Opinion dynamics with decaying confidence: application to community detection in graphs

Irinel-Constantin Morarescu, Antoine Girard

Opinion Dynamics with Decaying Confidence: Application to Community Detection in Graphs

Irinel-Constantin Morărescu*, Antoine Girard#

Abstract

We study a class of discrete-time multi-agent systems modeling opinion dynamics with decaying confidence. We consider a network of agents where each agent has an opinion. At each time step, each agent exchanges its opinion with its neighbors and updates its opinion by taking into account only its neighbors opinions that differs from its own opinion less than some confidence bound. This confidence bound is decaying: an agent gives repetitively confidence only to its neighbors that approach sufficiently fast its own opinion. Essentially, the agents try to reach an agreement with the constraint that it has to be approached no slower than a prescribed convergence rate. Under that constraint, global consensus may not be achieved and only local agreements may be reached. The agents reaching a local agreement form communities inside the network. In this paper, we analyze this opinion dynamics model: we show that communities correspond to asymptotically connected component of the network and give an algebraic characterization of communities in terms of eigenvalues of the matrix defining the collective dynamics. Finally, we apply our opinion dynamics model to address the problem of community detection in graphs. We propose a new formulation of the community detection problem based on eigenvalues of normalized Laplacian matrix of graphs and show that this problem can be solved using our opinion dynamics model. We provide experimental results that show that our opinion dynamics model provides an approach to community detection that is not only appealing but also effective.

Index Terms

Multi-agent systems, Consensus problem, Community detection, Algebraic graph theory

Laboratoire Jean Kuntzmann, Université de Grenoble, B.P. 53, 38041 Grenoble, France. This work was supported by UJF-MSTIC (CARESSE Project).

*`constantin.morarescu@imag.fr`

`antoine.girard@imag.fr`

I. Introduction

The analysis of multi-agent systems received an increasing interest in the past decades. In such systems, a set of agents interact according to simple local rules in order to achieve some global coordinated behavior. The most widely studied problem is certainly the consensus or agreement problem where each agent in the network maintains a value and repetitively averages its value with those of its neighbors, resulting in all the agents in the network reaching asymptotically a common value called consensus value. It is to be noted that the graph of interaction describing the network of agents is generally not fixed and may vary in time. Conditions ensuring consensus have been established by various authors including [JLM03], [BHOT05], [Mor05], [RB05] (see [OSFM07] for a survey). More recently, there have been several works providing estimations of the rate of convergence towards the consensus value [OT09], [AB08], [ZW09].

In this paper, we adopt a different point of view. We consider a discrete-time multi-agent system where the agents try to reach an agreement with the constraint that the consensus value must be approached no slower than a prescribed convergence rate. Under that constraint, global consensus may not be achieved and the agents may only reach local agreement. We call communities the subsets of agents reaching a consensus. Our model can be interpreted in terms of opinion dynamics. Each agent has an opinion. At each time step, the agent receives the opinions of its neighbors and then updates its opinion by taking a weighted average of its opinion and the opinions of its neighbors that are within some confidence range of its own opinion. The confidence ranges are getting smaller at each time step: an agent gives repetitively confidence only to the neighbors that approach sufficiently fast its own opinion. This can be seen as a model for negotiation process where an agent expects that its neighbors move significantly towards its opinion at each negotiation round in order to keep negotiating. Our model can be seen as an extension of the opinion dynamics with bounded confidence proposed by Krause in [Kra97] and studied in [HK02], [BHT06].

We analyze our opinion dynamics model by first studying the relation between asymptotic agreement of a subset of agents and the fact that they are asymptotically connected. We show that under suitable assumptions, these are actually equivalent (i.e. communities correspond to asymptotically connected component of the network) except for a set of initial opinions of measure 0. We then give an algebraic characterization of communities in terms of eigenvalues of the matrix defining the collective dynamics.

Finally, we apply our opinion dynamics model to address the problem of community detection in graphs. In the usual sense, communities in a graph are groups of vertices such that the concentration of edges inside communities is high with respect to the concentration of edges between communities. Given the increasing need of analysis tools for understanding complex networks in social sciences, biology, engineering or economics, the community detection problem has attracted a lot of attention in the recent years (see the extensive survey [For09]). We propose a new formulation of this problem based on eigenvalues of normalized Laplacian matrix of graphs. Then, we show that the communities that are obtained using our opinion dynamics model matches the classical notion of communities in graphs. Using two examples of network, we show that our opinion dynamics model not only provides an appealing approach to community detection but that it is also effective.

II. Opinion Dynamics with Decaying Confidence

A. Model Description

We study a discrete-time multi-agent model. We consider a set of n *agents*, $V = \{1, \dots, n\}$. A relation $E \subseteq V \times V$ models the *interactions* between the agents. We assume that the relation is symmetric ($(i,j) \in E$ iff $(j,i) \in E$) and anti-reflexive ($(i,i) \notin E$). V is the set of vertices and E is the set of edges of an undirected graph $G = (V, E)$, describing the network of agents. Each agent $i \in V$ has an *opinion* modeled by a real number $x_i(t) \in \mathbb{R}$. Initially, agent i has an opinion $x_i(0) = x_i^0$ independent from the opinions of the other agents. Then, at every time step, the agents update their opinion by taking a weighted average of its opinion and opinions of other agents:

$$x_i(t+1) = \sum_{j=1}^{n} p_{ij}(t) x_j(t) \tag{1}$$

with the coefficients $p_{ij}(t)$ satisfying

$$\forall i,j \in V,\ (p_{ij}(t) \neq 0 \iff j \in \{i\} \cup N_i(t)) \tag{2}$$

where $N_i(t)$ denotes the *confidence neighborhood* of agent i at time t:

$$N_i(t) = \left\{ j \in V \middle|\ ((i,j) \in E) \wedge \left(|x_i(t) - x_j(t)| \leq R\rho^t \right) \right\} \tag{3}$$

with $R > 0$ and $\rho \in (0,1)$ model parameters. We make the following additional assumptions:

Assumption 1: For $t \in \mathbb{N}$, the coefficients $p_{ij}(t)$ satisfy

(a) $p_{ij}(t) \in [0,1]$, for all $i, j \in V$.

(b) $\sum_{j=1}^{n} p_{ij}(t) = 1$, for all $i \in V$.

This model can be interpreted in terms of opinion dynamics. At each time step t, agent $i \in V$ receives the opinions of its neighbors in the graph G. If the opinion of i differs from the opinion of its neighbor j more than a certain threshold $R\rho^t$, then i does not give confidence to j and does not take into account the opinion of j when updating its own opinion. The parameter ρ characterizes the confidence decay of the agents. Agent i gives repetitively confidence only to neighbors whose opinion converges sufficiently fast to its own opinion. This model can be interpreted in terms of negotiations where agent i requires that, at each negotiation round, the opinion of agent j moves significantly towards its own opinion in order to keep negotiating with j.

Remark 1: We assume that $\rho \in (0,1)$, however, let us remark that for a complete graph G (every agent talks with all the other agents) and $\rho = 1$ (there is no confidence decay), our model would coincide with Krause model of opinion dynamics with bounded confidence [Kra97], [HK02], [BHT06].

Our first result states that the opinion of each agent converges to some limit value:

Proposition 1: Under Assumption 1, for all $i \in V$, the sequence $(x_i(t))_{t\in\mathbb{N}}$ is convergent. We denote x_i^* its limit. Furthermore, we have for all $t \in \mathbb{N}$,

$$|x_i(t) - x_i^*| \leq \frac{R}{1-\rho}\rho^t. \tag{4}$$

Proof: Let $i \in V$, $t \in \mathbb{N}$, we have from (1), Assumption 1 and (2)

$$\begin{aligned}
|x_i(t+1) - x_i(t)| &= \left| \left(\sum_{j=1}^{n} p_{ij}(t) x_j(t) \right) - x_i(t) \right| \\
&= \left| \sum_{j=1}^{n} p_{ij}(t)(x_j(t) - x_i(t)) \right| \\
&= \left| \sum_{j \in N_i(t)} p_{ij}(t)(x_j(t) - x_i(t)) \right| \\
&\leq \sum_{j \in N_i(t)} p_{ij}(t) |x_j(t) - x_i(t)|
\end{aligned}$$

Then, it follows from equation (3) that

$$|x_i(t+1) - x_i(t)| \leq \sum_{j \in N_i(t)} p_{ij}(t) R\rho^t$$

Finally, Assumption 1 gives for all $t \in \mathbb{N}$

$$|x_i(t+1) - x_i(t)| \leq (1 - p_{ii}(t)) R\rho^t \leq R\rho^t.$$

Let $t \in \mathbb{N}$, $\tau \in \mathbb{N}$, then

$$|x_i(t+\tau) - x_i(t)| \leq \sum_{k=0}^{\tau-1} |x_i(t+k+1) - x_i(t+k)| \leq \sum_{k=0}^{\tau-1} R\rho^{t+k}$$

Therefore,

$$|x_i(t+\tau) - x_i(t)| \leq \frac{R}{1-\rho}\rho^t(1-\rho^\tau) \leq \frac{R}{1-\rho}\rho^t \tag{5}$$

which shows, since $\rho \in (0,1)$, that the sequence $(x_i(t))_{t\in\mathbb{N}}$ is a Cauchy sequence in $\mathbb{R}$. Therefore, it is convergent. Equation (4) is obtained from (5) by letting τ go to $+\infty$. ■

The previous proposition allows us to complete the interpretation of our opinion dynamics model. The agents try to reach an agreement with the constraint that the consensus value must be approached no slower than $O(\rho^t)$. Under that constraint, global agreement may not be attainable and the agents may only reach local agreements. We refer to the sets of agents that asymptotically agree as communities.

