THE INSTITUTE FOR SYSTEMS RESEARCH

ISR TECHNICAL REPORT 2010-4

The asymptotic consensus problem on convex metric spaces

Ion Matei and John S. Baras

The Institute for Systems Research

A. JAMES CLARK SCHOOL OF ENGINEERING

ISR develops, applies and teaches advanced methodologies of design and analysis to solve complex, hierarchical, heterogeneous and dynamic problems of engineering technology and systems for industry and government.

ISR is a permanent institute of the University of Maryland, within the A. James Clark School of Engineering. It is a graduated National Science Foundation Engineering Research Center.

www.isr.umd.edu

The asymptotic consensus problem on convex metric spaces

Ion Matei*, John S. Baras†

*Department of Electrical and Computer Engineering t

University of Maryland, College Park, MD 20742

Email: imatei@umd.edu

†Department of Electrical and Computer Engineering

University of Maryland, College Park, MD 20742

Email: baras@umd.edu

Abstract

We consider the consensus problem of a group of dynamic agents whose communication network is modeled by a directed time-varying graph. In this paper we generalize the asymptotic consensus problem to convex metric spaces. A convex metric space is a metric space on which we define a convex structure. Using this convex structure we define convex sets and in particular the convex hull of a (finite) set. Under minimal connectivity assumptions, we show that if at each iteration an agent updates its state by choosing a point from a particular subset of the convex hull generated by the agent's current state and the states of his/her neighbors, then the asymptotic agreement is achieved. In addition, we give bounds on the distance between the consensus point(s) and the initial values of the agents. As application example, we use this framework to introduce an iterative algorithm for reaching consensus of opinion. In this example, the agents take values in the space of discrete random variable on which we define an appropriate metric and convex structure. For this particular convex metric space we provide a more detail analysis of the convex hull generated by a finite set points. In addition we give some numerical simulations of the consensus of opinion algorithm.

I. Introduction

A consensus problem consists of a group of dynamic agents who seek to agree upon certain quantities of interest by exchanging information among them according to a set of rules. This problem can model many phenomena involving information exchange between agents such as cooperative control of vehicles, formation control, flocking, synchronization, parallel computing, etc. Distributed computation over networks has a long history in control theory starting with the work of Borkar and Varaiya [1], Tsitsikils, Bertsekas and Athans [23], [24] on asynchronous agreement problems and parallel computing. A theoretical framework for solving consensus problems was introduced by Olfati-Saber and Murray in [16], [17], while Jadbabaie et al. studied alignment problems [5] for reaching an agreement. Relevant extensions of the consensus problem were done by Ren and Beard [14], by Moreau in [9] or, more recently, by Nedic and Ozdaglar in [12], [11].

Typically agents are connected via a network that changes with time due to link failures, packet drops, node failure, etc. Such variations in topology can happen randomly which motivates the investigation of consensus problems under a stochastic framework. Hatano and Mesbahi consider in [6] an agreement problem over random information networks, where the existence of an information channel between a pair of elements at each time instance is probabilistic and independent of other channels. In [13], Porfiri and Stilwell provide sufficient conditions for reaching consensus almost surely in the case of a discrete linear system, where the communication flow is given by a directed graph derived from a random graph process, independent of other time instances. Under a similar model of the communication topology, Tahbaz-Salehi and Jadbabaie give necessary and sufficient conditions for almost sure convergence to consensus in [21], while in [22], the authors extend the applicability of their necessary and sufficient conditions to strictly stationary ergodic random graphs. Extensions to the case where the random graph modeling the communication among agents is a Markovian random process are given in [7], [8].

A convex metric space is a metric space on which we define a convex structure. The main goal of this paper is to generalize the asymptotic consensus problem to the more general case of convex metric spaces and emphasize the fundamental role of convexity and in particular of the convex hull of a finite set of points. Tsitsiklis showed in [23] that, under some minimal connectivity assumptions on the communication network, if an agent updates its value by choosing a point from the (interior) of the convex hull of its current value and the current values of its neighbors, then asymptotic convergence to consensus is achieved. We will show that this idea extends naturally to the more general case of convex metric spaces.

Our main contributions are as follows. *First*, after citing relevant results concerning convex metric spaces, we study the properties of the distance between two points belonging to two, possible overlapping convex hulls of two finite sets of points. These properties will prove to be crucial in proving the convergence of the agreement algorithm. *Second*, we provide a dynamic equation for an upper bound of the vector of distances between the current values of the agents. We show that the agents asymptotically reach agreement, by showing that this upper bound asymptotically converges to zero. *Third*, we characterize the agreement point(s) compared to the initial values of the agents, be giving upper bounds on the distance between the agreement point(s) and the initial values in terms of the distances between the initial values of the agents. *Forth*, we emphasize the relevance of our framework, by providing an application under the form of a consensus of opinion algorithm. For this example we define a particular convex metric space and we study in more depth the properties of the convex hull of a finite set of points.

The paper is organized as follows. Section II introduces the main concepts related to the convex metric spaces and focuses in particular on the convex hull of a finite set. Section III formulates the problem and states our main theorem. Section IV gives the proof of our main theorem together with some auxiliary results. In Section VI we present an application of our main result by providing an iterative algorithm for reaching consensus of opinion.

Some basic notations: Given $W \in \mathbb{R}^{n\times n}$ by $[W]_{ij}$ we refer to the (i,j) element of the matrix. The *underlying graph* of W is a graph of order n for which every edge corresponds to a non-zero, non-diagonal entry of W. We will denote by $\mathbb{1}_{\{A\}}$ the indicator function of event A. Given some space $\mathcal{X}$ we denote by $\mathcal{P}(\mathcal{X})$ the set of all subsets of $\mathcal{X}$.

II. Convex Metric Spaces

The first part of this section deals with a set of definitions and basic results about convex metric spaces. The second part focuses on the convex hull of a finite set in convex metric spaces.

A. Definitions and Results on Convex Metric Spaces

For more details about the following definitions and results the reader is invited to consult [25],[26].

Definition 2.1: Let $(\mathcal{X},d)$ be a metric space and let $x,y,z \in \mathcal{X}$. We say that z *is between* x and y if $d(x,z)+d(z,y) = d(x,y)$. For any two points $x,y \in \mathcal{X}$, the set

$$\{z \in \mathcal{X} \mid d(x,z)+d(z,y) = d(x,y)\}$$

is called *metric segment* and is denoted by $[x,y]$.

Definition 2.2: Let $(\mathcal{X},d)$ be a metric space. A mapping $\psi:\mathcal{X}\times\mathcal{X}\times[0,1]\to\mathcal{X}$ is said to be a *convex structure* on $\mathcal{X}$ if

$$d(u,\psi(x,y,\lambda))\leq\lambda d(u,x)+(1-\lambda)d(u,y),\ \forall x,y,u\in\mathcal{X}\text{ and }\forall\lambda\in[0,1]. \tag{1}$$

Definition 2.3: The metric space $(\mathcal{X},d)$ together with the convex structure ψ is called *convex metric space*.

A Banach space and each of its subsets are convex metric space. There are examples of convex metric spaces not embedded in any Banach space. The following two examples are taken from [26].

Example 2.1: Let I be the unit interval $[0,1]$ and X be the family of closed intervals $[a_i,b_i]$ such that $0\leq a_i\leq b_i\leq 1$. For $I_i=[a_i,b_i]$, $I_j=[a_j,b_j]$ and $\lambda\in I$, we define a mapping ψ by $\psi(I_i,I_j,\lambda)=[\lambda a_i+(1-\lambda)a_j,\lambda b_i+(1-\lambda)b_j]$ and define a metric d in X by the Hausdorff distance, i.e.

$$d(I_i,I_j)=\sup_{a\in I}\{|\inf_{b\in I_i}\{|a-b|\}-\inf_{c\in I_j}\{|a-c|\}|\}.$$

Example 2.2: We consider a linear space L which is also a metric space with the following properties:

(a) For $x,y\in L$, $d(x,y)=d(x-y,0)$;

(b) For $x,y\in L$, and $\lambda\in[0,1]$,

$$d(\lambda x+(1-\lambda)y,0)\leq\lambda d(x,0)+(1-\lambda)d(y,0).$$

Definition 2.4: Let $\mathcal{X}$ be a convex metric space. A nonempty subset $K\subset\mathcal{X}$ is said to be *convex* if $\psi(x,y,\lambda)\in K$, $\forall x,y\in K$ and $\forall\lambda\in[0,1]$.

We define the set valued mapping $\tilde{\psi}:\mathcal{P}(\mathcal{X})\to\mathcal{P}(\mathcal{X})$ as

$$\tilde{\psi}(A)\triangleq\{\psi(x,y,\lambda)\mid\forall x,y\in A,\forall\lambda\in[0,1]\}, \tag{2}$$

where A is an arbitrary set in $\mathcal{X}$.

In [26] it is shown that, in a convex metric space, an arbitrary intersection of convex sets is also convex and therefore the next definition makes sense.

Definition 2.5: The *convex hull* of the set $A\subset\mathcal{X}$ is the intersection of all convex sets in $\mathcal{X}$ containing A and is denoted by $conv(A)$.

Another characterization of the convex hull of a set in $\mathcal{X}$ is given in what follows. By defining $A_m \triangleq \tilde{\psi}(A_{m-1})$ with $A_0 = A$ for some $A \subset \mathcal{X}$, it is discussed in [25] that the set sequence $\{A_m\}_{m\geq 0}$ is increasing and $\limsup A_m$ exits, and $\limsup A_m = \liminf A_m = \lim A_m = \bigcup_{m=0}^{\infty} A_m$.

