic process $\{\mathbf{X}_i^t\}_{t \in \mathbb{N}}$,

$$\Pr \left(\sum_{t \in \mathbb{N}} |\mathbf{X}_i^t| \text{ finite} \right) = 1$$

for all $i \in V$. The probability measure $\Pr(\cdot)$ is defined on the infinite product space $(\Omega_i^\infty, \mathcal{F}_i^\infty)$ of the underlying probability spaces $(\Omega_i^t, \mathcal{F}_i^t, P_i^t)$ of the random variables \mathbf{X}_i^t .¹³ The condition implies that $\{\mathbf{X}_i^t\}_t$ converges almost surely to a constant random variable that takes the value zero in all states. One example for a stochastic process satisfying the condition is the case where \mathbf{X}_i^t is uniformly distributed over the interval $[-\delta^t, \delta^t]$.

The following theorem establishes the robustness of asymptotic consensus in strongly connected networks to this type of probabilistic mistakes.

Theorem 2 *Let the network G be strongly connected. If the revision functions of all agents are constricting and continuous, all agents satisfy finite types and the error terms satisfy*

$$\Pr \left(\sum_{t \in \mathbb{N}} |\mathbf{X}_i^t| \text{ finite} \right) = 1$$

for all i , then asymptotic consensus holds with probability one for any initial opinion $\mathbf{a}^1 \in \mathbb{R}^v$.

The almost surely result and the corresponding probability measure relates to the product

¹³The infinite product measure theorem states conditions on the probability spaces $(\Omega_i^t, \mathcal{F}_i^t, P_i^t)$ corresponding to \mathbf{X}_i^t such that a probability measure on the infinite product space $(\Omega_i^\infty, \mathcal{F}_i^\infty)$ exists. I assume that these conditions are satisfied.

space (across agents) of the infinite product probability spaces $(\Omega_i^\infty, \mathcal{F}_i^\infty, P_i^\infty)$ of agents. Theorem 2 establishes that asymptotic consensus is robust against probabilistic mistakes under certain conditions. Note that independence of the error terms across time or agents is not necessary.

In order to prove the theorem I consider a deterministic sequence of error terms $\{\mathbf{x}^t\}_{t \in \mathbb{N}}$, $\mathbf{x}^t \in \mathbb{R}^v$, that satisfy absolute convergence. For the given error terms define the sequence of opinions $\{\mathbf{a}_\mathbf{x}^t\}_{t \in \mathbb{N}}$ recursively,

$$\begin{aligned}\mathbf{a}_\mathbf{x}^1 &= \mathbf{a}^1 + \mathbf{x}^1 \\ \mathbf{a}_\mathbf{x}^t &= \mathbf{f}^t(\mathbf{a}_\mathbf{x}^{t-1}) + \mathbf{x}^t.\end{aligned}$$

The proof of the theorem makes use of the following two lemmas.

Lemma 6 *If for all agents i , the series $\sum_{t \in \mathbb{N}} |\mathbf{x}_i^t|$ is finite, and f_i^t is constricting and continuous for all t , then the sequence $\{\mathbf{a}_\mathbf{x}^t\}_{t \in \mathbb{N}}$ is bounded. For every $t \in \mathbb{N}$ we have*

$$\mathbf{a}_\mathbf{x}^{t+m} \in \mathbf{A}_\mathbf{x}^t = \left\{ \mathbf{a} \in \mathbb{R}^v : \mathbf{a}_i \in \left[\mathbf{a}_-^t - \sum_{l=1}^{\infty} \max_i |\mathbf{x}_i^{t+l}|, \mathbf{a}_+^t + \sum_{l=1}^{\infty} \max_i |\mathbf{x}_i^{t+l}| \right] \quad \forall i \in V \right\}$$

for all $m \in \mathbb{N}$.

Lemma 7 closes the gap between the deterministic error sequences and the almost sure result of the theorem.

Lemma 7 *Consider a finite product probability space $(\Omega^\infty, \mathcal{F}^\infty, P^\infty)$,*

$$\Omega^\infty = \Omega_1^\infty \times \dots \times \Omega_v^\infty, \mathcal{F}^\infty = \mathcal{F}_1^\infty \times \dots \times \mathcal{F}_v^\infty,$$

and let P_i^∞ denote the i -marginal of P^∞ . Consider a set $\mathbf{S} = (S_1, \dots, S_v) \in \mathcal{F}^\infty$ such that $P_i^\infty(S_i) = 1$ for all $i = 1, \dots, v$. Then $P^\infty(\mathbf{S}) = 1$.

The outline of the proof of Theorem 2 is as follows. I use Lemma 6 and the Bolzano-Weierstrass Theorem to show that $\{\mathbf{a}_\mathbf{x}^t\}_{t \in \mathbb{N}}$ has a convergent subsequence. I then show that the limit of the subsequence lies on the diagonal of the Euclidean space by applying Lemma 4 and Lemma 6. Finally, I use Lemma 6 together with the fact that the limit of the subsequence

lies on the diagonal to show that $\{\mathbf{a}_x^t\}_{t \in \mathbb{N}}$ is a Cauchy sequence and therefore converges to the limit of its converging subsequence. Lemma 7 then concludes the proof by establishing that the error series converges absolutely for all agents with probability one in the product space (across agents).

5 Relation to the Literature

This section briefly addresses the main points of distinction of the paper to the vast literature on non-Bayesian opinion formation in computer science and electrical engineering.¹⁴ Some of the important contributions here are Blondel, Hendrickx and Tsitsiklis [5], Jadbabaie, Lin and Morse [13], Lobel, Ozdaglar and Feijer [15], Moreau [19], Touri and Nedic [30], and Tsitsiklis, Bertsekas and Athans [31]. The central question in these papers concerns consensus. Here opinions are typically real valued or elements of a finite dimensional vector space. This literature generalizes the DeGroot [7] framework by allowing the weights of agents to depend on the current time period and the current observed opinions. Necessary and sufficient conditions for consensus are then characterized in terms of the sequence of weight matrices. The main innovation of this paper relative to these parallel literatures consists in a generalization of the opinion space to a linear continuum (in which case a weighted average approach cannot be applied in general) and in considering probabilistic shocks on opinions.