Definition 1: Let $i, j \in V$, we say that agents i and j *asymptotically agree*, denoted $i \sim^* j$, if and only if $x_i^* = x_j^*$.

It is straightforward to verify that $\sim^*$ is an equivalence relation over V.

Definition 2: A *community* $C \subseteq V$ is an element of the quotient set $\mathcal{C} = V/\sim^*$.

Let us remark that the community structure is dependent on the initial distribution of opinions. In the following, we shall provide some insight on the structure of these communities. But first, we need to introduce some additional notations.

B. Notations and Preliminaries

We define the set of interactions at time t, $E(t) \subseteq V \times V$ as

$$E(t) = \left\{(i,j) \in E |\; |x_i(t) - x_j(t)| \leq R\rho^t \right\}.$$

Let us remark that $(i,j) \in E(t)$ if and only if $j \in N_i(t)$. The interaction graph at time t is then $G(t) = (V, E(t))$.

For a set of agents $I \subseteq V$, the subset of edges of G connecting the agents in I is $E_I = E \cap (I \times I)$. Let $E' \subseteq E_I$ be a symmetric relation over I, then the graph $G' = (I, E')$ is called a *subgraph* of G. If $I = V$, then the graph $G' = (V, E')$ is called a *spanning subgraph* of G. The set of spanning subgraphs of G is denoted $\mathcal{S}(G)$. For all $t \in \mathbb{N}$, $G(t) \in \mathcal{S}(G)$. Let us remark that the set $\mathcal{S}(G)$ is finite: it has $2^{|E|/2}$ elements where $|E|$ denotes the number of elements in E. Given a partition of the agents $\mathcal{P} = \{I_1, \dots, I_p\}$, we define the set of edges $E_{\mathcal{P}} = \bigcup_{I \in \mathcal{P}} E_I$ and the spanning subgraph of G, $G_{\mathcal{P}} = (V, E_{\mathcal{P}})$. Essentially, $G_{\mathcal{P}}$ is the spanning subgraph of G obtained by removing all the edges between agents belonging to different elements of the partition $\mathcal{P}$. An interesting such graph is the *graph of communities* $G_{\mathcal{C}} = (V, E_{\mathcal{C}})$ where:

$$E_{\mathcal{C}} = \{(i,j) \in E |\; i \sim^* j\}\,.$$

Let $G' = (V, E') \in \mathcal{S}(G)$, a *path* in G' is a finite sequence of edges $(i_1, i_2), (i_2, i_3), \dots, (i_p, i_{p+1})$ such that $(i_k, i_{k+1}) \in E'$ for all $k \in \{1, \dots, p\}$. Two vertices i, $j \in V$ are *connected* in G' if there exists a path in G' joining i and j (i.e. $i_1 = i$ and $j_p = j$). A subset of agents $I \subseteq V$ is a *connected component* of G' if for all i, $j \in I$ with $i \neq j$, i and j are connected in G' and for all $i \in I$, for all $j \in V \setminus I$, i and j are not connected in G'. The set of connected components of G' is denoted $\mathcal{K}(G')$. Let us remark that $\mathcal{K}(G')$ is a partition of V.

We define the vectors of opinions $x(t) = (x_1(t), \dots, x_n(t))^\top$ and of initial opinions $x^0 = (x_1^0, \dots, x_n^0)^\top$. The dynamics of the vector of opinions is then given by

$$x(t+1) = P(t)x(t)$$

where $P(t)$ is the row stochastic matrix with entries $p_{ij}(t)$. For a set of agents $I \subseteq V$, with $I = \{v_1, \dots, v_k\}$, we define the vector of opinions $x_I(t) = (x_{v_1}(t), \dots, x_{v_k}(t))^\top$. Given a $n \times n$ matrix A with entries a_{ij}, we define the $k \times k$ matrix A_I whose entries are the $a_{v_i v_j}$. In particular, $P_I(t)$ is the matrix with entries $p_{v_i v_j}(t)$. Let us remark that $P_I(t)$ is generally not row stochastic. We state the following preliminary result that will be useful in further sections.

Lemma 1: Let $I \subseteq V$ be a subset of agents such that no agent in I is connected to an agent in $V \setminus I$ in the graph $G(t)$, then under Assumption 1

$$x_I(t+1) = P_I(t)x_I(t)$$

and $P_I(t)$ is an aperiodic row stochastic matrix. Moreover, if I is a connected component of $G(t)$ then $P_I(t)$ is irreducible.

Proof: Let $i \in I$, then

$$x_i(t+1) = \sum_{j=1}^{n} p_{ij}(t)x_j(t) = \sum_{j\in I} p_{ij}(t)x_j(t) + \sum_{j\in V\setminus I} p_{ij}(t)x_j(t).$$

Since i is not connected to any agent in $V\setminus I$ in $G(t)$, it follows from equation (2) that $p_{ij}(t) = 0$ for all $j \in V \setminus I$. Therefore,

$$x_i(t+1) = \sum_{j\in I} p_{ij}(t)x_j(t)$$

which gives $x_I(t+1) = P_I(t)x_I(t)$. Similarly, we obtain from Assumption 1

$$1 = \sum_{j=1}^{n} p_{ij}(t) = \sum_{j\in I} p_{ij}(t)$$

Since in addition, all entries of $P_I(t)$ are nonnegative, it follows that $P_I(t)$ is a row stochastic matrix. It is aperiodic because equation (2) gives that for all $i \in I$, $p_{ii}(t) > 0$. By definition of irreducibility, if I is a connected component of $G(t)$ then $P_I(t)$ is irreducible. ∎

The following sections are devoted to the analysis of the community structure of the network of agents.

III. Asymptotic Connectivity and Agreement

In this section, we explore the relation between communities and asymptotically connected components of the network. Let us remark that the set of edges E can be classified into two subsets as follows:

$$E^f = \{(i,j) \in E |\ \exists t_{ij} \in \mathbb{N},\ \forall s \geq t_{ij},\ (i,j) \notin E(s)\}$$

and

$$E^\infty = \{(i,j) \in E |\ \forall t \in \mathbb{N},\ \exists s \geq t,\ (i,j) \in E(s)\}.$$

Intuitively, an edge (i,j) is in E^f if the agents i and j stop interacting with each other in finite time. E^∞ consists of the interactions between agents that are infinitely recurrent. It is clear that $E^f \cap E^\infty = \emptyset$ and $E = E^f \cup E^\infty$. Also, since E and thus E^f is a finite set, there exists $T \in \mathbb{N}$ such that

$$\forall (i,j) \in E^f,\ \forall s \geq T,\ (i,j) \notin E(s). \tag{6}$$

Let us remark that the sets E^f and E^∞ and the natural number T generally depend on the vector of initial opinions x^0. We define the graph $G^\infty = (V, E^\infty)$.

Definition 3: Let $i, j \in V$, we say that agents i and j are *asymptotically connected* if and only if i and j are connected in G^∞. We say that they are *asymptotically disconnected* if they are not asymptotically connected.

A. Asymptotic Connectivity Implies Asymptotic Agreement

Proposition 2: Under Assumption 1, if two agents $i, j \in V$ are asymptotically connected then they asymptotically agree.