Proposition 2.1 ([25]): Let $\mathcal{X}$ be a convex metric space. The convex hull of a set $A \subset \mathcal{X}$ is given by

$$conv(A) = \lim A_m = \bigcup_{m=0}^{\infty} A_m. \tag{3}$$

B. On the convex hull of a finite set

For a positive integer n, let $A = \{x_1, \ldots, x_n\}$ be a finite set in $\mathcal{X}$ with convex hull $conv(A)$ and let z belong to $conv(A)$. By Proposition 2.1 it follows that there exits a positive integer m such that $z \in A_m$. But since $A_m = \tilde{\psi}(A_{m-1})$ it follows that there exits $z_1, z_2 \in A_{m-1}$ and $\lambda_{(1,2)} \in [0,1]$ such that $z = \psi(z_1, z_2, \lambda_{(1,2)})$. Similarly, there exits $z_3, z_4, z_5, z_6 \in A_{m-2}$ and $\lambda_{(3,4)}, \lambda_{5,6} \in [0,1]$ such that $z_1 = \psi(z_3, z_4, \lambda_{(3,4)})$ and $z_2 = \psi(z_5, z_6, \lambda_{(5,6)})$. By further decomposing z_3,z_4,z_5 and z_6 and their followers until they are expressed as functions of elements of A and using a graph theory terminology, we note the z is the root of a weighted binary tree with leaves belonging to the set A. Each node η (except the leaves) has two children η_1 and η_2, and are related through the operator ψ in the sense $\eta = \psi(\eta_1, \eta_2, \lambda)$ for some $\lambda \in [0,1]$. The weights of the edges connecting η with η_1 and η_2 are given by λ and $1-\lambda$ respectively.

From the above discussion we note that for any point $z \in conv(A)$ there exits a non-negative integer m such that z is the root of a binary tree of height m, and has as leaves elements of A. The binary tree rooted at z may or may not be a *perfect binary tree*. That is because on some branches of the tree the points in A are reached faster then on others. Let n_i denote the number of times x_i appears as a leaf node, with $\sum_{i=1}^{n} n_i \leq 2^m$ and let m_{i_l} be the length of the i_l^{th} path from the root z to the node x_i, for $l = 1 \ldots n_i$. We formally describe the paths from the root z to x_i as the set

$$P_{z,x_i} \triangleq \left\{ \left(\{y_{i_l,j}\}_{j=0}^{m_{i_l}}, \{\lambda_{i_l,j}\}_{j=1}^{m_{i_l}} \right) \mid l = 1 \ldots n_i \right\}, \tag{4}$$

where $\{y_{i_l j}\}_{j=0}^{m_{i_l}}$ is the set of points forming the $i_l{}^{th}$ path, with $y_{i_l,0} = z$ and $y_{i_l,m_{i_l}} = x_i$ and where $\{\lambda_{i_l,j}\}_{j=1}^{m_{i_l}}$ is the set of weights corresponding to the edges along the paths, in particular $\lambda_{i_l,j}$ being the weight of the edge $(y_{i_l,j-1}, y_{i_l,j})$. We define the aggregate weight of the paths from root z to node x_i as

$$\mathcal{W}(P_{z,x_i}) \triangleq \sum_{l=1}^{n_i} \prod_{j=1}^{m_{i_l}} \lambda_{i_l,j}. \tag{5}$$

It is not difficult to note that all the aggregate weights of the a paths from the root z to the leaves $\{x_1, \ldots, x_n\}$ sum up to one, i.e.

$$\sum_{i=1}^{n} \mathcal{W}(P_{z,x_i}) = 1.$$

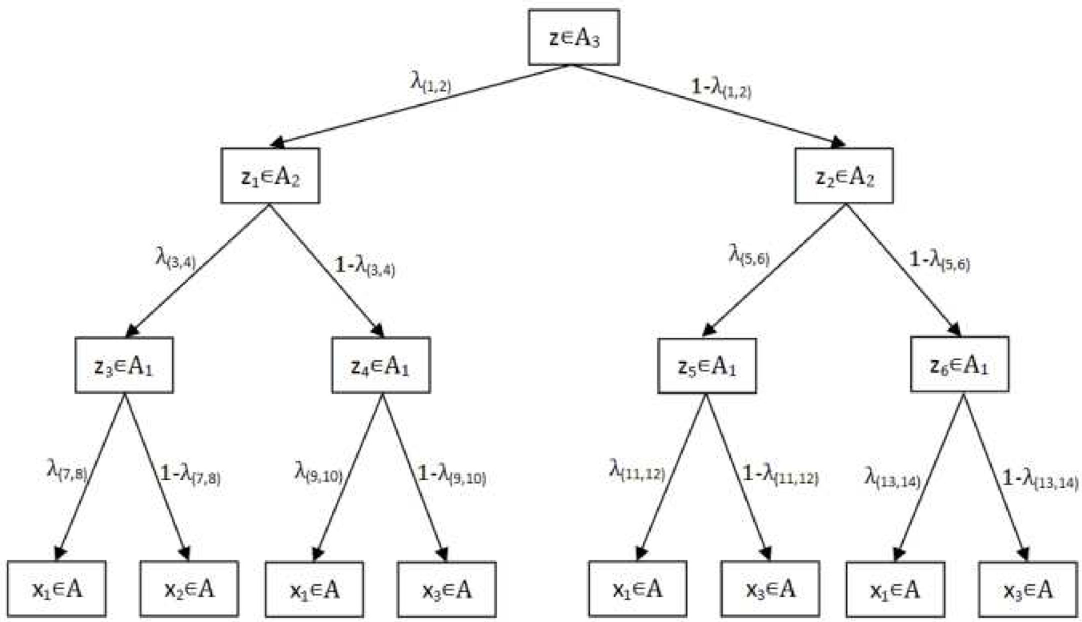

Fig. 1. The decomposition of a point $z \in A_3$ with $A = \{x_1, x_2, x_3\}$

Example 2.3: Figure 1 shows a binary tree corresponding to a point $z \in A_3$, where $A = \{x_1, x_2, x_3\}$. For this particular example, the paths from to root z to the leaves x_i are given by

$$P_{z,x_1} = \{(\{z, z_1, z_3, x_1\}, \{\lambda_{(1,2)}, \lambda_{(3,4)}, \lambda_{(7,8)}\}), (\{z, z_1, z_4, x_1\}, \{\lambda_{(1,2)}, (1-\lambda_{(3,4)}), \lambda_{(9,10)}\}),$$

$$\Big(\{z, z_2, z_5, x_1\}, \{(1-\lambda_{(1,2)}), \lambda_{(5,6)}, \lambda_{(11,12)}\}\Big), \Big(\{z, z_2, z_6, x_1\}, \{(1-\lambda_{(1,2)}), (1-\lambda_{(5,6)}), \lambda_{(13,14)}\}\Big)\Big\},$$

$$P_{z,x_2} = \{(\{z, z_1, z_3, x_2\}, \{\lambda_{(1,2)}, \lambda_{(3,4)}, (1-\lambda_{(7,8)})\})\}$$

$$P_{z,x_3} = \Big\{(\{z, z_1, z_4, x_3\}, \{\lambda_{(1,2)}, (1-\lambda_{(3,4)}), (1-\lambda_{(9,10)})\}), \Big(\{z, z_2, z_5, x_3\}, \{(1-\lambda_{(1,2)}), \lambda_{(5,6)}, (1-\lambda_{(11,12)})\}\Big),$$

$$\Big(\{z, z_2, z_6, x_3\}, \{(1-\lambda_{(1,2)}), (1-\lambda_{(5,6)}), (1-\lambda_{(13,14)})\}\Big)\Big\}$$

and the path weights are

$$\mathcal{W}(P_{z,x_1}) = \lambda_{(1,2)}\lambda_{(3,4)}\lambda_{(7,8)} + \lambda_{(1,2)}(1-\lambda_{(3,4)})\lambda_{(9,10)} + (1-\lambda_{(1,2)}), \lambda_{(5,6)}, \lambda_{(11,12)},$$

$$\mathcal{W}(P_{z,x_2}) = \lambda_{(1,2)}\lambda_{(3,4)}(1-\lambda_{(7,8)}),$$

$$\mathcal{W}(P_{z,x_3}) = \lambda_{(1,2)}(1-\lambda_{(3,4)})(1-\lambda_{(9,10)}) + (1-\lambda_{(1,2)})\lambda_{(5,6)}(1-\lambda_{(11,12)}) + (1-\lambda_{(1,2)})(1-\lambda_{(5,6)})(1-\lambda_{(13,14)}).$$

Definition 2.6: We say that a point z belongs to the *interior* of $conv(A)$ and we denote this by $z \in int(conv(A))$, if all elements of A belong to the set of leaves of the binary tree rooted at z.

Definition 2.7: Given a small enough positive scalar $\underline{\lambda} < 1$ we define the following sub-set of $int(conv(A))$ consisting in all points in $int(conv(A))$ whose aggregate weights are lower bounded by $\underline{\lambda}$, i.e.

$$C_{\underline{\lambda}}(A) \triangleq \{z \mid z \in int(conv(A)), \mathcal{W}(P_{z,x_i}) \geq \underline{\lambda}, \ \forall x_i \in A\}. \tag{6}$$

Remark 2.1: We can iteratively generate points for which we can make sure that they belong to the interior of the convex hull of a finite set $A = \{x_1, \ldots, x_n\}$. Given a set of positive scalars $\{\lambda_1, \ldots, \lambda_{n-1}\} \in (0,1)$, consider the iteration

$$y_{i+1} = \psi(y_i, x_{i+1}, \lambda_i) \text{ for } i = 1 \ldots n-1 \text{ with } y_1 = x_1. \tag{7}$$

It is not difficult to note that y_n is guaranteed to belong to the interior of $conv(A)$. In addition, if we impose the condition

$$\lambda_i \geq \sqrt[n-1]{\underline{\lambda}} \text{ for } i = 1 \ldots n-1, \tag{8}$$

then $y_n \in C_{\underline{\lambda}}(A)$

The next result characterizes the distance between two points $x, y \in \mathcal{X}$ belonging to the convex hulls of two (possible overlapping) finite sets X and Y.

Proposition 2.2: Let n_x and n_y be two positive integers, let $X = \{x_1, \ldots, x_{n_x}\}$ and $Y = \{y_1, \ldots, y_{n_y}\}$ be two finite sets on $\mathcal{X}$ and let $\underline{\lambda} < 1$ be a positive scalar small enough.