6 Discussion and Conclusion

The objective of this paper is to develop an understanding of long run non-Bayesian opinion formation in social networks. The existing theoretical models fall either under the full rationality assumption, as for example Gale and Kariv [11], Rosenberg, Solan and Vieille [29], and Mueller-Frank [22], or are grounded upon tractability considerations, as DeMarzo, Vayanos and Zwiebel [8] and Golub and Jackson [12]. This paper considers a general class of functions, encompassing weighted averages, whose defining property can be easily tested experimentally and empirically. The crucial property is that of constricting opinion formation under which agents position themselves in the (strict) interior of the smallest and largest observed opinions. The data gathered in a recent experiment by Lorenz et al. [16] indicates

¹⁴While writing the paper the author was unaware of this parallel literature.

that individuals indeed overwhelmingly satisfy constricting behavior in certain opinion formation environments. A rigorous analysis of their data and further experiments could be very insightful in regards to the range of environments in which subjects satisfy the constricting property.

Analyzing the implications of the individual properties of opinion formation, that is the constricting property, towards aggregate behavior we learn that even in widely dispersed (but strongly connected) societies, where all agents communicate only locally with a subset of society, consensus occurs in the long run. Therefore, under non-Bayesian observational learning in general environments failure of strong connectedness remains the leading explanation for failure of consensus. The tendency of individuals to uniform behavior that has been established in a wide array of theoretical contributions is shown to be valid for non-Bayesian opinion formation in social networks. This consensus result is also established to be robust against agents being subject to probabilistic (decreasing) errors.

The constricting property is introduced, and the consensus result derived, under the assumption that the set of opinions is linearly ordered. This restriction of the opinion space is not necessary. The constricting property carries forward to topological vector spaces. It is then defined as follows: f_i is constricting if $f_i(\mathbf{a})$ lies in the relative interior of the convex hull of the opinions agent i observes. That is

$$f_i(\mathbf{a}) \in \text{ri}(\text{co}\{a \in A : a = \mathbf{a}_i \text{ for some } j \in N_i \cup \{i\}\})$$

The consensus result holds and the same overall structure of the proof can be applied under the assumption that the topology on the opinion space is induced by a norm.¹⁵

It is important to emphasize that this paper makes two crucial assumptions that apply to many but clearly not all settings of opinion formation: (i) a rich opinion space, and (ii) a common order on the set of opinions, or more precisely the opinion set being a linear continuum. One natural setting in which these conditions apply is one of agents communicating probabilistic beliefs of some uncertain event occurring, as for example Germany winning the world cup in 2014. Another setting in which our assumptions are clearly not satisfied is one of agents discussing the likely winner of the world cup. A companion paper, Mueller-Frank and Neri [24], considers non-Bayesian opinion formation on a finite set of alternatives. It

¹⁵The proofs of the lemmas go through with minor adaptation. The main difference lies in the proof of the Theorem itself, particular in step 3. where one needs to establish that convergence of the subsequence to a non-consensus point leads to a contradiction. The proof is available on request.

is shown that while consensus is relatively easy to achieve in settings with a rich opinion space and a common order, for example for heterogeneous and time inconsistent agents, as established in this paper, under a finite set of alternatives agents need to be homogeneous and their revision function has to have a very specific form in order for consensus to occur.¹⁶

This paper establishes that constricting and continuous Markov revision functions lead to asymptotic consensus in strongly connected networks. The literature on Bayesian learning in networks, see Gale and Kariv [11], and Mueller-Frank [22], establishes (asymptotic) consensus as the outcome resulting from repeated interaction. Note however, that Bayesian revision functions can fail the constricting property.¹⁷ The Bayesian and non-Bayesian approach differs crucially in its assumption on the updating mechanism of agents. In particular, the revision of a Bayesian agent in period t depends on the whole realized and observed history leading up to period t , while the revision of a non-Bayesian agent depends only on the observed opinions of period $t - 1$. Therefore, no conclusions or implications can be drawn in neither direction.

The reader might be interested in the efficiency properties of the consensus opinion. Suppose that individuals start off with private information that they correctly incorporate in the first round of interaction. If agents subsequently act in a non-Bayesian constricting manner, under which conditions does the consensus opinion correctly represent some or even all of the private information of agents. In other words, how wise are unsophisticated crowds? Another companion paper, Mueller-Frank [23] establishes that generically in the class of weighted average revision functions the asymptotic consensus opinion fails to represent any subset of private information of agents. Based on the insight gained in the companion paper, there is little hope of achieving an efficient learning result in the general class of functions. While consensus in opinions might be useful for coordination purposes its information aggregation properties are undesirable.

The reader might also wonder how the results change if the constricting property is redefined in a sense that agents position themselves in the strict interior of the smallest and largest opinion of their neighbors, excluding their own last period opinion. This self-exclusive property allows to consider best response dynamics in games. For example, in a certain class of local interaction games, convention games, a deterministic best response dynamic as in Morris [20] satisfies the self-exclusive constricting property. In an earlier working paper, see Mueller-Frank [25], it is established that asymptotic consensus occurs in strongly connected

¹⁶ Agents are homogeneous if they use the same revision function.

¹⁷ For an example see pages 7-8 in Mueller-Frank [25].

and aperiodic networks if the revision functions of all agents are self-exclusive constricting and continuous, and all agents satisfy finite types.

One additional point to mention is that while the model and results provided in this paper are explicitly stated in the context of non-Bayesian opinion formation in networks they apply in other settings as well. Any dynamic process that satisfies the properties considered here, leads to asymptotic consensus. One example for a learning dynamic in games that satisfies the properties can be found in appendix B.

Finally, from a mathematical perspective Theorem 1 implies a fixed point convergence result for mappings whose components are constricting and continuous functions. The set of fixed points equals the set of consensus points. For every initial point, the recursive application of the mapping then leads to convergence to a fixed point. This fixed point convergence result might be of interest in other areas of economics beyond the applications analyzed here.