Proof: Let $(i_1, i_2), (i_2, i_3), \ldots, (i_p, i_{p+1})$ be a path in G^∞ joining $i = i_1$ and $j = i_{p+1}$. We can choose a path without loops (i.e. $i_k \neq i_{k'}$ for all $k \neq k'$) and therefore $p \leq n-1$. Let $t \in \mathbb{N}$, $k \in \{1, \ldots, p\}$, since $(i_k, i_{k+1}) \in E^\infty$ there exists $\tau_k \in \mathbb{N}$ such that $(i_k, i_{k+1}) \in E(t + \tau_k)$. Let us define in addition $\tau_0 = 0$, $\tau_{p+1} = 0$, then

$$\begin{aligned}
|x_i(t) - x_j(t)| &= |x_{i_1}(t + \tau_0) - x_{i_{p+1}}(t + \tau_{p+1})| \\
&\leq |x_{i_1}(t + \tau_0) - x_{i_1}(t + \tau_1)| \\
&\quad + \sum_{k=1}^{p} \left(|x_{i_k}(t + \tau_k) - x_{i_{k+1}}(t + \tau_k)| + |x_{i_{k+1}}(t + \tau_k) - x_{i_{k+1}}(t + \tau_{k+1})| \right) \\
&\leq \sum_{k=1}^{p+1} |x_{i_k}(t + \tau_{k-1}) - x_{i_k}(t + \tau_k)| + \sum_{k=1}^{p} |x_{i_k}(t + \tau_k) - x_{i_{k+1}}(t + \tau_k)|.
\end{aligned}$$

Let $\underline{\tau}_k = \min(\tau_{k-1}, \tau_k)$ for $k \in \{1, \ldots, p+1\}$, then from equation (5) we have for all $k \in \{1, \ldots, p+1\}$

$$|x_{i_k}(t + \tau_{k-1}) - x_{i_k}(t + \tau_k)| \leq \frac{R}{1-\rho} \rho^{t + \underline{\tau}_k} \leq \frac{R}{1-\rho} \rho^t.$$

Moreover, since $(i_k, i_{k+1}) \in E(t + \tau_k)$

$$|x_{i_k}(t + \tau_k) - x_{i_{k+1}}(t + \tau_k)| \leq R\rho^{t + \underline{\tau}_k} \leq R\rho^t.$$

Therefore,

$$|x_i(t) - x_j(t)| \leq (p+1) \frac{R}{1-\rho} \rho^t + pR\rho^t \leq n \frac{R}{1-\rho} \rho^t + (n-1)R\rho^t$$

which shows by letting t go to $+\infty$ that i and j asymptotically agree. ■

Remark 2: The notion of asymptotic connectivity has already been considered in several works (including [JLM03], [BHOT05], [Mor05]) for proving consensus in multi-agent systems.

Actually, the previous proposition could be proved using Theorem 3 in [Mor05]. However, for the sake of self-containment, we preferred to provide a simpler proof of the result that uses the specificities of our model.

B. Asymptotic Agreement Implies Asymptotic Connectivity

The converse result of Proposition 2 is much more challenging: it is clear that it cannot hold for all initial conditions. Indeed, if all the initial opinions x_i^0 are identical, then it is clear that the agents asymptotically agree independently from the fact that they are asymptotically connected or not. Hence, the better we can expect is the converse result to hold for almost all initial conditions. In this paragraph, we will need additional assumptions in order to be able to prove this result. The first one is the following:

Assumption 2: The sequence of matrices $P(t)$ satisfy the following conditions:

(a) For all $t \in \mathbb{N}$, $P(t)$ is invertible.

(b) For all $t \in \mathbb{N}$, $t' \in \mathbb{N}$, if $G(t') = G(t)$ then $P(t') = P(t)$.

Let us remark that the first assumption can be enforced, for instance, by choosing $p_{ii}(t) > 1/2$ for all $i \in V$, for all $t \in \mathbb{N}$, in that case $P(t)$ is a strictly diagonally dominant matrix and therefore it is invertible. The second assumption states that $P(t)$ only depends on the graph $G(t)$, then we shall write $P(t) = P(G(t))$ where $P(G')$ is the matrix associated to a graph $G' \in \mathcal{S}(G)$. From the first assumption, $P(G')$ must be invertible. Then, we can define for all $t \in \mathbb{N}$, the following set of matrices:

$$\mathcal{Q}_t = \left\{ P(G_0)^{-1} P(G_1)^{-1} \dots P(G_{t-1})^{-1} |\ G_k \in \mathcal{S}(G),\ 0 \le k \le t-1 \right\}. \tag{7}$$

Let us remark that since $\mathcal{S}(G)$ is finite, the set $\mathcal{Q}_t$ is finite: it has at most $2^{t \times |E|/2}$ elements.

We shall now prove the converse result of Proposition 2 in two different cases.

1) Average preserving dynamics: We first assume that the opinion dynamics preserves the average of the opinions:

Assumption 3: For all $t \in \mathbb{N}$, for all $j \in V$, $\sum_{i=1}^n p_{ij}(t) = 1$

This assumption simply means that the matrix $P(t)$ is doubly stochastic. It is therefore average preserving: the average of $x(t)$ is equal to the average of $x(t+1)$. Similar to Lemma 1, we can prove the following lemma:

Lemma 2: Let $I \subseteq V$ be a subset of agents such that no agent in I is connected to an agent in $V \setminus I$ in the graph $G(t)$, then under Assumption 3, $P_I(G(t))$ is average preserving.

Proof: It is sufficient to remark that for all $j \in I$, and $i \in V \setminus I$, equation (2) gives that $p_{ij}(t) = 0$. Therefore it follows that for all $j \in I$,

$$\sum_{i \in I} p_{ij}(t) = \sum_{i=1}^{n} p_{ij}(t) = 1$$

which proves that $P_I(G(t))$ is average preserving. ■

We now state the main result of the section:

Theorem 1: Under Assumptions 1, 2 and 3, for almost all vectors of initial opinions x^0, two agents $i, j \in V$ asymptotically agree if and only if they are asymptotically connected.

Proof: The if part of the theorem is a consequence of Proposition 2. To prove the only if part, let us define the following set

$$\mathcal{W} = \{(I,J)|\ (I \subseteq V) \wedge (I \neq \emptyset) \wedge (J \subseteq V) \wedge (J \neq \emptyset) \wedge (I \cap J = \emptyset)\}$$

Since V is a finite set, it is clear that $\mathcal{W}$ is finite (it has less than 2^{2n} elements). For all $(I,J) \in \mathcal{W}$, let $|I|$ and $|J|$ denote the number of elements of I and J respectively. We define the vector of $\mathbb{R}^n$, c_{IJ} whose coordinates $c_{IJ,k} = 1/|I|$ if $k \in I$, $c_{IJ,k} = -1/|J|$ if $k \in J$, and $c_{IJ,k} = 0$ otherwise. We define the $(n-1)$-dimensional subspace of $\mathbb{R}^n$:

$$H_{IJ} = \{x \in \mathbb{R}^n|\ c_{IJ} \cdot x = 0\}.$$

Finally, let us define the subset of $\mathbb{R}^n$:

$$X^0 = \bigcup_{t \in \mathbb{N}} \left(\bigcup_{(I,J) \in \mathcal{W}} \left(\bigcup_{Q \in \mathcal{Q}_t} Q H_{IJ} \right) \right)$$

where $\mathcal{Q}_t$ is the set of matrices defined in (7). Since $\mathcal{W}$ is a finite set and for all $t \in \mathbb{N}$, $\mathcal{Q}_t$ are finite sets, X^0 is a countable union of $(n-1)$-dimensional subspaces of $\mathbb{R}^n$. Therefore X^0 has Lebesgue measure 0.

Let $x^0 \in \mathbb{R}^n$ be a vector of initial opinions, let us assume that there exist two agents $i, j \in V$ that asymptotically agree but are asymptotically disconnected. Let us show that necessarily, x^0 belongs to the set X^0. Let I and J denote the connected components of G^∞ containing i and j respectively. Since i and j are asymptotically disconnected, $I \cap J = \emptyset$, therefore $(I,J) \in \mathcal{W}$. Let T be defined as in equation (6), since no agent in I is connected to an agent outside of I

in G^∞ (and hence in $G(t)$ for $t \geq T$), we have by Lemma 1, that for all $t \geq T$, $x_I(t+1) = P_I(G(t))x_I(t)$. From Lemma 2, we have that $P_I(G(t))$ is average preserving. Therefore, for all $t \geq T$, the average of $x_I(t)$ is the same as the average of $x_I(T)$. From Proposition 2, all agents in I asymptotically agree, then the limit value is necessarily the average of $x_I(T)$. Therefore $x_i^* = (\mathbf{1}_{|I|} \cdot x_I(T))/|I|$ where $\mathbf{1}_{|I|}$ denote the $|I|$-dimensional vector with all entries equal to 1. A similar discussion gives that $x_j^* = (\mathbf{1}_{|J|} \cdot x_J(T))/|J|$. Since i and j asymptotically agree, we have $(\mathbf{1}_{|I|} \cdot x_I(T))/|I| = (\mathbf{1}_{|J|} \cdot x_J(T))/|J|$. This means that $x(T) \in H_{IJ}$ and therefore

$$x^0 = P(G(0))^{-1}P(G(1))^{-1} \dots P(G(T-1))^{-1}x(T) \in \bigcup_{Q \in \mathcal{Q}_T} QH_{IJ}$$

which leads to $x^0 \in X^0$. ■

Hence, in the case of average preserving dynamics, asymptotic connectivity is equivalent to asymptotic agreement for almost all vectors of initial opinions. We shall now prove a similar result under different assumptions.