(a) If $x \in int(conv(X))$ and $y \in \mathcal{X}$ then

$$d(x,y) \leq \sum_{i=1}^{n_x} \lambda_i d(x_i, y), \tag{9}$$

for some $\lambda_i > 0$ with $\sum_{i=1}^{n_x} \lambda_i = 1$.

(b) If $x \in int(conv(X))$ and $y \in int(conv(Y))$ then

$$d(x,y) \leq \sum_{i=1}^{n_x} \sum_{j=1}^{n_y} \lambda_{ij} d(x_i, y_j), \tag{10}$$

for some $\lambda_{ij} > 0$ with $\sum_{i=1}^{n_x}\sum_{j=1}^{n_y}\lambda_{ij} = 1$.

(c) If $x \in \mathcal{F}_{\underline{\lambda}}(X)$, $y \in \mathcal{F}_{\underline{\lambda}}(Y)$, then

$$\lambda_i \geq \underline{\lambda} \text{ and } \lambda_{ij} \geq \underline{\lambda}^2, \ \forall \ i, j, \tag{11}$$

where λ_i and λ_{ij} where introduced in part (a) and part (b), respectively.

(d) If $x \in \mathcal{F}_{\underline{\lambda}}(X)$, $y \in \mathcal{F}_{\underline{\lambda}}(Y)$ and $X \cap Y \neq \emptyset$, then

$$\sum_{i=1}^{n_x}\sum_{j=1}^{n_y}\lambda_{ij}\mathbb{1}_{\{d(x_i,y_j)\neq 0\}} \leq 1 - \underline{\lambda}^2, \tag{12}$$

where λ_{ij} were introduced in part (b).

Proof:

(a) Mimicking the idea introduced at the beginning of this section, since $x \in conv(X)$ it follows that there exists a positive integer m such that $z \in X_m$, where $X_{m+1} = \tilde{\psi}(X_m)$ with $X_0 = X$. Farther, there exist $z_1, z_2 \in X_{m-1}$ and $\lambda_{12} \in [0,1]$ such that $z = \psi(z_1, z_2, \lambda_{12})$. Using the definition of the convex structure, it follows that the distance between z and y can be upper bounded by

$$d(x,y) \leq \lambda_{12} d(z_1,y) + (1-\lambda_{12}) d(z_2,y).$$

Inductively decomposing z_1, z_2 and their *children*, it can be easily argued that

$$d(x,y) \leq \sum_{i=1}^{n_x}\lambda_i d(x_i,y),$$

for some positive weights $\lambda_i \geq 0$ summing up to one. Since we assumed $x \in int(conv(X))$ we get that $\lambda_i > 0$, for $i = 1\ldots n$.

(b) To obtain (10) we proceed as in the previous lines and obtain upper bounds on $d(x_i,y)$. More precisely we get that

$$d(x_i,y) \leq \sum_{j=1}^{n_y}\mu_j d(x_i,y_j), \ \forall i,$$

with $\mu_j > 0$ and $\sum_{j=1}^{n_y}\mu_j = 1$, and it follows that

$$d(x,y) \leq \sum_{i=1}^{n_x}\sum_{j=1}^{n_y}\lambda_{ij} d(x_i,y_j),$$

where $\lambda_{ij} = \lambda_i\mu_j > 0$ and $\sum_{i=1}^{n_x}\sum_{j=1}^{n_y}\lambda_{ij} = 1$.

(c) We note that $\lambda_i = \mathcal{W}(P_{x,x_i})$ and $\mu_j = \mathcal{W}(P_{y,y_j})$, $\forall i, j$. But since $x \in \mathcal{F}_{\underline{\lambda}}(X)$ and $y \in \mathcal{F}_{\underline{\lambda}}(Y)$ it immediately follows that $\lambda_i \geq \underline{\lambda}$ and $\mu_j \geq \underline{\lambda}$, and therefore $\lambda_{ij} = \underline{\lambda}^2$.

(d) If $X \cap Y \neq \emptyset$ then there exists at least one pair (i, j) such that $d(x_i, y_j) = 0$. But since $\lambda_{ij} \geq \underline{\lambda}^2$ the inequality (12) follows. ■

III. Problem formulation and statement of the main result

We consider a convex metric space $(\mathcal{X}, d)$ and a set of n agents indexed by i which take values on $\mathcal{X}$. Denoting by k the time index, the agents exchange information based on a communication network modeled by a time varying graph $G(k) = (V, E(k))$, where V is the set of vertices (the agents) and $E(k)$ is the set of edges. An edge (communication link) $e_{ij}(k) \in E(k)$ exits if node i receives information from node j. Each agent has an initial value in $\mathcal{X}$. At each subsequent time-slot is adjusting his/her value based on the observations about the values of his/her neighbors. The goal of the agents is to asymptotically agree on the same value. In what follows we denote by $x_i(k) \in \mathcal{X}$ the value or *state* of agent i at time k.

Definition 3.1: We say that the agents asymptotically reach *consensus* (or agreement) if

$$\lim_{k\to\infty} d(x_i(k), x_j(k)) = 0, \ \forall i, j, \ i \neq j. \tag{13}$$

Similar to the communication models used in [24], [2], [10], we impose minimal assumptions on the connectivity of the communication graph $G(k)$. Basically these assumption consists in having the communication graph connected *infinitely often* and having *bounded intercommunication interval* between neighboring nodes.

Assumption 3.1 (Connectivity): The graph (V, E_∞) is connected, where E_∞ is the set of edges (i, j) representing agent pairs communicating directly infinitely many times, i.e.,

$$E_\infty = \{(i, j) \mid (j, i) \in E(k) \text{ for infinetly many indices } k\}$$

Assumption 3.2 (Bounded intercommunication interval): There exists an integer $B \geq 1$ such that for every $(i, j) \in E_\infty$ agent j sends his/her information to the neighboring agent i at least once every B consecutive time slots, i.e. at time k or at time $k+1$ or $\ldots$ or (at latest) at time $k+B-1$ for any $k \geq 0$.

Assumption 3.2 is equivalent to the existence of an integer $B \geq 1$ such that

$$(i, j) \in E(k) \cup E(k+1) \cup \ldots \cup E(k+B-1), \ \forall (i, j) \in E_\infty.$$

Let $\mathcal{N}_i(k)$ denote the communication neighborhood of agent i, i.e. the set of all nodes sending information to i at time k, which by convention contains the node i itself. We denote by $A_i(k) \triangleq \{x_j(k), \forall j \in \mathcal{N}_i(k)\}$ the set of the states of agent i's neighbors (its own included).

The following theorem states our main result regarding the asymptotic agreement problem on metric convex space.

Theorem 3.1: Let Assumptions 3.1 and 3.2 hold for $G(k)$ and let $\underline{\lambda} < 1$ be a positive scalar sufficiently small. If agents update their state according to the scheme

$$x_i(k+1) \in C_{\underline{\lambda}}(A_i(k)), \ \forall i, \tag{14}$$

then they asymptotically reach consensus, i.e.

$$\lim_{k\to\infty} d(x_i(k), x_j(k)) = 0, \ \forall i, j, \ i \neq j. \tag{15}$$

Remark 3.1: We would like to point out that the result refers strictly to the convergence of the distances between states and not to the convergence of the states themselves. It may be the case that the sequences $\{x_i(k)\}_{k\geq 0}$ $i = 1 \ldots n$ do not have a limit and still the distances $d(x_i(k), x_j(k))$ decrease to zero as k goes to infinity. In other words the agents asymptotically agree on the same value which may be very well variable.

Remark 3.2: A procedure for generating points for which is guaranteed to belong to $C_{\underline{\lambda}}(A_i(k))$ is described in Remark 2.1. The idea of picking $x_i(k+1)$ from $C_{\underline{\lambda}}(A_i(k))$ rather than $int(conv(A_i(k)))$ is in the same spirit of the assumption imposed on the non-zero consensus weights in [23], [10], [2], i.e. they are assumed lower bounded by a positive, sub-unitary scalar. Setting $x_i(k+1) \in int(conv(A_i(k)))$ may not necessarily guarantee asymptotic convergence to consensus. Indeed, consider the case where $\mathcal{X} = \mathbb{R}$ with the standard Euclidean distance. A convex structure on $\mathbb{R}$ is given by $\psi(x, y, \lambda) = \lambda x + (1-\lambda)y$, for any $x, y \in \mathbb{R}$ and $\lambda \in [0, 1]$. Assume that we have two agents which exchange information at all time slots and therefore $A_1(k) = \{x_1(k), x_2(k)\}$, $A_2(k) = \{x_1(k), x_2(k)\}$, $\forall k \geq 0$. Let $x_1(k+1) = \lambda(k)x_1(k) + (1-\lambda(k))x_2(k)$, where $\lambda(k) = 1 - 0.1e^{-k}$ and let $x_2(k+1) = \mu(k)x_1(k) + (1-\mu(k))x_2(k)$, where $\mu(k) = 0.1e^{-k}$. Obviously, $x_i(k+1) \in int(conv(A_i(k)))$, $i = 1, 2$ for all $k \geq 0$. It can be easily argued that

$$d(x_1(k+1), x_2(k+1)) \leq (\lambda(k)(1-\mu(k)) + \mu(k)(1-\lambda(k)))\, d(x_1(k, x_2(k))). \tag{16}$$

We note that $\lim_{K\to\infty} \prod_{k=0}^{K} (\lambda(k)(1-\mu(k)) + (1-\lambda(k))\mu(k)) = \lim_{K\to\infty} \prod_{k=0}^{K} (1 - 0.2e^{-k} + 0.02e^{-2k}) = 0.73$ and therefore under inequality (16) asymptotic convergence to consensus is not guaranteed. In fact it can be explicitly shown that the agents do not reach consensus. From the dynamic equation governing the evolution of $x_i(k)$, $i = 1,2$, we can write

$$\mathbf{x}(k+1) = \begin{pmatrix} \lambda(k) & 1-\lambda(k) \\ \mu(k) & 1-\mu(k) \end{pmatrix} \mathbf{x}(k),\ \ \mathbf{x}(0) = \mathbf{x}_0,$$

where $\mathbf{x}(k)^T = [x_1(k), x_2(k)]$, and we obtain that

$$\lim_{k\to\infty} \mathbf{x}(k) = \begin{pmatrix} 0.8540 & 0.1451 \\ 0.1451 & 0.8540 \end{pmatrix} \mathbf{x}_0$$

and therefore it can be easily seen that consensus is not reached from any initial states.