Appendix A

Proof of Lemma 1

Proof. The " \Rightarrow " direction follows immediately from the constricting property and the network being strongly connected. For the other direction consider a consensus profile $\mathbf{b} \in \mathbf{C}$ and a sequence $\{\mathbf{a}^t\}_{t \in \mathbb{N}}$ converging to \mathbf{b} such that

$$\mathbf{a}_i^t \neq \mathbf{a}_j^t$$

for all t and $i, j \in V$. Note that since the order topological space is Hausdorff the limit is unique. Consider a neighborhood \mathbf{V}_{t_0} of \mathbf{b} ,

$$\mathbf{V}_{t_0} = \{\mathbf{a} \in \mathbf{A} : \mathbf{a}_-^{V_0} \prec \mathbf{a}_i \prec \mathbf{a}_+^{V_0} \text{ for all } i\}$$

such that all $t \geq t_0$ we have $\mathbf{a}_t \in \mathbf{V}_{t_0}$. As f_i is constricting we have $\mathbf{a}_-^{V_0} \prec f_i(\mathbf{a}_t) \prec \mathbf{a}_+^{V_0}$ and therefore $\mathbf{f}^t(\mathbf{a}_t) \in \mathbf{V}_0$ for all $t \geq t_0$. Hence, $\{\mathbf{a}^t\}_{t \in \mathbb{N}}$ converging to \mathbf{b} implies that $\{\mathbf{f}(\mathbf{a}_t)\}_{t \in \mathbb{N}}$ converges to \mathbf{b} . Since \mathbf{f} is continuous (by continuity of all f_i) and $\{\mathbf{a}^t\}_{t \in \mathbb{N}}$ converges to \mathbf{b} we have that $\{\mathbf{f}(\mathbf{a}^t)\}_{t \in \mathbb{N}}$ converges to $\mathbf{f}(\mathbf{b})$ which implies $\mathbf{f}(\mathbf{b}) = \mathbf{b}$. ■

Proof of Lemma 3

Proof. Consider \mathbf{a}^1 with $\mathbf{a}_-^1 \prec \mathbf{a}_+^1$. The proof is presented in a number of simple steps. 1.) As f_i^t is constricting and continuous for all $i \in V$ and $t \in \mathbb{N}$, by Lemma 2 we have $\mathbf{a}_-^1 \preceq \mathbf{a}_i^t \preceq \mathbf{a}_+^1$ for all $i \in V$ and $t \in \mathbb{N}$. 2.) f_i^t constricting implies that if $\mathbf{a}_j^{t-1} \neq \mathbf{a}_i^{t-1}$ or $\mathbf{a}_j^{t-1} \notin \{\mathbf{a}_-^1, \mathbf{a}_+^1\}$ for some $j \in N_i$ then $\mathbf{a}_-^1 \prec \mathbf{a}_i^{t+k} \prec \mathbf{a}_+^1$ for all $k \in \mathbb{N}$. 3.) G strongly connected and $\mathbf{a}_-^1 \prec \mathbf{a}_+^1$ implies that there exists i with $\mathbf{a}_j \neq \mathbf{a}_i$ for some $j \in N_i$. For such i we have $\mathbf{a}_-^1 \prec \mathbf{a}_i^t = f_i^t(\mathbf{a}^{t-1}) \prec \mathbf{a}_+^1$ for all $t \geq 2$ by step two. 4.) Next consider the agents j such that $i \in N_j$. By step three $\mathbf{a}_-^1 \prec \mathbf{a}_i^t \prec \mathbf{a}_+^1$ for all $t \geq 2$ which implies $\mathbf{a}_-^1 \prec \mathbf{a}_j^t \prec \mathbf{a}_+^1$ for all $t \geq 3$ by step two. Let $d_G(j, i)$ denote the length of the shortest path from j to i in graph G . Applying the above reasoning inductively implies $\mathbf{a}_-^1 \prec \mathbf{a}_j^t \prec \mathbf{a}_+^1$ for all $t \geq d_G(j, i) + 1$. As G is strongly connected and finite there exists a finite $k \in \mathbb{N}$ such that such that for all $t \geq k$ we have $\mathbf{a}_-^1 \prec \mathbf{a}_i^t \prec \mathbf{a}_+^1$ for all $i \in V$. The cutoff k is bounded above by the diameter of the graph plus one. ■

Proof of Lemma 4

Proof. In order to show that \mathbf{A}^t is compact all I need is to show that the set A^t ,

$$A^t = \{a \in A : \mathbf{a}_-^t \preceq a \preceq \mathbf{a}_+^t\}$$

is compact since the finite product of compact sets is compact. Note that a simply ordered set endowed with the order topology is compact if and only if the set is a complete lattice. So for compactness of \mathbf{A}^t it is sufficient to establish that A^t is a complete lattice. Since A^t is bounded each of its subsets is bounded. A being a linear continuum implies that every bounded, non-empty subset has the least upper bound property. Since each non-empty subset of A^t is bounded above by \mathbf{a}_+^t and below by \mathbf{a}_-^t each non-empty subset has the least upper bound property (and greatest lower bound property) and A^t is therefore a complete lattice. ■

Proof of Lemma 5

Proof. The proof consists of two parts. The first part establishes, in five steps, that for any open set $\mathbf{V}_{\mathbf{a}^*}$ containing \mathbf{a}^* there exists t_λ such that $\mathbf{A}^{t_\lambda} \subset \mathbf{V}_{\mathbf{a}^*}$ for all $t_\lambda \in \Lambda$, $t_\lambda > t_\lambda$. 1.) $\mathbf{V}_{\mathbf{a}^*}$ equals the union of a collection of basis elements of $\tau_{\mathbf{A}}$. Therefore there exists a basis element $\mathbf{B} \subset \mathbf{V}_{\mathbf{a}^*}$ such that $\mathbf{a}^* \in \mathbf{B}$. 2.) As A is a linear continuum it is connected which implies that the basis of the order topology on A contains no singletons. 3.) The basis element \mathbf{B} equals the finite product of open sets $\prod_{i \in V} B_i$ where (i) each B_i is a base