2) Fast convergence assumption: We now replace the average preserving assumption by another assumption. From Proposition 1, we know that the opinion of each agent converges to its limit value no slower than $O(\rho^t)$. This is an upper bound, numerical experiments show that in practice the convergence to the limit value is often slightly faster than $O(\rho^t)$. This observation motivates the following assumption:

Assumption 4: There exists $\underline{\rho} < \rho$ and $M \geq 0$ such that for all $i \in V$, for all t$\in \mathbb{N}$,

$$|x_i(t) - x_i^*| \leq M\underline{\rho}^t.$$

Remark 3: The previous assumption always holds unless there exists $i \in V$ such that

$$\limsup_{t \to +\infty} \frac{1}{t} \log(|x_i(t) - x_i^*|) = \log(\rho).$$

It should be noted that unlike Assumptions 1, 2 and 3, it is generally not possible to check a priori whether Assumption 4 holds. However, numerical experiments tend to show that in practice, Assumption 4 holds.

The previous assumption allows us to state the following result:

Lemma 3: Under Assumptions 1 and 4, there exists $T' \in \mathbb{N}$ such that for all $t \geq T'$, $G(t) = G^\infty$. Moreover, $G^\infty = G_{\mathcal{C}}$.

Proof: We shall prove the lemma by showing that there exists $T' \in \mathbb{N}$ such that for all $t \geq T'$, $E(t) \subseteq E^\infty \subseteq E_{\mathcal{C}} \subseteq E(t)$. Firstly, let $T_1 \geq T$ where T is defined as in equation (6), then for all

$t \geq T_1$, $E(t) \subseteq E^\infty$. Secondly, let $(i,j) \in E^\infty$, then agents i and j are asymptotically connected. From Proposition 2, it follows that i and j asymptotically agree. Therefore, $(i,j) \in E_{\mathcal{C}}$. Thirdly, let $(i,j) \in E_{\mathcal{C}}$, then $x_i^* = x_j^*$ and for all $t \in \mathbb{N}$

$$\begin{aligned} |x_i(t) - x_j(t)| &\leq |x_i(t) - x_i^*| + |x_i^* - x_j^*| + |x_j(t) - x_j^*| \\ &\leq |x_i(t) - x_i^*| + |x_j(t) - x_j^*| \end{aligned}$$

From Assumption 4, we have for all $t \in \mathbb{N}$,

$$|x_i(t) - x_j(t)| \leq 2M\underline{\rho}^t.$$

Since $\underline{\rho} < \rho$, there exists $T_2 \in \mathbb{N}$, such that for all $t \geq T_2$, $2M\underline{\rho}^t \leq R\rho^t$. Then, for all $t \geq T_2$, $(i,j) \in E(t)$. Let $T' = \max(T_1, T_2)$, then for all $t \geq T'$, $E(t) = E^\infty = E_{\mathcal{C}}$ and thus $G(t) = G^\infty = G_{\mathcal{C}}$. ■

The previous result states that after a finite number of steps, the graph of interactions between agents remains always the same. Then, we can state a result similar to Theorem 1:

Theorem 2: Under Assumptions 1, 2 and 4, for almost all vectors of initial opinions x^0, two agents $i, j \in V$ asymptotically agree if and only if they are asymptotically connected.

Proof: The if part of the theorem is a consequence of Proposition 2. To prove the only if part, let us define the following set associated to a spanning subgraph $G' \in \mathcal{S}(G)$:

$$\mathcal{W}(G') = \{(I,J)|\ (I \subseteq V) \wedge (J \subseteq V) \wedge (I \neq J) \wedge (I \in \mathcal{K}(G')) \wedge (J \in \mathcal{K}(G'))\}$$

Since V is a finite set, it is clear that $\mathcal{W}(G')$ is finite (it has less than 2^{2n} elements). Let $(I,J) \in \mathcal{W}(G')$, $I = \{v_1, \dots, v_{|I|}\}$, $J = \{w_1, \dots, w_{|J|}\}$. Since I and J are connected components of G', we have from Lemma 1 that $P_I(G')$ and $P_J(G')$ are aperiodic irreducible row stochastic matrices. Let $e_I(G')$ and $e_J(G')$ be the left Perron eigenvectors of $P_I(G')$ and $P_J(G')$, respectively:

$$e_I(G')^\top P_I(G') = e_I(G')^\top \text{ and } e_I(G') \cdot \mathbf{1}_{|I|} = 1$$

and

$$e_J(G')^\top P_J(G') = e_J(G')^\top \text{ and } e_J(G') \cdot \mathbf{1}_{|J|} = 1.$$

We define the vector of $\mathbb{R}^n$, c_{IJ} whose coordinates are given by $c_{IJ,v_k} = e_{I,k}$ if $v_k \in I$, $c_{IJ,w_k} = -e_{J,k}$ if $w_k \in J$ and $c_{IJ,k} = 0$ if $k \in V \setminus (I \cup J)$. We define the $(n-1)$-dimensional subspace of $\mathbb{R}^n$:

$$H_{IJ}(G') = \{x \in \mathbb{R}^n |\ c_{IJ}(G') \cdot x = 0\}.$$

Finally, let us define the subset of $\mathbb{R}^n$:

$$X^0 = \bigcup_{t\in\mathbb{N}} \left(\bigcup_{G'\in\mathcal{S}(G)} \left(\bigcup_{(I,J)\in\mathcal{W}(G')} \left(\bigcup_{Q\in\mathcal{Q}_t} QH_{IJ}(G') \right)\right)\right) \tag{8}$$

where $\mathcal{Q}_t$ is the set of matrices defined in (7). $\mathcal{S}(G)$ is a finite set and for all $G' \in \mathcal{S}(G)$, $\mathcal{W}(G')$ is a finite set. Moreover for all $t \in \mathbb{N}$, $\mathcal{Q}_t$ is a finite set. Then, X^0 is a countable union of $(n-1)$-dimensional subspaces of $\mathbb{R}^n$. Therefore X^0 has Lebesgue measure 0.

Let $x^0 \in \mathbb{R}^n$ be a vector of initial opinions, let us assume that there exist two agents $i, j \in V$ that asymptotically agree but are asymptotically disconnected. Let us show that necessarily, x^0 belongs to the set X^0. Let I and J denote the connected components of G^∞ containing i and j respectively. Since i and j are asymptotically disconnected, $I \neq J$, therefore $(I,J) \in \mathcal{W}(G^\infty)$. Since I is a connected component of G^∞, it follows from Lemma 1 and Lemma 3 that for all $t \geq T'$, $x_I(t+1) = P_I(G^\infty)x_I(t)$. Moreover, $P_I(G^\infty)$ is an aperiodic irreducible row stochastic matrix and from the Perron-Frobenius Theorem (see e.g. [Sen81]), it follows that 1 is a simple eigenvalue of $P_I(G^\infty)$ and all other eigenvalues of $P_I(G^\infty)$ have modulus strictly smaller than 1. Therefore,

$$\lim_{t\to+\infty} x_I(t) = (e_I(G^\infty)\cdot x_I(T'))\mathbf{1}_{|I|}$$

and $x_i^* = e_I(G^\infty)\cdot x_I(T')$. A similar discussion gives that $x_j^* = e_J(G^\infty)\cdot x_J(T')$. Since i and j asymptotically agree, we have $e_I(G^\infty)\cdot x_I(T') = e_J(G^\infty)\cdot x_J(T')$. This means that $x(T') \in H_{I,J}(G^\infty)$ and therefore

$$x^0 = P(G(0))^{-1}P(G(1))^{-1}\dots P(G(T'-1))^{-1}x(T') \in \bigcup_{Q\in\mathcal{Q}_{T'}} QH_{IJ}(G^\infty)$$

which leads to $x^0 \in X^0$. ■

In this section, we showed that asymptotic connectivity of agents implies asymptotic agreement and that under additional reasonable assumptions these are actually equivalent except for a set of vectors of initial opinions of Lebesgue measure 0. In other words, we can consider almost surely that the communities of agents correspond to the connected components of the graph G^∞.