IV. Proof of the main result

This section is divided in three parts. In the first part we use the results of Section II-B regarding the convex hull of a finite set and show that the entries of the vector of distances between the states of the agents at time $k+1$ are upper bounded by linear combinations of the entries of the same vector but at time k. The coefficients of the linear combinations are the entries of a time varying matrix for which we prove a number of properties (Lemma 4.1). In the second part we analyze the properties of the transition matrix of the aforementioned time varying matrix (Lemma 4.2). The last part is reserved to the proof of Theorem 3.1.

Lemma 4.1: Given a small enough positive scalar $\underline{\lambda} < 1$, assume that agents update their states according to the scheme $x_i(k+1) \in C_{\underline{\lambda}}(A_i(k))$, for all i. Let $\mathbf{d}(k) \triangleq (d(x_i(k), x_j(k)))$ for $i \neq j$ be the N dimensional vector of all distances between the states of the agents, where $N = \frac{n(n-1)}{2}$. Then we obtain that

$$\mathbf{d}(k+1) \leq \mathbf{W}(k)\mathbf{d}(k),\ \ \mathbf{d}(0) = \mathbf{d}_0, \tag{17}$$

where the $N \times N$ dimensional matrix $\mathbf{W}(k)$ has the following properties:

(a) $\mathbf{W}(k)$ is non-negative and there exits a positive scalar $\eta \in (0,1)$ such that

$$[\mathbf{W}(k)]_{\bar{i}\bar{i}} \geq \eta,\ \ \forall\ \bar{i}, k \tag{18}$$

$$[\mathbf{W}(k)]_{\bar{i}\bar{j}} \geq \eta,\ \ \forall\ [\mathbf{W}(k)]_{\bar{i}\bar{j}} \neq 0,\ \ \bar{i} \neq \bar{j},\ \ \forall\ k. \tag{19}$$

(b) If $\mathcal{N}_i(k) \cap \mathcal{N}_j(k) \neq \emptyset$, then the row $\bar{i}$ of matrix $\mathbf{W}(k)$, corresponding to the pair of agents (i, j), has the property

$$\sum_{\bar{j}=1}^{N} [\mathbf{W}(k)]_{\bar{i}\bar{j}} \leq 1-\eta, \tag{20}$$

where η is the same as in part (a).

(c) If $\mathcal{N}_i(k) \cap \mathcal{N}_j(k) = \emptyset$ then the row $\bar{i}$ corresponding to the pair of agents (i, j) sums up to one, i.e.

$$\sum_{\bar{j}=1}^{N} [\mathbf{W}(k)]_{\bar{i}\bar{j}} = 1. \tag{21}$$

In particular if $G(k)$ is completely disconnected (i.e. agents do not send any information), then $\mathbf{W}(k) = I$.

(d) the rows of $\mathbf{W}(k)$ sum up to a value smaller or equal then one, i.e.

$$\sum_{\bar{j}=1}^{N} [\mathbf{W}(k)]_{\bar{i}\bar{j}} \leq 1, \ \forall \ \bar{i}, k. \tag{22}$$

Proof: Given two agents i and j, by part (b) of Proposition 2.2 the distance between their states can be upper bounded by

$$d(x_i(k+1), x_j(k+1)) \leq \sum_{p \in \mathcal{N}_i(k), q \in \mathcal{N}_j(k)} w_{pq}^{ij}(k) d(x_p(k), x_q(k)), \ i \neq j, \tag{23}$$

where $w_{pq}^{ij}(k) > 0$ and $\sum_{p \in \mathcal{N}_i(k), q \in \mathcal{N}_j(k)} w_{pq}^{ij}(k) = 1$. By defining $\mathbf{W}(k) \triangleq (w_{pq}^{ij}(k))$ for $i \neq j$ and $p \neq q$ (where the pairs (i, j) and (p, q) refer to the rows and columns of $\mathbf{W}(k)$, respectively), inequality (17) follows. We continue with proving the properties of matrix $\mathbf{W}(k)$.

(a) Since all $w_{pq}^{ij}(k) > 0$ for all $i \neq j$, $p \in \mathcal{N}_i(k)$ and $q \in \mathcal{N}_j(k)$ we obtain that $\mathbf{W}(k)$ is non-negative. By part (c) of Proposition 2.2, there exists $\eta \triangleq \underline{\lambda}^2$ such that $w_{pq}^{ij}(k) \geq \eta$ for all non-zero entries of $\mathbf{W}(k)$. Also, since $i \in \mathcal{N}_i(k)$ and $j \in \mathcal{N}_j(k)$ for all $k \geq 0$ it follows that the term $w_{ij}^{ij}(k) d(x_i(k), x_j(k))$, with $w_{ij}^{ij}(k) \geq \eta$ will always be present in the right-hand side of the inequality (23), and therefore $\mathbf{W}(k)$ has positive diagonal entries.

(b) Follows from part (d) of Proposition 2.2, with $\eta = \underline{\lambda}^2$.

(c) If $\mathcal{N}_i(k) \cap \mathcal{N}_j(k) = \emptyset$ then no terms of the form $w_{pp}^{ij}(k) d(x_p(k), x_p(k))$ will appear in the sum of the right hand side of inequality (23). Hence $\sum_{p \in \mathcal{N}_i(k), q \in \mathcal{N}_j(k)} w_{pq}^{ij}(k) = 1$ and therefore

$$\sum_{\bar{j}=1}^{N} [\mathbf{W}(k)]_{\bar{i}\bar{j}} = 1.$$

If $G(k)$ is completely disconnected, then the sum of the right hand side of inequality (23) will have only the term $w_{ij}^{ij}(k) d(x_i(k), x_j(k))$ with $w_{ij}^{ij}(k) = 1$, for all $i, j = 1 \ldots n$. Therefore $\mathbf{W}(k)$ is the identity matrix.

(d) The result follows from parts (b) and (c). ■

Let $\bar{G}(k) = (\bar{V}, \bar{E}(k))$ be the underlying graph of $\mathbf{W}(k)$ and let $\bar{i}$ and $\bar{j}$ refer to the rows and columns of $\mathbf{W}(k)$, respectively. Note the under this notation, index $\bar{i}$ corresponds to a pair (i, j) of distinct agents. It is not difficult to see that the set of edges of $\bar{G}(k)$ is given by

$$\bar{E}(k) = \{((i, j),(p,q)) \mid (i, p) \in E(k),(j,q) \in E(k), i \neq j, \;\; p \neq q\}. \tag{24}$$

Proposition 4.1: Let Assumptions 3.1 and 3.2 hold for $G(k)$. Then, similar properties hold for $\bar{G}(k)$ as well, i.e.

(a) the graph $(\bar{V}, \bar{E}_\infty)$ is connected, where

$$\bar{E}_\infty = \{(\bar{i}, \bar{j}) \mid (\bar{i}, \bar{j}) \in \bar{E}(k) \text{ infinetly many indices } k\};$$

(b) there exists an integer $\bar{B} \geq 1$ such that every $(\bar{i}, \bar{j}) \in \bar{E}_\infty$ appears at least once every $\bar{B}$ consecutive time slots, i.e. at time k or at time $k+1$ or … or (at latest) at time $k+\bar{B}-1$ for any $k \geq 0$.

Proof:

It is not difficult to observe that similar to (24), $\bar{E}_\infty$ is given by

$$\bar{E}_\infty = \{((i, j),(p,q)) \mid (i, p) \in E_\infty,(j, p) \in E_\infty, p \neq q, i \neq j\}. \tag{25}$$

(a) Showing that $(\bar{V}, \bar{E}_\infty)$ is connected is equivalently to showing that for any two pairs (i, j) and (p,q) there exits a path connecting them. Since (V, E_∞) is assumed connected, there exits a path $i_0 \to i_1 \to \ldots, \to i_{l-1} \to i_l$, for some $l \leq n$, such that $i_0 = p$ and $i_l = i$. From (25), it is easily argued that $(i_0, j) \to (i_1, j) \to \ldots \to (i_{l-1}, j) \to (i_l, j)$ represents a path connecting (i, j) with (p, j). Similarly, there exits a path $j_0 \to j_1 \to \ldots \to j_{m-1} \to j_m$ for some $m \leq n$, such that $j_0 = q$ and $j_m = j$. Therefore, $(p, j_0) \to (p, j_1) \to \ldots \to (p, j_{m-1}) \to (p, j_m)$ is a path connecting (p, j) with (p,q) and it follows that (i, j) and (p,q) are connected.

(b) Let $((i, j),(p,q))$ be an edge in $\bar{E}_\infty$ or equivalently $(i, p) \in E_\infty$ and $(j,q) \in E_\infty$. By Assumption 3.2, we have that for any $k \geq 0$

$$(i, p) \in E(k) \cup E(k+1) \ldots \cup E(k+B-1),$$

$$(j,q) \in E(k) \cup E(k+1) \ldots \cup E(k+B-1),$$

where the scalar B was introduced in Assumption 3.2. But this also implies that

$$(\bar{i},\bar{j}) \in \bar{E}(k) \cup \bar{E}(k+1) \cup \ldots \cup \bar{E}(k+B-1),\ \forall (\bar{i},\bar{j}) \in \bar{E}_\infty.$$

Choosing $\bar{B} \triangleq B$, the result follows. ■

Let $\Phi(k,s) \triangleq \mathbf{W}(k-1)\mathbf{W}(k-2)\cdots\mathbf{W}(s)$, with $\Phi(k,k) = \mathbf{W}(k)$ denote the transition matrix of $\mathbf{W}(k)$ for any $k \geq s$. It should be obvious from the properties of $\mathbf{W}(k)$ that $\Phi(k,s)$ is a non-negative matrix with positive diagonal entries and $\|\Phi(k,s)\|_\infty \leq 1$ for any $k \geq s$.