element of the order topology on A , (ii) each B_i contains \mathbf{a}_i^* , and (iii) $B_i \neq \mathbf{a}_i^*$ by step 2. $\mathbf{B} \subset \mathbf{V}_{\mathbf{a}^*} \subset \mathbf{A}^1$ implies that each B_i is bounded below by \mathbf{a}_-^1 and above by \mathbf{a}_+^1 . 4.) A being a linear continuum implies that every bounded, non-empty subset has an infimum and supremum. Let b_{i-} and b_{i+} denote the infimum and supremum of B_i . Note that by step 3 $b_{i-} \prec \mathbf{a}_i^* \prec b_{i+}$ for all i . 5.) $\bigcap_{t_\lambda} \mathbf{A}^{t_\lambda} = \mathbf{a}^*$ and $\mathbf{A}^{t_\lambda} \subset \mathbf{A}^{t_{\lambda-1}}$ implies that for each i there exists a $t_{\lambda_{i-}}, t_{\lambda_{i+}}$ such that $b_{i-} \notin \mathbf{A}_i^{t_{\lambda_{i-}}}$ and $b_{i+} \notin \mathbf{A}_i^{t_{\lambda_{i+}}}$ for all $t_\lambda > t_{\lambda_{i-}}$ and $t_\lambda > t_{\lambda_{i+}}$ which implies that for all $t_\lambda > \max\{t_{\lambda_{1-}}, t_{\lambda_{1+}}, \dots, t_{\lambda_{v-}}, t_{\lambda_{v+}}\}$ we have $\mathbf{A}^{t_\lambda} \subset \mathbf{B} \subset \mathbf{V}_{\mathbf{a}^*}$. This establishes that for any open set $\mathbf{V}_{\mathbf{a}^*}$ containing \mathbf{a}^* there exists $t_{\lambda'}$ such that $\mathbf{A}^{t_\lambda} \subset \mathbf{V}_{\mathbf{a}^*}$ for all $t_\lambda \in \Lambda$, $t_\lambda > t_{\lambda'}$. To conclude the proof, note that by Lemma 2 for all $t \geq t_{\lambda'}$ we have $\mathbf{a}^t \in \mathbf{A}^{t_{\lambda'}}$. Convergence of $\{\mathbf{a}^{t_\lambda}\}_{t_\lambda \in \Lambda}$ to \mathbf{a}^* hence implies convergence of $\{\mathbf{a}^t\}_{t \in \mathbb{N}}$ to \mathbf{a}^* . ■

Proof of Lemma 6

Proof. For $\mathbf{a}_{\mathbf{x}}^t \in \mathbb{R}^v$ denote the maximal and minimal element by $\mathbf{a}_{\mathbf{x}+}^t = \max_{i \in V} \mathbf{a}_{\mathbf{x}+,i}^t$ and $\mathbf{a}_{\mathbf{x}-}^t = \min_{i \in V} \mathbf{a}_{\mathbf{x}-,i}^t$ respectively. As f_i is constricting and continuous we have $f_i(\mathbf{a}_{\mathbf{x}}^t) \in [\mathbf{a}_{\mathbf{x}-}^t, \mathbf{a}_{\mathbf{x}+}^t]$ for all $i \in V$. This implies the following bounds for $\mathbf{a}_{\mathbf{x}-}^{t+1}, \mathbf{a}_{\mathbf{x}+}^{t+1}$

$$\begin{aligned} \mathbf{a}_{\mathbf{x}-}^{t+1} &\geq \mathbf{a}_{\mathbf{x}-}^t - \max_{i \in v} |\mathbf{x}_i^{t+1}| \\ \mathbf{a}_{\mathbf{x}+}^{t+1} &\leq \mathbf{a}_{\mathbf{x}+}^t + \max_{i \in v} |\mathbf{x}_i^{t+1}|. \end{aligned}$$

As $\sum_{t \in \mathbb{N}} |\mathbf{x}_i^t|$ is finite for all i and V finite, there exists $\bar{u} \in \mathbb{R}_+$ such that

$$\sum_{t \in \mathbb{N}} \max_i |\mathbf{x}_i^t| < \bar{u}.$$

Therefore proceeding inductively from period one we have the following bounds on $\mathbf{a}_{\mathbf{x}-}^t, \mathbf{a}_{\mathbf{x}+}^t$

$$\begin{aligned} \mathbf{a}_{\mathbf{x}-}^t &> \mathbf{a}_-^1 - \bar{u} = \mathbf{a}_{\mathbf{x}-} \\ \mathbf{a}_{\mathbf{x}+}^t &< \mathbf{a}_+^1 + \bar{u} = \mathbf{a}_{\mathbf{x}+} \end{aligned}$$

for all $t \in \mathbb{N}$. Hence $\{\mathbf{a}_{\mathbf{x}}^t\}$ lies in $\mathbf{A}_{\mathbf{x}}^1 \subset \mathbb{R}^v$,

$$\mathbf{A}_{\mathbf{x}}^1 = \{\mathbf{a} \in \mathbb{R}^v : \mathbf{a}_i \in [\mathbf{a}_{\mathbf{x}-}, \mathbf{a}_{\mathbf{x}+}] \ \forall i \in V\}.$$