In the following, under Assumptions 1, 2 and 4, we show that an algebraic characterization of communities can be given in terms of eigenvalues of the matrix associated to the graph of communities $P(G_\mathcal{C})$.

IV. Algebraic Characterization of Communities

Let $G' \in \mathcal{S}(G)$, let $I \subseteq V$ be a subset of agents such that no agent in I is connected to an agent in $V \setminus I$ in the graph G', then from Lemma 1, $P_I(G')$ is a row stochastic matrix. Let $\lambda_1(P_I(G')), \dots, \lambda_{|I|}(P_I(G'))$ denote the eigenvalues of $P_I(G')$ with $\lambda_1(P_I(G')) = 1$ and

$$|\lambda_1(P_I(G'))| \geq |\lambda_2(P_I(G'))| \geq \dots \geq |\lambda_{|C|}(P_I(G'))|.$$

Let $C \in \mathcal{C}$, then no agent in C is connected to an agent in $V \setminus C$ in the graph $G_{\mathcal{C}}$. The following theorem gives a characterization of the communities in terms of the eigenvalues $\lambda_2(P_C(G_{\mathcal{C}}))$ for $C \in \mathcal{C}$.

Theorem 3: Under Assumptions 1, 2 and 4, for almost all vectors of initial opinions x^0, for all communities $C \in \mathcal{C}$, such that $|C| \geq 2$,

$$|\lambda_2(P_C(G_{\mathcal{C}}))| < \rho.$$

Proof: Let us consider a spanning subgraph $G' \in \mathcal{S}(G)$, let $I = \{v_1, \dots, v_{|I|}\}$, with $|I| \geq 2$, be a connected component of G' then from Lemma 1, $P_I(G')$ is an aperiodic irreducible row stochastic matrix. Then, from the Perron-Frobenius Theorem, it follows that 1 is a simple eigenvalue of $P_I(G')$. Therefore, $\lambda_2(P_I(G')) \neq 1$. Let $f_I(G')$ be a left eigenvector of $P_I(G')$ associated to eigenvalue $\lambda_2(P_I(G'))$. Let us define the vector of $\mathbb{R}^n$, $c_I(G')$ whose coordinates are given by $c_{I,v_k}(G') = f_{I,k}(G')$ if $v_k \in I$ and $c_{I,k}(G') = 0$ if $k \in V \setminus I$. We define the $(n-1)$-dimensional subspace of $\mathbb{R}^n$:

$$H_I(G') = \{x \in \mathbb{R}^n |\ c_I(G') \cdot x = 0\}.$$

Finally, let us define the subset of $\mathbb{R}^n$:

$$Y^0 = \bigcup_{t \in \mathbb{N}} \left(\bigcup_{G' \in \mathcal{S}(G)} \left(\bigcup_{I \in \mathcal{K}(G'),\, |I| \geq 2} \left(\bigcup_{Q \in \mathcal{Q}_t} Q H_I(G') \right) \right) \right).$$

where $\mathcal{Q}_t$ is the set of matrices defined in (7). $\mathcal{S}(G)$ is a finite set and for all $G' \in \mathcal{S}(G)$, $\mathcal{K}(G')$ is a finite set. Moreover, for all $t \in \mathbb{N}$, $\mathcal{Q}_t$ is a finite set. Then, Y^0 is a countable union of $(n-1)$-dimensional subspaces of $\mathbb{R}^n$. Therefore Y^0 has Lebesgue measure 0.

Let X^0 be given as in equation (8), let $x^0 \in \mathbb{R}^n \setminus X^0$ be a vector of initial opinions. Let us assume there is a community $C \in \mathcal{C}$ with $|C| \geq 2$, such that $|\lambda_2(P_C(G_{\mathcal{C}}))| \geq \rho$. Let us show that necessarily, x^0 belongs to the set Y^0. First, since $x^0 \notin X^0$, we have from the proof of Theorem 2 that C is a connected component of $G^\infty = G_\mathcal{C}$. Therefore, from Lemma 1 and Lemma 3, for all $t \geq T'$, $x_C(t+1) = P_C(G_\mathcal{C})x_C(t)$ and $P_C(G_\mathcal{C})$ is an aperiodic irreducible row stochastic matrix. From the Perron-Frobenius Theorem, it follows that 1 is a simple eigenvalue of $P_C(G_\mathcal{C})$ and all other eigenvalues of $P_C(G_\mathcal{C})$ have modulus strictly smaller than 1. Let $e_C(G_\mathcal{C})$ be the left Perron eigenvector of $P_C(G_\mathcal{C})$:

$$e_C(G_\mathcal{C})^\top P_C(G_\mathcal{C}) = e_C(G_\mathcal{C})^\top \text{ and } e_C(G_\mathcal{C}) \cdot \mathbf{1}_{|C|} = 1$$

Then

$$\lim_{t \to +\infty} x_C(t) = x_C^* \quad \text{where} \quad x_C^* = (e_C(G_\mathcal{C}) \cdot x_C(T'))\mathbf{1}_{|C|}.$$

Let us remark that for all $t \geq T'$,

$$x_C(t+1) - x_C^* = P_C(G_\mathcal{C})(x_C(t) - x_C^*). \tag{9}$$

Let $f_C(G_\mathcal{C})$ be a left eigenvector of $P_C(G_\mathcal{C})$ associated to eigenvalue $\lambda_2(P_C(G_\mathcal{C}))$:

$$f_C(G_\mathcal{C})^\top P_C(G_\mathcal{C}) = \lambda_2(P_C(G_\mathcal{C})) f_C(G_\mathcal{C})^\top.$$

Then, it follows from equation (9) that for all $t \geq T'$,

$$f_C(G_\mathcal{C}) \cdot (x_C(t) - x_C^*) = f_C(G_\mathcal{C}) \cdot (x_C(T') - x_C^*) \lambda_2(P_C(G_\mathcal{C}))^{(t-T')}.$$

Therefore, by the Cauchy-Schwarz inequality, we have for all $t \geq T'$

$$\begin{aligned} \|x_C(t) - x_C^*\| &\geq \frac{|f_C(G_\mathcal{C}) \cdot (x_C(t) - x_C^*)|}{\|f_C(G_\mathcal{C})\|} \\ &\geq \frac{|f_C(G_\mathcal{C}) \cdot (x_C(T') - x_C^*)|}{\|f_C(G_\mathcal{C})\|} |\lambda_2(P_C(G_\mathcal{C}))|^{(t-T')}. \end{aligned}$$

Since we assumed $|\lambda_2(P_C(G_\mathcal{C}))| \geq \rho$, we have for all $t \geq T'$

$$\|x_C(t) - x_C^*\| \geq \frac{|f_C(G_\mathcal{C}) \cdot (x_C(T') - x_C^*)|}{\|f_C(G_\mathcal{C})\| \rho^{T'}} \rho^t. \tag{10}$$

Now, let us remark that it follows from Assumption 4 that for all $t \in \mathbb{N}$

$$\|x_C(t) - x_C^*\| \leq \sqrt{|C|} M \underline{\rho}^t. \tag{11}$$

Inequalities (10) and (11) give for all $t \geq T'$

$$\frac{|f_C(G_\mathcal{C}) \cdot (x_C(T') - x_C^*)|}{\|f_C(G_\mathcal{C})\| \rho^{T'}} \rho^t \leq \sqrt{|C|} M \underline{\rho}^t.$$

Since $\underline{\rho} < \rho$, the previous inequality holds for all $t \geq T'$ if and only if $|f_C(G_{\mathcal{C}}) \cdot (x_C(T') - x_C^*)| = 0$. Therefore, $f_C(G_{\mathcal{C}}) \cdot x_C(T') = f_C(G_{\mathcal{C}}) \cdot (x_C(T') - x_C^*) = 0$ which means that $x(T') \in H_C(G_{\mathcal{C}})$. Therefore,

$$x^0 = P(G(0))^{-1} P(G(1))^{-1} \dots P(G(T'-1))^{-1} x(T') \in \bigcup_{Q \in \mathcal{Q}_{T'}} Q H_C(G_{\mathcal{C}})$$

which leads to $x^0 \in Y^0$. Therefore, we have proved that for all vectors of initial opinions $x^0 \in \mathbb{R}^n \setminus (X^0 \cup Y^0)$, for all communities $C \in \mathcal{C}$ such that $|C| \geq 2$, $|\lambda_2(P_C(G_{\mathcal{C}}))| < \rho$. We conclude by remarking that $X^0 \cup Y^0$ is a set of Lebesgue measure 0. ■

In this section, we showed that the community structure $\mathcal{C}$ satisfies some properties related to the eigenvalues of the matrix $P_C(G_{\mathcal{C}})$, for $C \in \mathcal{C}$. In the following, we use this result to address the problem of community detection in graphs.