Lemma 4.2: Let $\mathbf{W}(k)$ be the matrix introduced in Lemma 4.1. Let Assumptions 3.1 and 3.2 hold for $G(k)$. Then there exits a row index $\bar{i}^*$ such that

$$\sum_{\bar{j}=1}^{N} [\Phi(s+m,s)]_{\bar{i}^*\bar{j}} \leq 1-\eta^m\ \forall\ s,m \geq \bar{B}-1, \tag{26}$$

where η is the lower bound on the non-zero entries of $\mathbf{W}(k)$ and $\bar{B}$ is the positive integer from the part (b) of the Proposition 4.1.

Proof: Let $(i^*,j^*) \in E_\infty$ be a pair of agents. By Assumptions 3.1 and 3.2, there exits a positive integer $s' \in \{s,s+1,\ldots,s+\bar{B}-1\}$ such that agent j^* sends information to agent i^*. This implies that $\mathcal{N}_{i^*}(k) \cap \mathcal{N}_{j^*}(k) \neq \emptyset$ and by part (b) of Lemma 4.1, we have that

$$\sum_{\bar{j}=1}^{N} [\mathbf{W}(s')]_{\bar{i}^*\bar{j}} \leq 1-\eta,$$

where $\bar{i}^*$ is the index corresponding to the pair (i^*,j^*). The sum of the $\bar{i}^*$ row of transition matrix $\Phi(s'+1,s)$ can be expressed as

$$\sum_{\bar{j}=1}^{N} [\Phi(s'+1,s)]_{\bar{i}^*\bar{j}} = \sum_{\bar{j}=1}^{N} [\mathbf{W}(s')]_{\bar{i}^*\bar{j}} \sum_{\bar{h}=1}^{N} [\Phi(s',s)]_{\bar{j}\bar{h}}.$$

But since $\|\Phi(k,s)\|_\infty \leq 1$ for any $k \geq s$, we have that $\sum_{\bar{h}=1}^{N} [\Phi(s',s)]_{\bar{j}\bar{h}} \leq 1$ for any $\bar{j}$, and therefore

$$\sum_{\bar{j}=1}^{N} [\Phi(s'+1,s)]_{\bar{i}^*\bar{j}} \leq 1-\eta. \tag{27}$$

We can write $\Phi(s'+2,s) = \mathbf{W}(s'+1)\Phi(s'+1,s)$ and it follows that the $\bar{i}^*$ row sum of $\Phi(s'+2,s)$ can be expressed as

$$\sum_{\bar{j}=1}^{N} [\Phi(s'+2,s)]_{\bar{i}^*\bar{j}} = \sum_{\bar{j}=1}^{N} [\mathbf{W}(s'+1)]_{\bar{i}^*\bar{j}} \sum_{\bar{h}=1}^{N} [\Phi(s'+1,s)]_{\bar{j}\bar{h}}.$$

Since $\sum_{h=1}^{N}[\Phi(s'+1,s)]_{\bar{j}\bar{h}} \leq 1$ for any $\bar{j}$ it follows that

$$\sum_{\bar{j}=1}^{N}[\Phi(s'+2,s)]_{\bar{i}^*\bar{j}} \leq [\mathbf{W}(s'+1)]_{\bar{i}^*\bar{i}}\sum_{\bar{h}=1}^{N}[\Phi(s'+1,s)]_{\bar{i}^*\bar{h}} + \sum_{\bar{j}=1,\bar{j}\neq\bar{i}^*}[\mathbf{W}(s'+1)]_{\bar{i}^*\bar{j}} \leq$$

$$\leq [\mathbf{W}(s'+1)]_{\bar{i}^*\bar{i}^*}(1-\eta) + \sum_{\bar{j}=1,\bar{j}\neq\bar{i}^*}[\mathbf{W}(s'+1)]_{\bar{i}^*\bar{j}} \leq \sum_{\bar{j}=1}^{N}[\mathbf{W}(s'+1)]_{\bar{i}^*\bar{j}} - \eta[\mathbf{W}(s'+1)]_{\bar{i}^*\bar{i}^*} \leq 1-\eta^2,$$

since $[\mathbf{W}(s'+1)]_{\bar{i}^*\bar{i}^*} \geq \eta$. By induction it can be easily argued that

$$\sum_{\bar{j}=1}^{N}[\Phi(s'+m,s)]_{\bar{i}^*\bar{j}} \leq 1-\eta^m,\ \forall m \geq 0. \tag{28}$$

Note that by Assumption 3.2, a pair (i,j) can exchange information at $s'=s$ the earliest or at $s'=s+B-1$ the latest. From (28) we obtain that for $s'=s+B-1$

$$\sum_{\bar{j}=1}^{N}[\Phi(s+B-1+m,s)]_{\bar{i}^*\bar{j}} \leq 1-\eta^m,\ \forall m \geq 0, \tag{29}$$

and for $s'=s$

$$\sum_{\bar{j}=1}^{N}[\Phi(s+m,s)]_{\bar{i}^*\bar{j}} \leq 1-\eta^m,\ \forall m \geq 0,$$

or

$$\sum_{\bar{j}=1}^{N}[\Phi(s+B-1+m,s)]_{\bar{i}^*\bar{j}} \leq 1-\eta^{m+B-1},\ \forall m \geq 0, \tag{30}$$

From (29) and (30) we get

$$\sum_{\bar{j}=1}^{N}[\Phi(s+B-1+m,s)]_{\bar{i}^*\bar{j}} \leq 1-\eta^{m+B-1},\ \forall s,m \geq 0,$$

or equivalently

$$\sum_{\bar{j}=1}^{N}[\Phi(s+m,s)]_{\bar{i}^*\bar{j}} \leq 1-\eta^m,\ \forall m \geq B-1. \tag{31}$$

■

Corollary 4.1: Let $\mathbf{W}(k)$ be the matrix introduced in Lemma 4.1 and let Assumptions 3.1 and 3.2 hold for $G(k)$. We then have

$$[\Phi(s+(N-1)\bar{B}-1,s)] \geq \eta^{(N-1)\bar{B}}\ \forall s,i,j, \tag{32}$$

where η is the lower bound on the non-zero entries of $\mathbf{W}(k)$ and $\bar{B}$ is the positive integer from the part (b) of the Proposition 4.1.

Proof: By Proposition 4.1 and Lemma 4.1 all the assumptions of Lemma 2, [10] are satisfied, from which the result follows. ■

We are now ready to prove **Theorem 3.1**.

Proof: We have that the vector of distances between the states of the agents respects the inequality

$$\mathbf{d}(k+1) \leq \mathbf{W}(k)\mathbf{d}(k),$$

where the properties of $\mathbf{W}(k)$ are described by Lemma 4.1.

It immediately follows that

$$\|\mathbf{d}(k+1)\|_\infty \leq \|\mathbf{d}(k)\|_\infty, \text{ for } k \geq 0. \tag{33}$$

Let $\bar{B}_0 \triangleq (N-1)\bar{B}-1$, where $\bar{B}$ is the positive integer from the part (b) of the Proposition 4.1. It the following we show that all row sums of $\Phi(s+2B_0, s)$ are upper-bounded by a positive scalar strictly less than one. Indeed since $\Phi(s+2\bar{B}_0, s) = \Phi(s+2\bar{B}_0, s+\bar{B}_0)\Phi(s+\bar{B}_0, s)$ we obtain that

$$\sum_{\bar{j}=1}^{N}[\Phi(s+2\bar{B}_0, s)]_{\bar{i}\bar{j}} = \sum_{\bar{j}=1}^{N}[\Phi(s+2\bar{B}_0, s+\bar{B}_0)]_{\bar{i}\bar{j}}\sum_{\bar{h}=1}^{N}[\Phi(s+\bar{B}_0, s)]_{\bar{j}\bar{h}}, \ \forall \bar{i}.$$

By Lemma 4.2 we have that there exists a row $\bar{j}^*$ such that

$$\sum_{\bar{h}=1}^{N}[\Phi(s+\bar{B}_0, s)]_{\bar{j}^*\bar{h}} \leq 1-\eta^{\bar{B}_0}, \forall s,$$

and since $\sum_{\bar{h}=1}^{N}[\Phi(s+\bar{B}_0, s)]_{\bar{j}\bar{h}} \leq 1$ for any $\bar{j}$, we get

$$\sum_{\bar{j}=1}^{N}[\Phi(s+2\bar{B}_0, s)]_{\bar{i}\bar{j}} \leq \sum_{\bar{j}=1, \bar{j}\neq \bar{j}^*}^{N}[\Phi(s+2\bar{B}_0, s+\bar{B}_0)]_{\bar{i}\bar{j}} + [\Phi(s+2\bar{B}_0, s+\bar{B}_0)]_{\bar{i}\bar{j}^*}(1-\eta^{\bar{B}_0}) =$$

$$= \sum_{\bar{j}=1}^{N}[\Phi(s+2\bar{B}_0, s+\bar{B}_0)]_{\bar{i}\bar{j}} - [\Phi(s+2\bar{B}_0, s+\bar{B}_0)]_{\bar{i}\bar{j}^*}\eta^{\bar{B}_0}.$$

By Corollary 4.1 it follows that

$$[\Phi(s+2\bar{B}_0, s+\bar{B}_0)]_{\bar{i}\bar{j}} \geq \eta^{\bar{B}_0+1}, \ \forall \bar{i}, \bar{j}, s,$$

and since $\sum_{\bar{j}=1}^{N}[\Phi(s+2\bar{B}_0, \bar{B}_0)]_{\bar{i}\bar{j}} \leq 1$ we get that

$$\sum_{\bar{j}=1}^{N}[\Phi(s+2\bar{B}_0, s)]_{\bar{i}\bar{j}} \leq 1-\eta^{2\bar{B}_0+1} \ \forall \bar{i}, s.$$

Therefore

$$||\Phi(s+2\bar{B}_0,s)||_\infty \leq 1-\eta^{2\bar{B}_0+1} \ \forall s. \tag{34}$$

It follows that

$$||\mathbf{d}(t_k)||_\infty \leq \left(1-\eta^{2\bar{B}_0+1}\right)^k ||\mathbf{d}(0)||_\infty, \ \forall k \geq 0, \tag{35}$$

where $t_k = k\bar{B}_0$ which shows that the subsequence $\{||\mathbf{d}(t_k)||_\infty\}_{k\geq 0}$ asymptotically converges to zero. Combined with inequality (33) we farther obtain that the sequence $\{||\mathbf{d}(k)||_\infty\}_{k\geq 0}$ asymptotically converges to zero. Therefore the agents asymptotically reach consensus.