Therefore the sequence $\{\mathbf{a}_x^t\}$ is bounded. Moreover, for a given \mathbf{a}_x^t we have

$$\mathbf{a}_{x-}^{t+k} > \mathbf{a}_{x-}^t - \sum_{l=1}^{\infty} \max_{i \in V} |\mathbf{x}_i^{t+l}|$$

and

$$\mathbf{a}_{x+}^{t+k} < \mathbf{a}_{x+}^t + \sum_{l=1}^{\infty} \max_{i \in V} |\mathbf{x}_i^{t+l}|$$

for all $k \in \mathbb{N}$. ■

Proof of Lemma 7

Proof. By assumption and definition of the marginal probability measure we have

$$P_i^\infty(S_i) = P^\infty(\Omega_1^\infty, \dots, \Omega_{i-1}^\infty, S_i, \Omega_{i+1}^\infty, \dots, \Omega_v^\infty) = P^\infty(\mathbf{S}_i) = 1$$

for all $i = 1, \dots, n$. As a first step,

$$P^\infty(S_1, S_2, \Omega_3^\infty, \dots, \Omega_v^\infty) = 1$$

is established. Let $\mathbf{S}_1^C = \{S_1^C, \Omega_2^\infty, \Omega_3^\infty, \dots, \Omega_v^\infty\}$ denote the complement of \mathbf{S}_1 in Ω^∞ where S_1^C denotes the complement of S_1 in Ω_1^∞ . $P^\infty(\mathbf{S}_1) = 1$ implies $P^\infty(\mathbf{S}_1^C) = 0$. Partition the set \mathbf{S}_2 into the following two sets $\mathbf{S}_{12} = \{S_1, S_2, \Omega_3^\infty, \dots, \Omega_v^\infty\}$ and $\{S_1^C, S_2, \Omega_3^\infty, \dots, \Omega_v^\infty\}$. Countable additivity of the probability measure P^∞ implies

$$\begin{aligned} P^\infty(\mathbf{S}_2) &= P^\infty(\mathbf{S}_{12}) + P^\infty(S_1^C, S_2, \Omega_3^\infty, \dots, \Omega_v^\infty) \\ P^\infty(\mathbf{S}_{12}) &= P^\infty(\mathbf{S}_2) - P^\infty(S_1^C, S_2, \Omega_3^\infty, \dots, \Omega_v^\infty). \end{aligned}$$

As $\{S_1^C, S_2, \Omega_3^\infty, \dots, \Omega_v^\infty\} \subset \mathbf{S}_1^C$, $P^\infty(\mathbf{S}_1^C) = 0$ and $P^\infty(\mathbf{S}_2) = 1$ we have

$$P^\infty(S_1, S_2, \Omega_3^\infty, \dots, \Omega_v^\infty) = 1.$$

Proceeding inductively concludes the proof. ■

Proof of Theorem 2

Proof. Consider a deterministic sequence $\{\mathbf{x}^t\}_{t \in \mathbb{N}}$ of error terms and the corresponding sequence of opinions $\{\mathbf{a}_x^t\}_{t \in \mathbb{N}}$. The sequence of opinions is bounded by Lemma 6 and therefore has a converging subsequence $\{\mathbf{a}_x^{t_\lambda}\}_{t_\lambda \in \Lambda}$ by the Bolzano-Weierstrass Theorem. I employ the

same notation as in the proof of Theorem 1. Consider the subsequence $\{\mathbf{a}_x^{t_\lambda}\}_{t_\lambda \in \Lambda}$ and the corresponding sequence of functions $\{\mathbf{F}^{t_\lambda+1, t_\lambda+k}\}_{t_\lambda \in \Lambda}$. As $\mathbf{F}^{t+1, t+k} \in \mathcal{F}^k$ for all $t \in \mathbb{N}$ and \mathcal{F}^k is finite, implied by finite types and finite agents, there exists a $\mathbf{F} \in \mathcal{F}^k$ denoted as

$$\mathbf{F} = \hat{\mathbf{f}}^m \circ \hat{\mathbf{f}}^{m-1} \circ \dots \circ \hat{\mathbf{f}}^1$$

and an infinite subsequence $\{\mathbf{a}_x^{t_\gamma}\}_{t_\gamma \in \Gamma}$ of $\{\mathbf{a}_x^{t_\lambda}\}_{t_\lambda \in \Lambda}$ such that $\mathbf{F}^{t_\gamma+1, t_\gamma+m} = \mathbf{F}$ for all $t_\gamma \in \Gamma$. Consider the subsequence $\{\mathbf{a}_x^{t_\gamma}\}_{t_\gamma \in \Gamma}$ and for each $t_\gamma \in \Gamma$ the functions $\mathbf{G}_{t_\gamma}^k : \mathbb{R}^v \rightarrow \mathbb{R}^v$ for $m = 1, \dots, k$, where

$$\mathbf{G}_{t_\gamma}^1(\mathbf{a}) = \hat{\mathbf{f}}^1(\mathbf{a}) + \mathbf{x}^{t_\gamma+1}$$

and \mathbf{G}^k is defined recursively

$$\mathbf{G}_{t_\gamma}^m(\mathbf{a}) = \hat{\mathbf{f}}^k \left(\mathbf{G}_{t_\gamma}^{k-1}(\mathbf{a}) \right) + \mathbf{x}^{t_\gamma+m}.$$

Note that $\mathbf{G}_{t_\gamma}^k(\mathbf{a}_x^{t_\gamma}) = \mathbf{a}_x^{t_\gamma+k}$. Applying the chain rule for limits inductively it can be shown that the sequence of functions $\{\mathbf{G}_{t_\gamma}^m\}_{t_\gamma \in \Gamma}$ converges for every $\mathbf{a} \in \mathbb{R}^v$. We have

$$\lim_{t_\gamma \rightarrow \infty} \mathbf{G}_{t_\gamma}^k(\mathbf{a}) = \mathbf{F}(\mathbf{a}).$$

for every \mathbf{a} . Let \mathbf{a}^* denote the limit of $\{\mathbf{a}_x^{t_\lambda}\}_{t_\lambda \in \Lambda}$ and hence the limit of $\{\mathbf{a}_x^{t_\gamma}\}_{t_\gamma \in \Gamma}$. By $\mathbf{G}_{t_\gamma}^k(\mathbf{a}_x^{t_\gamma}) = \mathbf{a}_x^{t_\gamma+k}$, the reasoning above and continuity of \mathbf{F} we have

$$\lim_{t_\gamma \rightarrow \infty} \mathbf{a}_x^{t_\gamma+k} = \lim_{t_\gamma \rightarrow \infty} \lim_{\mathbf{a}_x^{t_\gamma} \rightarrow \mathbf{a}^*} \mathbf{G}_{t_\gamma}^k(\mathbf{a}_x^{t_\gamma}) = \lim_{\mathbf{a}_x^{t_\gamma} \rightarrow \mathbf{a}^*} \lim_{t_\gamma \rightarrow \infty} \mathbf{G}_{t_\gamma}^k(\mathbf{a}_x^{t_\gamma}) = \mathbf{F}(\mathbf{a}^*) = \hat{\mathbf{a}}^*.$$