V. Application: Community Detection in Graphs

In this section, we propose to use the model of opinion dynamics with decaying confidence to address the problem of community detection in graphs.

A. The Community Detection Problem

In the usual sense, communities in a graph are groups of vertices such that the concentration of edges inside one communitiy is high and the concentration of edges between communities is comparatively low. Because of the increasing need of analysis tools for understanding complex networks in social sciences, biology, engineering or economics, the community detection problem has attracted a lot of attention in the recent years. The problem of community detection is however not rigorously defined mathematically. One reason is that community structures may appear at different scales in the graph: there can be communities inside communities. Another reason is that communities are not necessarily disjoint. A formalization of the community detection problem has been proposed in terms of optimization of a quality function called modularity [NG04]. However, it has been shown that this optimization problem is NP-complete [BDG$^+$08]. Therefore, approaches for community detection rely mostly on heuristic methods. We refer the reader to the excellent survey [For09] and the references therein for more details.

In this section, we formulate the community detection problem using a measure of connectivity of graphs given by the eigenvalues of their *normalized Laplacian matrix*. Let $G = (V, E)$ be an undirected graph with $V = \{1, \dots, n\}$, with $n \geq 2$. For a vertex $i \in V$, the degree $d_i(G)$ of i is the number of neighbors of i in G. The normalized Laplacian of the graph G is the matrix $L(G)$ given by

$$L_{ij}(G) = \begin{cases} 1 & \text{if } i = j \text{ and } d_i(G) \neq 0, \\ \frac{-1}{\sqrt{d_i(G)d_j(G)}} & \text{if } (i,j) \in E, \\ 0 & \text{otherwise.} \end{cases}$$

Let us review some of the properties of the normalized Laplacian matrix (see e.g. [Chu97]). $\sigma_1(L(G)) = 0$ is always an eigenvalue of $L(G)$, it is simple if and only if G is connected. All other eigenvalues are real and belong to the interval $[0, 2]$. The second smallest eigenvalue of the normalized Laplacian matrix is denoted $\sigma_2(L(G))$. It can serve as an algebraic measure of the connectivity: $\sigma_2(L(G)) = 0$ if the graph G has two distinct connected components, $\sigma_2(L(G)) = n/(n-1)$ if the graph is the complete graph (for all $i, j \in V$, $i \neq j$, $(i,j) \in E$), in the other cases $\sigma_2(L(G)) \in (0, 1]$.

Remark 4: The second smallest eigenvalue of the (non-normalized) Laplacian matrix is called algebraic connectivity of a graph. In this paper, we prefer to use the eigenvalues of the normalized Laplacian matrix because it is less sensitive to the size of the graph. For instance, if G is the complete graph then $\sigma_2(L(G)) = n/(n-1)$ whereas its algebraic connectivity is n.

Let $\mathcal{P}$ be a partition of the set of vertices V. For all $I \in \mathcal{P}$, with $|I| \geq 2$, $L(G_I)$ denotes the normalized Laplacian matrix of the graph $G_I = (I, E_I)$ consisting of the set of vertices I and of the set of edges of G between elements of I.

We now propose a formulation of the community detection problem:

Problem 1: Given a graph $G = (V, E)$ and a real number $\delta \in (0, 1]$, find a partition $\mathcal{P}$ of V such that for all $I \in \mathcal{P}$, such that $|I| \geq 2$, $\sigma_2(L(G_I)) > \delta$.

If $\sigma_2(L(G)) > \delta$, it is sufficient to choose the trivial partition $\mathcal{P} = \{V\}$. If $\delta \geq \sigma_2(L(G))$, then we want to find groups of vertices that are more densely connected than the global graph. This coincides with the notion of community. The larger δ the more densely connected the communities. This makes it possible to search for communities at different scales of the graph.

Let us remark that the problem is not really well posed as it may have several solutions. Indeed, the trivial partition $\mathcal{P} = \{\{1\}, \dots, \{n\}\}$ is always a solution of the proposed problem. In order to evaluate the quality of the computed partition, we propose to consider the following measure:

$$\mu(G,\mathcal{P}) = \begin{cases} \max_{I\in\mathcal{P}} \sigma_2(L(G_I)) & \text{if for all } I \in \mathcal{P},\ |I| \geq 2 \\ +\infty & \text{otherwise.} \end{cases}$$

Note that if $\mathcal{P}$ is a solution of Problem 1, then $\mu(G,\mathcal{P}) > \delta$. In order to solve the problem in an effective way, $\mu(G,\mathcal{P})$ should be as close of δ as possible.

In the following, we show how to solve Problem 1 using an opinion dynamics with decaying confidence model.

B. Opinion Dynamics for Community Detection

Let $\alpha \in (0, 1/2)$, we consider the opinion dynamics with decaying confidence model given by:

$$x_i(t+1) = \begin{cases} x_i(t) + \dfrac{\alpha}{|N_i(t)|} \displaystyle\sum_{j\in N_i(t)} (x_j(t) - x_i(t)) & \text{if } N_i(t) \neq \emptyset \\ x_i(t) & \text{if } N_i(t) = \emptyset \end{cases} \tag{12}$$

where $N_i(t)$ is given by equation (3). It is straightforward to check that this model is a particular case of the model given by equations (1) and (2) and that Assumption 1 holds. Moreover, since $\alpha \in (0, 1/2)$ it follows that for all $i \in V$, $t \in \mathbb{N}$, $p_{ii}(t) > 1/2$. Therefore the matrix $P(t)$ is strictly diagonally dominant and hence it is invertible. Also, $P(t) = P(G(t))$, where for a subgraph G', $P(G') = Id - \alpha Q(G')$ where Id is the identity matrix and

$$Q_{ij}(G') = \begin{cases} 1 & \text{if } i = j \text{ and } d_i(G') \neq 0, \\ \frac{-1}{d_i(G')} & \text{if } (i,j) \in E', \\ 0 & \text{otherwise.} \end{cases} \tag{13}$$

where $d_i(G')$ denotes the degree of i in the graph G'. Therefore, Assumption 2 holds as well. Let us remark that the matrix $P(t)$ is generally not average preserving and therefore Assumption 3 does not hold.

Before stating the main result of this section, we need to prove the following lemma :

Lemma 4: Let $\mathcal{P}$ be a partition of V, $I \in \mathcal{P}$ such that $|I| \geq 2$. Then, λ is an eigenvalue of $P_I(G_{\mathcal{P}})$ if and only if $\sigma = (1-\lambda)/\alpha$ is an eigenvalue of $L(G_I)$.

Proof: First, let us remark that $P_I(G_{\mathcal{P}}) = Id - \alpha Q(G_I)$ where $Q(G_I)$ is defined as in equation (13). Then, let us introduce the matrices $R(G_I)$ and $D(G_I)$ defined by

$$R_{ij}(G_I) = \begin{cases} \frac{1}{\sqrt{d_i(G_I)}} & \text{if } i = j \text{ and } d_i(G_I) \neq 0, \\ \frac{-1}{d_i(G_I)\sqrt{d_j(G_I)}} & \text{if } (i,j) \in E_I, \\ 0 & \text{otherwise.} \end{cases}$$

and

$$D_{ij}(G_I) = \begin{cases} \sqrt{d_i(G_I)} & \text{if } i = j, \\ 0 & \text{otherwise.} \end{cases}$$

Then, let us remark that $L(G_I) = D(G_I)R(G_I)$ and $Q(G_I) = R(G_I)D(G_I)$. It follows that $L(G_I)$ and $Q(G_I)$ have the same eigenvalues. The stated result is then obtained from the fact that the matrix $Q(G_I) = (Id - P_I(G_{\mathcal{P}}))/\alpha$. ■

We now state the main result of the section which is a direct consequence of Theorem 3 and Lemma 4:

Corollary 1: Let $\rho = 1 - \alpha\delta$, under Assumption 4, for almost all vectors of initial opinions x^0, the set of communities $\mathcal{C}$ obtained by the opinion dynamics model (12) is a solution to Problem 1.