■

V. Distance between the consensus points and the initial points

In this section we analyze the evolution of the distance between the states of the agents and their initial values under the scheme described by Theorem 3.1. This analysis will give us upper bounds on the distance between the consensus point(s) and the initial values of the agents.

Consider distance $d(x_i(k),x_l(0))$ for some i,l and let us assume that $x_i(k+1)$ is chosen according to the scheme described by Theorem 3.1, i.e. $x_i(k+1) \in C_{\underline{\lambda}}(A_i(k))$. By part (a) of Proposition 2.2 we can express this distance as

$$d(x_i(k+1),x_l(0)) \leq \sum_{j\in\mathcal{N}_i(k)} \lambda_{ij}(k) d(x_j(k),x_l(0)), \tag{36}$$

where $\lambda_{ij}(k) \geq \underline{\lambda}$ and $\sum_{j\in\mathcal{N}_i(k)} \lambda_{ij}(k) = 1$. By defining the n dimensional vector $\eta^l(k) = (d(x_i(k),x_l(0)))$ (where i varies) and the $n\times n$ dimensional matrix $\Lambda(k) = (\lambda_{ij}(k))$, inequality (36) can be compactly written as

$$\eta^l(k+1) \leq \Lambda(k)\eta^l(k), \ \eta^l(0) = \eta^l_0. \tag{37}$$

where $\Lambda(k)$ is a row stochastic matrix. It is not difficult to note that the underlying graph of $\Lambda(k)$ is $G(k)$ and that in fact inequality (37) is valid for any l. In the following proposition we give upper bounds on the distance between the consensus states and the initial values of the states.

Proposition 5.1: Let Assumptions 3.1 and 3.2 hold for $G(k)$ and let the states of the agents be updates according to the scheme given by Theorem 3.1. We then have that

$$\lim_{k\to\infty} d(x_i(k),x_l(0)) \leq \sum_{j=1}^n \nu_j d(x_j(0),x_l(0)), \ \forall \ i,l, \tag{38}$$

where $v = (v_j)$ is a vector with positive entries summing up to one satisfying

$$\lim_{k\to\infty} \Lambda(k)\Lambda(k-1)\cdots\Lambda(0) = \mathbb{1}v^T, \tag{39}$$

and where $\Lambda(k)$ is the matrix defined in inequality (37).

Proof: Our assumptions fit the assumptions of Lemmas 3 and 4 of [10], from where (39) follows. Therefore by inequality (37) the result follows. ■

Remark 5.1: If in addition to the assumptions of Proposition 5.1 we also assume that $\Lambda(k)$ is doubly stochastic, then by Proposition 1 of [10] we get that

$$\lim_{k\to\infty} \Lambda(k)\Lambda(k-1)\cdots\Lambda(0) = \frac{1}{n}\mathbb{1}\mathbb{1}^T.$$

Therefore, inequality (38) gets simplified to

$$\lim_{k\to\infty} d(x_i(k), x_l(0)) \leq \frac{1}{n}\sum_{j=1}^{n} d(x_j(0), x_l(0)), \ \forall i.$$

The assumptions in this remark correspond to the assumptions for the average consensus problem in Euclidean spaces. For the aforementioned case, the consensus point is given by the average of the initial points, i.e. $x_{av} = \frac{1}{n}\sum_{i=1}^{n} x_i(0)$. It can be easily check that indeed x_{av} satisfies

$$||x_{av} - x_l(0)|| \leq \frac{1}{n}\sum_{j=1}^{n} ||x_j(0) - x_l(0)||,$$

where $||\cdot||$ represents the euclidean norm.

VI. Application - Asymptotic consensus of opinion

Social networks play a central role in the sharing of information and formation of opinions. This is true in the context of advising friends on which movies to see, relaying information about the abilities and fit of a potential new employee in a firm, debating the merits of politicians. In the following we consider a scenario in which a group of agents try to agree on a common opinion. Assume for example that a group of friends would like to go to see a movie. Different members of the group may suggest different movies. A member of the group discusses with all or just some of his/her friends to find out about their opinions. This member gives some weight (importance) to the opinion of his friends based on the trust in their *expertize*. For instance some members of the group are more informed about the quality of the proposed movies, and therefore there opinions may have a heavier influence on the final decision. By repeatingly discussing among themselves, the group of friends have to choose one of the movies.

In the following we mathematically formalize the scenario described above and show that we can use the framework introduced in the previous sections to give an algorithm which ensures asymptotic consensus on opinions. We model the opinion of a member of the group (agent) as a discrete random variable. Under an appropriate metric we show that the metric space of discrete random variable is convex by providing a convex structure. In addition, we analyze in more detail the convex hull of a finite set; this analysis is possible since the convex structure is given explicitly. We give an iterative algorithm that ensures agreement of opinion, which is based on Theorem 3.1 and provide some numerical simulations.

A. Geometric framework

Let s be a positive integer and let $S = \{1,2,\ldots,s\}$ be a finite set. Consider the sample space $\Omega = \{\omega_1,\omega_2,\ldots,\omega_s\}$ and let $(\Omega,\mathcal{F},\mathcal{P})$ be a probability space. We denote by $\mathcal{X}$ the space of discrete random variable defined on $(\Omega,\mathcal{F},\mathcal{P})$, given by $\mathcal{X} \triangleq \{X \mid X : \Omega \to S\}$.

We introduce the operator $d : \mathcal{X}\times\mathcal{X} \to \mathbb{R}$, defined as

$$d(X,Y) = E[\rho(X,Y)],$$

where $\rho : \mathbb{R}\times\mathbb{R} \to \{0,1\}$ is the discrete metric, i.e.

$$\rho(x,y) = \begin{cases} 1 & x \neq y \\ 0 & x = y \end{cases}$$

It is not difficult to note that the operator d can also be written as $d(X,Y) \triangleq E[\mathbb{1}_{\{X\neq Y\}}] = Pr(X \neq Y)$, where $\mathbb{1}_{\{X\neq Y\}}$ is the indicator function of the event $\{X \neq Y\}$.

We note that the operator d satisfies the following properties

1) For any $X,Y \in \mathcal{X}$, $d(X,Y) = 0$ if and only if $X = Y$ with probability one.
2) For any $X,Y,Z \in \mathcal{X}$, $d(X,Z) + d(Y,Z) \geq d(X,Y)$ with probability one,

and therefore is an (almost) metric on $\mathcal{X}$. The set $\mathcal{X}$ together with the operator d define the *(almost) metric space* $(\mathcal{X},d)$. We use the attribute *almost*, to emphasize that the two properties of the operator d are satisfied with probability one.

Let $\psi : \mathcal{X}\times\mathcal{X}\times[0,1] \to \mathcal{X}$ be a mapping given by

$$\psi(X_1,X_2,\lambda) = \mathbb{1}_{\{\theta=1\}}X_1 + \mathbb{1}_{\{\theta=2\}}X_2, \tag{40}$$

where $X_1, X_2 \in \mathcal{X}$ and where $\theta \in \{1,2\}$ is a random variable independent on any event on the σ-field $\mathcal{F}$, with probability mass function $Pr(\theta = 1) = \lambda$ and $Pr(\theta = 2) = 1-\lambda$, where $\lambda \in [0,1]$.

Proposition 6.1: The mapping ψ is a convex structure on $\mathcal{X}$.

Proof: For any $U, X_1, X_2 \in \mathcal{X}$ and $\lambda \in [0,1]$ we have

$$d(U,\psi(X_1,X_2,\lambda)) = E[\rho(U,\psi(X_1,X_2,\lambda))] = E[E[\rho(U,\psi(X_1,X_2,\lambda))|U,X_1,X_2]] =$$

$$= E[E[\rho(U,\mathbb{1}_{\{\theta=1\}}X_1 + \mathbb{1}_{\{\theta=2\}}X_2)]|U,X_1,X_2] == E[\lambda\rho(U,X_1)+(1-\lambda)\rho(U,X_2)] =$$

$$= \lambda d(U,X_1) + (1-\lambda)d(U,X_2).$$

■

From the above proposition it follows that $(\mathcal{X}, d)$ is a *convex metric space*.

The next theorem characterizes the convex hull of a finite set in $\mathcal{X}$.

Theorem 6.1: Let n be a positive integer and let $A = \{X_1,\dots,X_n\}$ be a set of points in $\mathcal{X}$. Consider a sample space Ω_θ and a random variable $\theta : \Omega_\theta \to \{1,\dots,n\}$, independent of any $X \in \mathcal{X}$, with probability measure given by $Pr(\omega : \theta(\omega) = i) = w_i$, for some non-negative scalars w_i, with $\sum_{i=1}^n w_i = 1$. We define the set

$$\mathcal{K}(A) \triangleq \left\{ Z \in \mathcal{X} \mid Z = \sum_{i=1}^{n} \mathbb{1}_{\{\theta=i\}} X_i,\ \forall w_i \geq 0, \sum_{i=1}^{n} w_i = 1 \right\}. \tag{41}$$

Then

$$conv(A) = \mathcal{K}(A).$$

Proof: We recall from Proposition 2.1 that the convex hull of A is given by

$$conv(A) = \lim A_m = \bigcup_{m=1}^{\infty} A_m,$$

where $A_m = \tilde{\psi}(A_{m-1})$, with $A_1 = \tilde{\psi}(A)$. Also, since A_m is an increasing sequence, clearly $A \subset A_m$ for all $m \geq 1$. The proof is structured in two parts. In the first part we show that any point in $\mathcal{K}(A)$ belongs to the convex hull of A, while in the second part we show that any point in $conv(A)$ belongs to $\mathcal{K}(A)$ as well.