Suppose that $\mathbf{a}^* \notin \mathbf{C}$. Lemma 4 then implies $\hat{\mathbf{a}}_-^* > \mathbf{a}_-^*$. Denote the difference by Δ , $\Delta = \hat{\mathbf{a}}_-^* - \mathbf{a}_-^*$. As $\{\mathbf{a}_x^{t_\gamma+k}\}_{t_\gamma \in \Gamma}$ converges to $\hat{\mathbf{a}}^*$, there exists a time period \bar{t} such that for all t_γ satisfying $t_\gamma + k > \bar{t}$ we have

$$\mathbf{a}_{\mathbf{x}-}^{t_\gamma+k} \geq \hat{\mathbf{a}}_-^* - \frac{\Delta}{2}.$$

The assumption $\sum_{t=1}^{\infty} |\mathbf{x}_i^t|$ being finite for all i implies that $\sum_{t=1}^{\infty} \max_i |\mathbf{x}_i^t|$ is finite. Therefore there exists \hat{t} such that $\sum_{t=\hat{t}}^{\infty} \max_i |\mathbf{x}_i^t| < \frac{\Delta}{4}$. By Lemma 6, for all $t > \hat{t}$ we have

$$\mathbf{a}_{\mathbf{x}-}^t > \mathbf{a}_{\mathbf{x}-}^{\hat{t}} - \frac{\Delta}{4}.$$

Consider a period t' such that $t' = t_\gamma + k$ for some t_γ and $t' > \max\{\bar{t}, \hat{t}\}$. We have

$$\mathbf{a}_{\mathbf{x}-}^{t'} \geq \hat{\mathbf{a}}_-^* - \frac{\Delta}{2}.$$

Furthermore we have for all period $t > t'$

$$\mathbf{a}_{\mathbf{x}-}^t > \mathbf{a}_{\mathbf{x}-}^{t'} - \frac{\Delta}{4}.$$

Combining the two yields

$$\mathbf{a}_{\mathbf{x}-}^t \geq \hat{\mathbf{a}}_-^* - \frac{3}{4}\Delta$$

for all $t > t'$ contradicting the convergence of $\{\mathbf{a}_{\mathbf{x}}^{t_\gamma}\}_{t_\gamma \in \Gamma}$ to \mathbf{a}^* . Therefore $\mathbf{a}^* \in \mathbf{C}$. So far it has been established that there exists a subsequence of $\{\mathbf{a}_{\mathbf{x}}^t\}_{t \in \mathbb{N}}$ that converges to a consensus vector. I show that the sequence $\{\mathbf{a}_{\mathbf{x}}^t\}_{t \in \mathbb{N}}$ converges to a consensus vector by establishing that $\{\mathbf{a}_{\mathbf{x}}^t\}_{t \in \mathbb{N}}$ is a Cauchy sequence. For each $t_\gamma \in \Gamma$ consider the sequence $\{\mathbf{a}_{\mathbf{x}}^{t_\gamma+m}\}_{m \in \mathbb{N}}$ which contains all but finitely many elements of $\{\mathbf{a}_{\mathbf{x}}^t\}_{t \in \mathbb{N}}$. By Lemma 6 the sequence $\{\mathbf{a}_{\mathbf{x}}^{t_\gamma+m}\}_{m \in \mathbb{N}}$ is bounded. Moreover, the sequence lies in the set $\mathbf{A}_{\mathbf{x}}^{t_\gamma}$ where

$$\mathbf{A}_{\mathbf{x}}^{t_\gamma} = \left\{ \mathbf{a} \in \mathbb{R}^v : \mathbf{a}_i \in \left[\mathbf{a}_{\mathbf{x}-}^{t_\gamma} - \sum_{l=1}^{\infty} \max |\mathbf{x}_i^{t_\gamma+l}|, \mathbf{a}_{\mathbf{x}+}^{t_\gamma} + \sum_{l=1}^{\infty} \max |\mathbf{x}_i^{t_\gamma+l}| \right] \right\}$$

But the diameter of $\mathbf{A}_{\mathbf{x}}^{t_\gamma}$ converges to zero with t_γ going to infinity as $\{\mathbf{a}_{\mathbf{x}}^{t_\gamma}\}_{t_\gamma \in \Gamma}$ converges to $\mathbf{a}^* \in \mathbf{C}$ and by the assumption $\sum_{t \in \mathbb{N}} |\mathbf{x}_i^t|$ being finite for all i . Therefore $\{\mathbf{a}_{\mathbf{x}}^t\}_{t \in \mathbb{N}}$ is a Cauchy sequence. As a Cauchy sequence in the Euclidean space converges to a limit \mathbf{a}^* if and only if there exists a subsequence with limit \mathbf{a}^* , the sequence of opinion vectors $\{\mathbf{a}_{\mathbf{x}}^t\}_{t \in \mathbb{N}}$ converges to $\mathbf{a}^* \in \mathbf{C}$.

For each agent i consider the probability space $(\Omega_i^\infty, \mathcal{F}_i^\infty, P_i^\infty)$ corresponding to the random variable $\sum_{t \in \mathbb{N}} |\mathbf{X}_i^t|$. Let S_i denote the subset of Ω_i^∞ where $\sum_{t \in \mathbb{N}} |\mathbf{X}_i^t|$ is finite. By assumption we have $P_i^\infty(S_i) = 1$ for all $i \in V$. Let $(\Omega^\infty, \mathcal{F}^\infty, P^\infty)$ denote the product space of the individual sample spaces of agents and let $S = (S_1, \dots, S_v)$ denote the set of states in Ω^∞ where the random variables $\sum_{t \in \mathbb{N}} |X_i^t|$ of all agents are finite. So far it has been established that asymptotic consensus holds for all states $\omega \in S$. In order to prove the theorem it remains to show that the set S has probability one in the product space $(\Omega^\infty, \mathcal{F}^\infty, P^\infty)$. If the random variables are independent across agents this immediately follows from $P_i^\infty(S_i) = 1$. If not,

then it follows from Lemma 7. ■

Appendix B

A Constricting Optimal Action Dynamic

This example is based on a combination of several standard frameworks of (network-based) repeated interaction under uncertainty. The idea behind the example is that agents receive initial private information regarding the realized state of the world, and repeatedly guess the state of the world while observing the history of choices of their neighbors.