In the next section, we propose to evaluate experimentally the validity of our approach.

C. Case Study 1: Zachary Karate Club

We propose to evaluate our approach on a standard benchmark for community detection: the karate club network initially studied by Zachary in [Zac73]. This is a social network with 34 agents shown on the top left part of Figure 1. The original study shows the existence of two communities represented on the figure by squares and triangles.

We propose to use our opinion dynamics model (12) to uncover the community structure of this network. We chose 4 different values for δ. The parameters of the model where chosen as follows: $\alpha = 0.1$, $R = 1$ and $\rho = 1 - \alpha\delta$. For each different value of δ, the model was simulated for 100 different vectors of initial opinions chosen randomly in $[0,1]^{34}$. Simulations were performed as long as enabled by floating point arithmetics.

The experimental results are reported on Figure 1 and Table I. For each different value of δ, we indicate the different partitions in communities obtained after running the opinion

δ	$\lvert\mathcal{C}\rvert$	$\min_{C\in\mathcal{C},\lvert C\rvert\geq 2} \sigma_2(L(G_C))$	$\mu(G,\mathcal{P})$	Occurrences
0.1	1	0.132	0.132	100
0.2	2	0.250	0.363	100
0.3	3	0.334	0.667	67
0.3	3	0.363	0.678	31
0.3	4	0.336	2.000	2
0.4	4	0.566	0.667	92
0.4	3	0.334	0.667	6
0.4	5	0.566	2.000	1
0.4	5	0.566	$+\infty$	1

TABLE I

QUANTITATIVE PROPERTIES OF THE PARTITIONS OF THE KARATE CLUB NETWORK OBTAINED BY THE OPINION DYNAMICS MODEL (100 DIFFERENT VECTORS OF INITIAL OPINIONS FOR EACH VALUE OF δ.

dynamics model. For each partition $\mathcal{C}$, we give the number of communities in the partition, the minimal value of $\sigma_2(L(G_C))$ for $C \in \mathcal{C}$, this value being greater than δ indicates that Problem 1 has been solved. We computed the measure $\mu(G,\mathcal{P})$ in order to evaluate the quality of the obtained partition. We also indicate the number of times that each partition occurred over the 100 simulations of the opinion dynamics model.

We can check in Table I that all the computed partitions are solutions of Problem 1 with the exception of the partition reported in line 7 that occurred 6 times over 100 simulations for $\delta = 0.4$. This partition is actually the same than the one reported in line 3. A closer inspection of the simulation data shows that in the case of $\delta = 0.4$, this partition would have been further subdivided if the model had been simulated beyond what is allowed by floating point arithmetics.

Let us remark that in general the computed partition depends on the initial vector of opinions, this is the case for $\delta = 0.3$ and $\delta = 0.4$. However, it is interesting to note that the partitions that are obtained the most frequently have comparatively a good quality measure $\mu(G,\mathcal{P})$. This shows that our approach not only allows to solve Problem 1 but solves it in an effective way.

In Figure 1, we represented the graphs of communities $G_{\mathcal{C}}$ that are the most frequently obtained for the different values of δ. It is interesting to remark that for $\delta = 0.2$ we almost obtained the communities that were reported in the original study [Zac73]. Only one agent has been classified

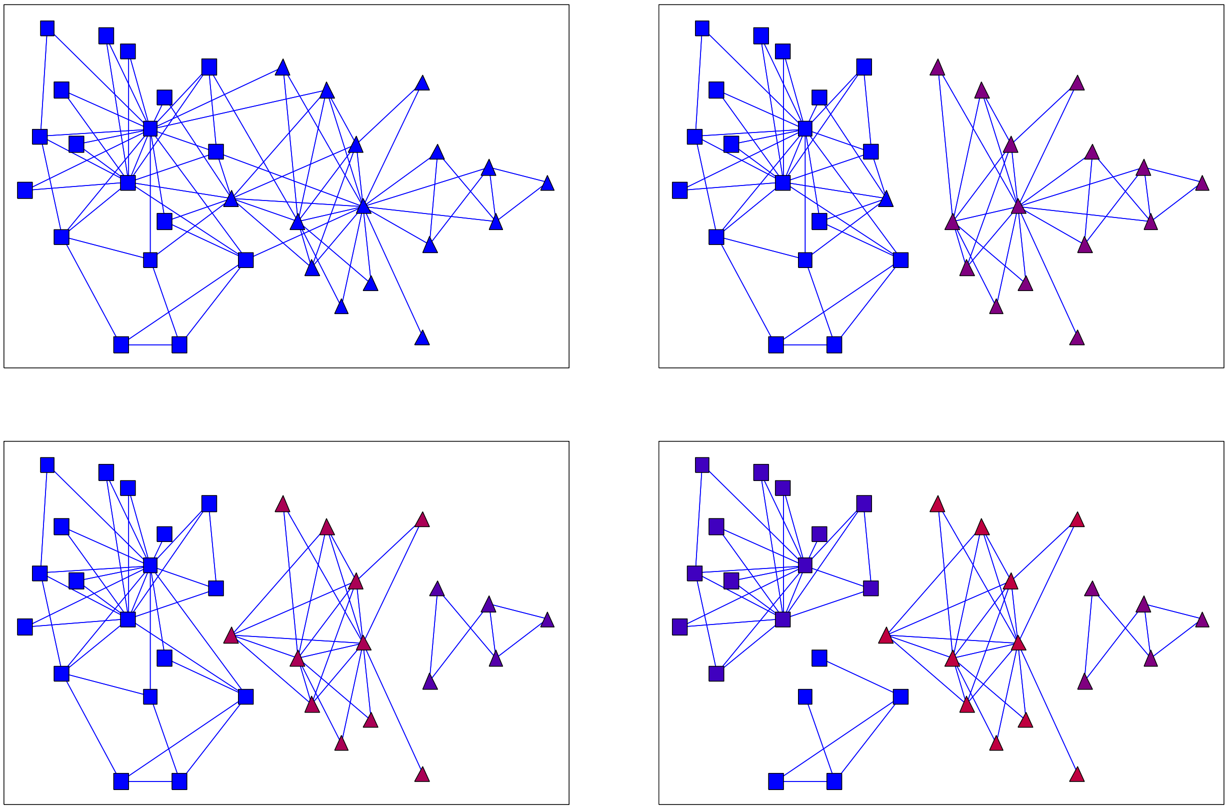

Fig. 1. Graphs $G_{\mathcal{C}}$ for the most frequently obtained partition of the karate club network for $\delta = 0.1$ (top left), $\delta = 0.2$ (top right), $\delta = 0.3$ (bottom left), $\delta = 0.4$ (bottom right).

differently. One may argue that this agent has originally 4 neighbors in each community so it could be classified in one or the other. It is also interesting to see that our approach allows us to search for communities at different scales of the graph. When δ increases, the communities become smaller but more densely connected.

D. Case Study 2: Books on American Politics

We propose to use our approach on a larger example consisting of a network of 105 books on politics [NG04], initially compiled by V. Krebs (unpublished, see `www.orgnet.com`). In this network, each vertex represents a book on American politics bought from Amazon.com. An edge between two vertices means that these books are frequently purchased by the same buyer. The network is presented on the top left part of Figure 2 where the shape of the vertices represent the political alignment of the book (liberal, conservative, centrist).

We used our opinion dynamics model (12) to uncover the community structure of this network. We chose 3 different values for δ. The parameters of the model are the same than in the previous example: $\alpha = 0.1$, $R = 1$ and $\rho = 1 - \alpha\delta$. For each different value of δ, the model was simulated for 100 different vectors of initial opinions chosen randomly in $[0, 1]^{105}$. Simulations were performed as long as enabled by floating point arithmetics. The experimental results are reported on Figure 2 and Table II.

Let us remark that all the computed partitions are solutions to the Problem 1. Also, for the same value of δ, the quality measure is almost the same for all partitions. Actually, all the partitions obtained for the same value of δ are almost the same. In Figure 2, we represented the graphs of communities $G_{\mathcal{C}}$ that are the most frequently obtained for the different values of δ. Let us remark that even though the information on the political alignment of the books is not used by the algorithm, our approach allows to uncover this information. Indeed, for $\delta = 0.1$, we obtain 2 communities that are essentially liberal and conservative. For $\delta = 0.2$, we then obtain 4 communities: liberal, conservative, centrist-liberal, centrist-conservative.