Let $Z \in \mathcal{K}(A)$ i.e. $Z = \sum_{i=1}^n \mathbb{1}_{\{\theta=i\}} X_i$ where $Pr(\theta = i) = w_i$, for some $w_i \geq 0$, $\sum_{i=1}^n w_i = 1$. The random variable θ is defined such that $\theta(\omega_i) = i$ and $Pr(\omega_i) = w_i$. Let $\Omega_i = \{\omega_1^i, \omega_2^i\}$, $i = 1\dots n-1$ be a set of sample spaces whose events are independent among themselves (i.e. ω_j^i are independent for $i = 1 \dots n-1,\ j = 1,2$) and of any event on the σ-algebra $\mathcal{F}$. We define the probability measure for each of the events in Ω_i as

$$Pr(\omega_1^i) = \frac{w_1 + \dots + w_{i-1}}{w_1 + \dots + w_i},$$
$$Pr(\omega_2^i) = \frac{w_i}{w_1 + \dots + w_i},$$

for $i = 1 \dots n-1$. We consider the following succession of events from Ω_i

$$\begin{aligned} S_1 &= \left\{\omega_1^1 \omega_1^2 \dots \omega_1^{n-1}\right\}, \\ S_2 &= \left\{\omega_2^1 \omega_1^2 \dots \omega_1^{n-1}\right\}, \\ S_i &= \bigcup_{j_1 \dots j_{i-2}=1}^{2} \left\{\omega_{j_1}^1 \dots \omega_{j_{i-2}}^{i-2} \omega_2^{i-1} \omega_1^i \dots \omega_1^{n-1}\right\}, i = 3 \dots n-1, \\ S_n &= \bigcup_{j_1 \dots j_{n-2}=1}^{2} \left\{\omega_{j_1}^1 \dots \omega_{j_{n-2}}^{n-2} \omega_2^{n-1}\right\}. \end{aligned} \tag{42}$$

For example, for $n = 4$ (42) becomes

$$\begin{aligned} S_1 &= \{\omega_1^1 \omega_1^2 \omega_1^3\}, \\ S_2 &= \{\omega_2^1 \omega_1^2 \omega_1^3\}, \\ S_3 &= \{\omega_1^1 \omega_2^2 \omega_1^3\} \cup \{\omega_2^1 \omega_2^2 \omega_1^3\} \\ S_4 &= \{\omega_1^1 \omega_1^2 \omega_2^3\} \cup \{\omega_2^1 \omega_1^2 \omega_2^3\} \cup \{\omega_1^1 \omega_2^2 \omega_2^3\} \cup \{\omega_2^1 \omega_2^2 \omega_2^3\}. \end{aligned}$$

Using the independence assumption on the events from Ω_i is not difficult no see that

$$Pr(S_i) = w_i,\ \ i = 1 \dots n.$$

Assume that each event ω_i that we observe can be decomposed in a succession of independent events from Ω_i, which are invisible to the observer. In particular let

$$\omega_i = S_i, i = 1 \dots n.$$

The particular decomposition of event ω_i in a set of intermediate, independent events given by S_i makes sense since both ω_i and S_i have the same probability measure. It immediately follows that

$$\mathbb{1}_{\{\omega:\theta(\omega)=i\}} = \mathbb{1}_{\{\omega_i\}} = \mathbb{1}_{\{S_i\}}. \tag{43}$$

Let us now define the random variables $\theta_i : \Omega_i \to \{i, i+1\}$, where

$$\theta_i(\omega_1^i) = i,\ \ \theta_i(\omega_2^i) = i+1,$$

for $i = 1 \dots n-1$. Obviously

$$Pr(\theta_i = i) = \frac{w_1 + \dots + w_{i-1}}{w_1 + \dots + w_i},\ Pr(\theta_i = i+1) = \frac{w_i}{w_1 + \dots + w_i},$$

and θ_i are independent from each other and from any event from $\mathcal{F}$.

From (42) and (43) together with the independence of the random variables θ_i the following equalities in terms of the indicator function are satisfied

$$\begin{aligned} &\mathbb{1}_{\{\theta=1\}} = \Pi_{j=1}^{n-1} \mathbb{1}_{\{\theta_j = j\}} \\ &\mathbb{1}_{\{\theta=i\}} = \mathbb{1}_{\{\theta_{i-1}=i\}} \Pi_{j=i}^{n-1} \mathbb{1}_{\{\theta_j=j\}},\ i = 2 \dots n-1 \\ &\mathbb{1}_{\{\theta=n\}} = \mathbb{1}_{\{\theta_{n-1}=n\}}. \end{aligned} \tag{44}$$

From (44) it follows that Z is the result of the n^{th} step of the iteration

$$Y_{i+1} = \mathbb{1}_{\{\theta_i=i\}} Y_i + \mathbb{1}_{\{\theta_i=i+1\}} X_{i+1},$$

for $i = 1 \dots n$, with $Y_1 = X_1$, i.e. $Z = Y_n$. It can be easily argued that $Y_i \in A_{i-1}$, $i = 2 \dots n$ and therefore $Z \in A_{n-1}$ or $Z \in conv(A)$ which implies that $\mathcal{K}(A) \subset conv(A)$.

We now begin the second part of the proof and show that any point in $conv(A)$ belongs to $\mathcal{K}(A)$ as well. If $Z \in conv(A)$, from Section II-B we have that there exits a positive integer m such that $Z \in A_m$ and therefore Z is the root of a binary tree of height m with leaves from the set A. Using the same notations as in Section II-B for each of the leaf nodes X_i, there exists $n_i \geq 1$ paths from Z to X_i, of lengths m_{i_l}, $l = 1 \dots n_i$ which are denoted by

$$P_{Z,X_i} \triangleq \left\{ \left(\{Y_{i_l,j}\}_{j=0}^{m_{i_l}}, \{\lambda_{i_l,j}\}_{j=1}^{m_{i_l}} \right) \mid l = 1 \dots n_i \right\},$$

where $Y_{i_l,j-1} = \psi\left(Y_{i_l,j}, *, \lambda_{i_l,j}\right)$ for $j = 1 \dots m_{i_l}$, $l = 1 \dots n_i$ and where we denoted by $*$ some intermediate node in the tree. We introduce the independent, random variables $\theta_{i_l,j}$ such that $Pr(\theta_{i_l,j} = i_l, j) = \lambda_{i_l,j}$ and $Pr(\theta_{i_l,j} = *) = 1 - \lambda_{i_l,j}$. It follows that Z can be expressed as

$$Z = \sum_{i=1}^{n} \left(\sum_{l=1}^{n_i} \prod_{j=1}^{m_{i_l}} \mathbb{1}_{\{\omega: \theta_{i_l,j} = i_l, j\}} \right) X_i$$

Using again the independence of $\theta_{i_l,j}$ we have that

$$\sum_{l=1}^{n_i} \prod_{j=1}^{m_{i_l}} \mathbb{1}_{\{\omega:\theta_{i_l,j}=i_l,j\}} = \mathbb{1}_{\{\bigcup_{l=1}^{n_i} \bigcap_{j=1}^{m_{i_l}} \{\omega:\theta_{i_l,j}=i_l,j\}\}}$$

Let $S_i \triangleq \left\{\bigcup_{l=1}^{n_i} \bigcap_{j=1}^{m_{i_l}} \{\omega : \theta_{i_l,j} = i_l, j\}\right\}$ and let us interpret the events in S_i as the set of underlying $sub-events$ generating ω_i i.e. $\omega_i = S_i$. It is not difficult to see that

$$Pr(\omega_i) = Pr(S_i) = \mathcal{W}(P_{Z,X_i}).$$

By defining $w_i \triangleq \mathcal{W}(P_{Z,X_i})$ we get that $\sum_{i=1}^{n} Pr(\omega_i) = 1$. Note that if there exits an i^* such that X_{i^*} is not among the leaves of the binary tree rooted at Z, the measure of the event ω_i is zero. Therefore we have that Z can be expressed as

$$Z = \sum_{i=1}^{n} \mathbb{1}_{\{\omega_i\}} X_i = \sum_{i=1}^{n} \mathbb{1}_{\{\theta=i\}} X_i,$$

where $Pr(\theta = i) = w_i$ and hence it follows that $Z \in \mathcal{K}(A)$ and consequently $conv(A) \subset \mathcal{K}(A)$. From part one and part two of our proof, the result follows.

■

Remark 6.1: We note that any point Z belonging to the convex hull of X_1, X_2 is between X_1 and X_2 in the sense of Definition 2.1. Indeed, if $Z \in conv(\{X_1, X_2\})$ then $Z = \mathbb{1}_{\{\theta=1\}} X_1 + \mathbb{1}_{\{\theta=2\}} X_2$, where $Pr(\theta = 1) = \lambda$ and $Pr(\theta = 2) = 1 - \lambda$, for some $\lambda \in [0, 1]$. It follows that

$$d(X_1, Z) + d(Z, X_2) = E[\rho(X_1, Z) + \rho(Z, X_2)] = E[E[\rho(X_1, Z) + \rho(Z, X_2)]|X_1, X_2] =$$

$$= E[\lambda\rho(X_1, X_2) + (1 - \lambda)\rho(X_1, X_2)] = d(X_1, X_2).$$

However, not any point belonging to the metric segment $[X_1, X_2]$ belongs to $conv(\{X_1, X_2\})$. Indeed, assume for example that $X_1, X_2 \in \{1, 2\}$ and consider a random variable $Z \in \{1, 2\}$ whose probability mass function, conditioned on the values of X_1 and X_2 is given by $Pr(Z = 2|X_1 = 2, X_2 = 1) = \lambda$, $Pr(Z = 1|X_1 = 2, X_2 = 1) = 1 - \lambda$, $Pr(Z = 1|X_1 = 1, X_2 = 2) = \tilde{\lambda}$, $Pr(Z = 1|X_1 = 1, X_2 = 2) = 1 - \tilde{\lambda}$ and $Pr(Z = 2|X_1 = 2, X_2 = 2) = Pr(Z = 1|X_1 = 1, X_2 = 1) = 1$, for some $\lambda \neq \tilde{\lambda} \in (0, 1)$. Since $Pr(Z = 2|X_1 = 2, X_2 = 1) \neq Pr(Z = 1|X_1 = 1, X_2 = 2)$ it follows that $Z \notin conv(\{X_1, X_2\})$. However it can be easily checked that $Z \in [X_1, X_2]$. In fact any random variable Z whose probability mass function conditioned on the values of X_1 and X_2 satisfies

$$\sum_{z \neq x_1 \neq x_2} Pr(Z = z|X_1 = x_1, X_2 = x_2) = 0, \ \sum_{z \neq x} Pr(Z = z|X_1 = x_1, X_2 = x_2) = 0,$$

belongs to the metric segment $[X_1, X_2]$.