Suppose that the prior distribution of the state of the world is standard normal. The state of the world θ is realized and agents observe a noisy signal about the realized state, $s_i = \theta + \varepsilon_i$ where the noise term ε_i is standard normal and independently distributed across agents. This distributional assumptions on the state of the world and signals are based upon the canonical framework for learning from normal random variables by Vives [32]. In each of countable rounds all agents select an action $\mathbf{a}_i^t \in \mathbb{R}$ that maximizes their expected utility in the given round conditioning on their private information. Hence agents are myopic in the way they select actions. This type of myopic behavior is commonly assumed in the literature on Bayesian learning in networks for tractability reasons. For examples see Bala and Goyal [4], Gale and Kariv [11] and Mueller-Frank [22]. The (commonly known) stage utility function of each agent i is given by

$$u_i(\mathbf{a}, \theta) = K - (\theta - \mathbf{a}_i)^2.$$

The utility functions are modeled as in a recent paper by Acemoglu, Bimpikis and Ozdaglar [1] on information exchange in social networks. In the first round of interaction the private information of agents is given by their private signal s_i . Note that the assumptions on the utility functions and the distribution on the state and signals imply that the expected utility maximizing action of an agent equals his private signal. At the beginning of the second round the private information of each agent i is based upon Bayesian inferences regarding the private signals of his neighbors given their first period actions. Since the agents perfectly reveal their private signal through their first period choice, the information of agent i at the beginning of round two equals the set of signals of his neighbors and himself. The expected utility maximizing action is then given by the equally weighted average of the signals of his neighbors and himself. Hence the optimal action of each agent in period two derives from a continuous and constricting revision function. Note that while the revision function to equal

the equally weighted average relies on the variance of the state and noise distribution being one, the optimal action for any assumption on the variances will be a weighted average of the signals as long as the mean of the distribution of θ and ε_i is zero. Hence for all such distributions of θ and ε_i the optimal choice in round two is based on a continuous and constricting revision function.

In regards to the private information of agents in round $t > 2$, I assume that all agents have a bounded memory of one period and in period t they behave as if period $t - 1$ was the first period. Under these assumptions the process of optimal actions $\{\mathbf{a}^t\}_{t \in \mathbb{N}}$ is based upon continuous and constricting revision functions.

The assumption of bounded memory can be justified with the large degree of complexity involved in fully Bayesian inferences in networks. It is common in the literature on non-Bayesian and even Bayesian learning to make assumption reducing the complexity of inferences. In Bala and Goyal [4] it is assumed that agents do not make inferences regarding the signals and actions of non-observed neighbors, and as mentioned previously the non-Bayesian papers by DeMarzo, Vayanos and Zwiebel [8], and Acemoglu, Ozdaglar and ParandehGheibi [2] assume that agents use a simple revision rule, the weighted average, when updating their opinion. Finally, there are a number of papers dealing with decision making in dynamic environments under bounded memory. For examples see Miller and Rozen [17], Monte and Said [18], Mullainathan [26], and Wilson [33].

The example also highlights that both continuity and the constricting property are satisfied for Bayesian learning in certain circumstances as here for the revision of actions from period one to period two.

Counter-Example: Relaxing Continuity and its Effect on Asymptotic Consensus

First note that the constricting definition imposes conditions only for the case where disagreement is observed by the agent. The natural way to relax continuity is to assume lower or upper semicontinuous functions. Consider a complete network consisting of two agents and the following upper semicontinuous revision functions of the agents

$$\begin{aligned} f_1(\mathbf{a}) &= \begin{cases} \frac{\mathbf{a}_1 + \mathbf{a}_2}{2} & \text{if } \mathbf{a}_1 \neq \mathbf{a}_2 \\ \mathbf{a}_1 + 1 & \text{if } \mathbf{a}_1 = \mathbf{a}_2 \end{cases} \\ f_2(\mathbf{a}) &= \begin{cases} \frac{\mathbf{a}_1 + \mathbf{a}_2}{2} & \text{if } \mathbf{a}_1 \neq \mathbf{a}_2 \\ \mathbf{a}_1 + 2 & \text{if } \mathbf{a}_1 = \mathbf{a}_2 \end{cases} . \end{aligned}$$

Assume that initial announcements vector equals $\mathbf{a}^1 = (0, 1)$. Then opinions evolve as follows $\{(0, 1); (.5, .5); (1.5, 2.5); (2, 2); (3, 4); (3.5, 3.5); \dots\}$. Opinions fail to converge and asymptotic consensus fails. The example also establishes failure of asymptotic consensus for lower semi-continuous functions if we change f_1 by replacing $+1$ with -1 and f_2 by replacing $+2$ with -2 .

Counter-Example: Relaxing Finite Types and its Effect on Asymptotic Consensus

A counter-example for failure of convergence under infinite types can be constructed in the following way. Consider a complete, two agent networks with initial announcement vector $\mathbf{a}^1 = (1, 0)$. Suppose agents use weighted average updating functions, with weights varying over time. Let α_t, β_t be the weights agents 1 and 2 assign to themselves in period t . Now consider a monotone sequence $\{\mathbf{a}^t\}_{t \in \mathbb{N}}$ in \mathbb{R}^2 that converges to $(\frac{3}{4}, \frac{1}{4})$. From the sequence of announcements $\{\mathbf{a}^t\}_{t \in \mathbb{N}}$ one can construct continuous and constricting updating functions $\{f_1^t\}_{t \in \mathbb{N}}$ and $\{f_2^t\}_{t \in \mathbb{N}}$ that generate the opinion process $\{\mathbf{a}^t\}_{t \in \mathbb{N}}$. The subsequent announcement vectors \mathbf{a}^{t-1} and \mathbf{a}^t yield the updating function in period t ,

$$\begin{aligned}\mathbf{a}_1^t &= \alpha_t \mathbf{a}_1^{t-1} + (1 - \alpha_t) \mathbf{a}_2^{t-1} \\ \mathbf{a}_2^t &= \beta_t \mathbf{a}_2^{t-1} + (1 - \beta_t) \mathbf{a}_1^{t-1}.\end{aligned}$$

Asymptotic consensus fails as the opinion process yields different opinions across agents in the limit. An apparent reason for failure of asymptotic consensus is the fast convergence of α_t and β_t to one which induces agents to approximately disregard the opinion of the other agent. Moreover, the limit revision function is not constricting.