δ	$\lvert\mathcal{C}\rvert$	$\min_{C\in\mathcal{C}, \lvert C\rvert\geq 2} \sigma_2(L(G_C))$	$\mu(G, \mathcal{P})$	Occurrences
0.1	2	0.134	0.187	99
0.1	2	0.129	0.182	1
0.15	3	0.187	0.329	89
0.15	3	0.182	0.322	11
0.2	4	0.269	0.791	71
0.2	4	0.266	0.810	18
0.2	4	0.269	0.810	5
0.2	4	0.268	0.791	4
0.2	4	0.269	0.750	2

TABLE II

QUANTITATIVE PROPERTIES OF THE PARTITIONS OF THE BOOKS NETWORK OBTAINED BY THE OPINION DYNAMICS MODEL (100 DIFFERENT VECTORS OF INITIAL OPINIONS FOR EACH VALUE OF PARAMETER δ.

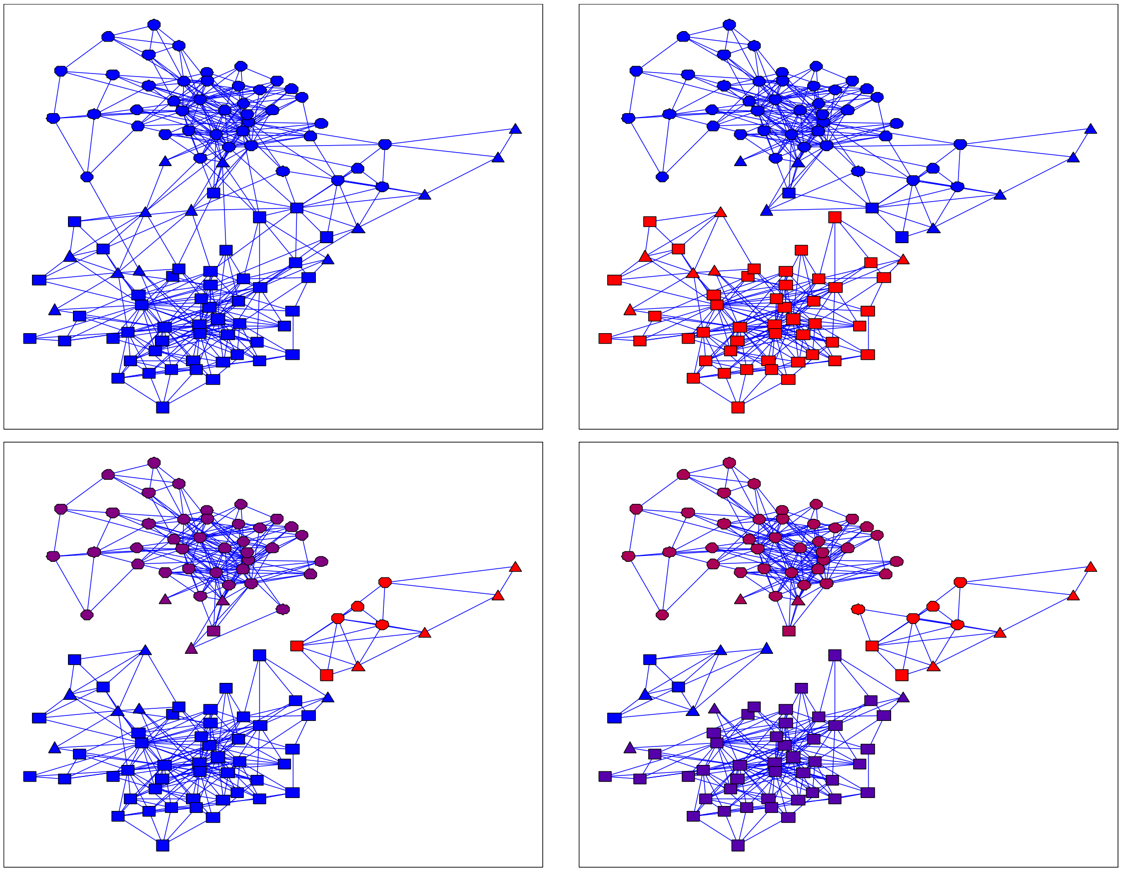

Fig. 2. Graphs $G_{\mathcal{C}}$ for the most frequently obtained partition of the books network: initial graph (top left), $\delta = 0.1$ (top right), $\delta = 0.15$ (bottom left), $\delta = 0.2$ (bottom right). Shapes represent political alignment of the books: circles are liberal, squares are conservative, triangles are centrist.

VI. Conclusion and Future Work

In this paper, we introduced and analyzed a model of opinion dynamics with decaying confidence where agents may only reach local agreements organizing themselves in communities. Under suitable assumptions, we have shown that these communities correspond to asymptotically connected component of the network. We have also provided an algebraic characterization of communities in terms of eigenvalues of the matrix defining the collective dynamics. To complete the analysis of our model, future work should focus on relaxing Assumption 4 by studying the model behavior when there is an agent $i \in V$ that approaches its limit value at a rate exactly ρ:

$$\limsup_{t \to +\infty} \frac{1}{t} \log(|x_i(t) - x_i^*|) = \log(\rho).$$

In the last part of the paper, we have applied our opinion dynamics model to address the problem of community detection in graphs. We believe that this new approach offers an appealing interpretation of community detection: communities are sets of agents that succeed to reach an agreement under some convergence rate constraint. We have shown on two examples that this approach is not only appealing but is also effective. In the future, we shall work on a distributed implementation of our approach. Let us remark that this should be feasible since our approach is by nature based on distributed computations. Then, we shall use our approach to analyze a number of networks including large scale networks.

REFERENCES

[AB08] D. Angeli and P. A. Bliman. Tight estimates for convergence of some non-stationary consensus algorithms. *Systems and Control Letters*, 57(12):996–1004, 2008.

[BDG+08] U. Brandes, D. Delling, M. Gaertler, R. Görke, M. Hoefer, Z. Nikoloski, and D. Wagner. On modularity clustering. *IEEE Trans. on Knowledge and Data Engineering*, 20(2):172–188, 2008.

[BHOT05] V. D. Blondel, J. M. Hendrickx, A. Olshevsky, and J.N. Tsitsiklis. Convergence in multiagent coordination, consensus, and flocking. In *Proc. IEEE Conf. on Decision and Control*, pages 2996–3000, 2005.

[BHT06] V. D. Blondel, J. M. Hendrickx, , and J.N. Tsitsiklis. On the 2R conjecture for multi-agent systems. In *Proc. European Control Conference*, pages 874–881, 2006.

[Chu97] F. Chung. *Spectral Graph Theory*. American Mathematical Society, 1997.

[For09] S. Fortunato. Community detection in graphs. Technical report, 2009. arXiv:0906.0612v1.

[HK02] R. Hegselmann and U. Krause. Opinion dynamics and bounded confidence models, analysis, and simulation. *Journal of Artificial Societies and Social Simulation*, 5(3), 2002.

[JLM03] A. Jadbabaie, J. Lin, and A. S. Morse. Coordination of groups of mobile autonomous agents using nearest neighbor rules. *IEEE Trans. on Automatic Control*, 48(6):988–1001, 2003.

[Kra97] U. Krause. Soziale Dynamiken mit vielen Interakteuren. Eine Problemskizze. In *Modellierung und Simulation von Dynamiken mit vielen interagierenden Akteuren*, pages 37–51, 1997.

[Mor05] L. Moreau. Stability of multiagent systems with time-dependent communication links. *IEEE Trans. on Automatic Control*, 50(2):169–182, 2005.

[NG04] M. E. J. Newman and M. Girvan. Finding and evaluating community structure in networks. *Phys. Rev. E*, 69(2):026113.1–15, 2004.

[OSFM07] R. Olfati-Saber, J. A. Fax, and R. M. Murray. Consensus and cooperation in networked multi-agent systems. *Proceedings of the IEEE*, 95(1):215–233, 2007.

[OT09] A. Oshelvsky and J. N. Tsitsiklis. Convergence speeds in distributed consensus and averaging. *SIAM J. Control and Optimization*, 48(1):33–55, 2009.

[RB05] W. Ren and R. W. Beard. Consensus seeking in multiagent systems under dynamically changing interaction topologies. *IEEE Trans. on Automatic Control*, 50(5):655–661, 2005.

[Sen81] E. Seneta. *Non-Negative Matrices and Markov Chains*. Springer-Verlag, 1981.

[Zac73] W. W. Zachary. An information flow model for conflict and fission in small groups. *Journal of anthropological research*, 33(4):1977, 452-473.

[ZW09] J. Zhou and Q. Wang. Convergence speed in distributed consensus over dynamically switching random networks. *Automatica*, 45(6):1455–1461, 2009.