References

- [1] D. Acemoglu, K. Bimpikis, A. Ozdaglar (2011), "Dynamics of Information Exchange in Endogenous Social Networks", working paper, MIT.
- [2] D. Acemoglu, A. Ozdaglar, A. ParandehGheibi (2010), "Spread of (Mis)Information in Social Networks", *Games and Economic Behavior* 70, 194-227.
- [3] C. D. Aliprantis, K. C. Border, Infinite Dimensional Analysis, Second Edition, Springer Verlag, New York, 1999.

- [4] V. Bala, S. Goyal (1998), "Learning from Neighbours", *Review of Economic Studies* 65, 595-621.
- [5] V.D. Blondel, J.M. Hendrickx, J.N. Tsitsikilis (2009), "On Krause's multi-agent consensus model with state dependent connectivity", *IEEE Transactions on Automatic Control* 54, 2586-2597.
- [6] S. Choi, D. Gale, S. Kariv (2005), Behavioral Aspects of Learning in Social Networks: An Experimental Study, in John Morgan., ed., "Advances in Applied Microeconomics", BE Press, Berkeley.
- [7] M. H. DeGroot (1974), "Reaching a Consensus", *Journal of the American Statistical Association* 69, 118-121.
- [8] P. DeMarzo, D. Vayanos, J. Zwiebel (2003), "Persuasion Bias, Social Influence and Uni-Dimensional Opinions", *Quarterly Journal of Economics* 118, 909-968.
- [9] R. Diestel, Graph Theory, Third Edition, Springer Verlag, New York, 2005.
- [10] G. Ellison, D. Fudenberg (1995), "Word-of-Mouth Communication and Social Learning", *Quarterly Journal of Economics* 110, 93-125.
- [11] D. Gale, S. Kariv (2003), "Bayesian learning in social networks", *Games and Economic Behavior* 45, 329-346.
- [12] B. Golub, M. Jackson (2010), "Naive Learning in Social Networks and the Wisdom of Crowds", *American Economic Journal: Microeconomics* 2:1, 112-149.
- [13] A. Jadbabaie, J. Lin, S. Morse (2003), "Coordination of groups of mobile autonomous agents using nearest neighbor rules", *IEEE Transactions on Automatic Control* 48, 988-1001.
- [14] A. Jadbabaie, P. Molavi, A. Sandroni, A. Tahbaz-Salehi (2012), "Non-Bayesian Social Learning", *Games and Economic Behavior* 76, 210-225.
- [15] I. Lobel, A. Ozdaglar, D. Feijer (2011), "Distributed multi-agent optimization with state-dependent communication", *Mathematical Programming*, 129, 255-284.
- [16] J. Lorenz, H. Rauhut, F. Schweitzer, D. Helbing (2011), "How social influence can undermine the wisdom of crowd effect", *Proceedings of the National Academy of Sciences* 108, 9020-9025.

- [17] D.A. Miller, K. Rozen (2010), "Monitoring with Collective Memory: Forgiveness for Optimally Empty Promises", working paper, Yale University.
- [18] D. Monte, M. Said (2010), "Learning in hidden Markov models with bounded memory", working paper, Microsoft Research.
- [19] L. Moreau (2005), "Stability of multiagent systems with time-dependent communication links", *IEEE Transactions on Automatic Control* 50, 169–182.
- [20] S. Morris (2000), "Contagion", *Review of Economic Studies* 67, 57-78.
- [21] E. Mossel, A. Sly, O. Tamuz (2012), "On Agreement and Learning", working paper, Weizmann Institute.
- [22] M. Mueller-Frank (2012), "A General Framework for Rational Learning in Social Networks", *Theoretical Economics*, forthcoming.
- [23] M. Mueller-Frank (2012), "Does one Bayesian make a Difference?", working paper, University of Oxford.
- [24] M. Mueller-Frank, C. Neri (2013), "Social Learning in Networks: Theory and Experiments", working paper, University of Oxford.
- [25] M. Mueller-Frank (2011), "The Robustness of Consensus in Network-Based Non-Bayesian Learning Models", working paper, University of Oxford.
- [26] S. Mullainathan (2002), "A Memory-Based Model of Bounded Rationality", *Quarterly Journal of Economics* 117(3), 735-774.
- [27] J.R. Munkres, *Topology*, Second Edition, Prentice Hall, 2000.
- [28] E. Rogers, F. Shoemaker, *Communication of Innovations: A Cross-Cultural Approach*, Second Edition, New York, Free Press, 1971.
- [29] D. Rosenberg, E. Solan, E. Vieille (2009), "Information externalities and emergence of consensus," *Games and Economic Behavior* 66, 979-994.
- [30] B. Touri, A. Nedic (2012), "On backward product of stochastic matrices", *Automatica* 48, 1477-1488.
- [31] J. N. Tsitsiklis, D. P. Bertsekas, M. Athans (1986), "Distributed asynchronous deterministic and stochastic gradient optimization algorithms", *IEEE Transactions on Automatic Control* 31, 803–812.

- [32] X. Vives (1993), "How fast do rational agents learn?", *Review of Economic Studies* 60, 329-347.
- [33] A. Wilson (2004), "Bounded Memory and Biases in Information Processing", unpublished manuscript, Harvard University.