Corollary 6.1: Let n be a positive integer, let $A = \{X_1, \ldots, X_n\}$ be a set of points in $\mathcal{X}$ and let $\underline{\lambda} < 1$ be a positive scalar sufficiently small. Consider a sample space Ω_θ and a random variable

$\theta : \Omega_\theta \to \{1, \ldots, n\}$, independent of any $X \in \mathcal{X}$, with probability measure given by $Pr(\omega : \theta(\omega) = i) = w_i$, for some positive scalars w_i, with $\sum_{i=1}^n w_i = 1$. We define the set

$$\mathcal{K}_{\underline{\lambda}}(A)) \triangleq \left\{ Z \in \mathcal{X} \mid Z = \sum_{i=1}^{n} \mathbb{1}_{\{\theta = i\}} X_i, \ \forall w_i \geq \underline{\lambda}, \sum_{i=1}^{n} w_i = 1 \right\}. \tag{45}$$

Then

$$C_{\underline{\lambda}}(A) = \mathcal{K}_{\underline{\lambda}}(A).$$

Proof: Follows immediately from Definitions 2.6, 2.7 and Theorem 6.1. ■

B. Asymptotic Consensus of Opinion Algorithm and Numerical Simulations

We assume that each agent of a group of n agents has an *initial opinion*. We model the set of opinions by a finite set of distinct integers, say $S = \{1, 2, \ldots, s\}$ for some positive integer s, where each element of S indicates an opinion. The goal of the agents is to reach the same opinion by continuously discussing among themselves.

Denoting as before by k the time-index and by $G(k) = (V, E(k))$ the time varying graph modeling the communication network among the n agents, we model the evolution of the opinion of an agent i as a random process $X_i(k)$ on the probability space $(\Omega, \mathcal{F}, \mathcal{P})$, where $X_i(k) \in \mathcal{X}$ for all $k \geq 0$. Each agent i has an initial opinion $X_i(0) = x_{il}^0 \in S$ with probability $p_{il} \geq 0$, with $\sum_{l=1}^s p_{il} = 1$.

Corollary 6.2: Let Assumptions 3.1 and 3.2 hold for $G(k)$. Given a small enough, positive scalar $\underline{\lambda} < 1$, assume that at every time-slot each agent i rolls an imaginary dice with $|\mathcal{N}_i(k)|$ facets numbered from 1 to $|\mathcal{N}_i(k)|$. The probability that the result of a dice roll is $j \in \mathcal{N}_i(k)$, is $w_{ij}(k)$ with $w_{ij}(k) \geq \underline{\lambda}$ and $\sum_{j \in \mathcal{N}_i(k)} w_{ij}(k) = 1$. The agent i updates is state according to the following scheme. If the result of the dice roll is j then agent i chooses the opinion of agent j. We then have that the agents asymptotically agree on the same opinion, i.e.

$$\lim_{k \to \infty} d(X_i(k), X_j(k)) = 0, \forall i, j$$

Proof: By modeling the dice of agent i as an i.i.d. random process $\theta_i(k) \in \{1, 2, \ldots, |\mathcal{N}_i(k)|\}$ such that $Pr(\theta_i(k) = j) = w_{ij}(k)$ for all $j \in \mathcal{N}_i(k)$ and for all $i, k \geq 0$, the update scheme of agent i can be formally written as

$$X_i(k+1) = \sum_{j \in \mathcal{N}_i(k)} \mathbb{1}_{\{\theta_i(k) = j\}} X_j(k).$$

However this implies that $X_i(k+1) \in C_{\underline{\lambda}}(A_i(k)),\ \forall i,k$ and the result follows from Theorem 3.1. ■

Remark 6.2: The above corollary shows that under the proposed scheme the distances between the agents states converge to zero. In general we can not say anything about the convergence of the states themselves. However for this particular metric space more can be said about the convergence of the states. Recall that we define the distance between two points $X_1, X_2 \in \mathcal{X}$ as

$$d(X_1, X_2) = E[\rho(X_1, X_2)] = Pr(X_1 \neq X_2).$$

From the above corollary we have that

$$\lim_{k\to\infty} d(X_i(k), X_j(k)) = 0,$$

or equivalently

$$\lim_{k\to\infty} Pr(X_i(k) \neq X_j(k)) = 0.$$

This says that the measure of the set on which $X_i(k)$ and $X_j(k)$ are different converges to zero as k goes to infinity. Noting that

$$\{\omega : X_i(k) \text{ are not all equal}\} = \bigcup_{i,j,i\neq j} \left\{\omega : X_i(k) \neq X_j(k)\right\},$$

it follows that

$$Pr\left(\omega : X_i(k) \text{ are not all equal}\right) \leq \sum_{i,j,i\neq j} Pr\left(\omega : X_i(k) \neq X_j(k)\right),$$

and therefore

$$\lim_{k\to\infty} Pr\left(\omega : X_i(k) \text{ are not all equal}\right) = 0,$$

or equivalently

$$\lim_{k\to\infty} Pr\left(\bigcup_{o\in S} \{\omega : X_i(k) = o, \forall i\}\right) = 1.$$

However, this does not necessarily imply that $X_i(k)$, $i = 1 \ldots n$ converge to the same value with probability one, i.e.

$$Pr\left(\bigcup_{o\in S} \left\{\omega : \lim_{k\to\infty} X_i(k) = o, \forall i\right\}\right) = 1.$$

It turns out that still this is the case. Recall that by inequality (35), $d(X_i(k), X_j(k)) = Pr(X_i(k) \neq X_j(k))$, $\forall i, j$ converge at least geometrically to zero. Therefore the sum

$$\sum_{k\geq 0}\left(1 - P\left(\bigcup_{o\in S}\{\omega : X_i(k) = o, \forall i\}\right)\right) < \infty,$$

which by the Borel-Cantelli lemma implies that

$$Pr\left(\bigcup_{o\in S}\left\{\omega : \lim_{k\to\infty} X_i(k) = o, \forall i\right\}\right) = 1,$$

and hence the agents converge to the same value with probability one. □

In the following we show that the same result can be obtained by using purely probability theory arguments. For simplicity we assume that the communication network remains constant and connected and that the coefficients w_{ij} are constant as well.

Proposition 6.2: Let G be the graph modeling the communication network assumed connected and let the agents update their state according to the scheme described in Corollary 6.2, where $w_{ij} > 0$ are assumed constant for all $k \geq 0$. We then have that the agents converge to the same value with probability one, i.e.

$$Pr\left(\bigcup_{o\in S}\left\{\omega : \lim_{k\to\infty} X_i(k) = o, \forall i\right\}\right) = 1. \quad (46)$$

Proof: We define the random process $Z(k) = (X_1(k), X_2(k), \ldots, X_n(k))$ which has a maximum of s^s states and we introduce the *agreement space* as

$$\mathcal{A} \triangleq \{(o, o, \ldots, o) \mid o \in S\}.$$

The state update dynamics is given by

$$X_i(k+1) = \sum_{j\in\mathcal{N}_i} \mathbb{1}_{\{\theta_i(k)=j\}} X_j(k),$$

where $Pr(\theta_i(k) = j) = w_{ij}$, for all $j \in \mathcal{N}_i$ and for all i. The conditional probability of $X_i(k+1)$ conditioned on $X_j(k)$, $j \in \mathcal{N}_i$ is given by

$$Pr(X_i(k+1) = o_i | X_j(k) = o_j, j \in \mathcal{N}_i) = \sum_{j\in\mathcal{N}_i} w_{ij} \mathbb{1}_{\{o_i=o_j\}}. \quad (47)$$

It is not difficult to note that $Z(k)$ is a finite state, homogeneous Markov chain. We will show that $Z(k)$ has s absorbing states and all other $s^s - s$ states are transient, where the absorbing states correspond to the states in agreement space $\mathcal{A}$. The entries of the probability transition matrix of $Z(k)$ can be derived from (47) and are given by

$$Pr(X_1(k+1) = o_{l_1}, \ldots, X_n(k+1) = o_{l_n} | X_1(k) = o_{p_1}, \ldots, X_n(k) = o_{p_n}) = \tag{48}$$
$$= \prod_{i=1}^{n} \sum_{j \in \mathcal{N}_i} w_{ij} \mathbb{1}_{\{l_i = p_j\}}.$$

We note from (48) that once the process reaches an agreement state it will stay there indefinitely, i.e.

$$Pr(X_1(k+1) = o, \ldots, X_n(k+1) = o | X_1(k) = o, \ldots, X_n(k) = o) = 1, \ \forall o \in S,$$

and hence the agreement states are absorbing states. We will show next that under the connectivity assumption the agreement space $\mathcal{A}$ is reachable from any other state, and therefore all other states are transient. We are not saying that all agreement states are reachable from any other state, but that from any state at least one agreement state is reachable. Assume that at time zero $Z(0) = (o_{i_1}, \ldots, o_{i_n}) \notin \mathcal{A}$, with $i_j \in \{1, \ldots, N\}$, for $j = 1 \ldots n$. Note that from $Z(0)$ only states with a smaller number distinct entries, compared to $Z(0)$, are reachable. Let j be an agent with an initial choice $o \in \mathcal{O}$, i.e. $X_j(0) = o$. We show that with positive probability the agreement vector $(o, o, \ldots, o)$ can be reached. At time one, with probability w_{jj} agent j keeps its initial choice, while its neighbors can choose o, i.e. $X_i(1) = o$ with probability w_{ij}, $i \in \mathcal{N}_j$. Due to the connectivity assumption $\mathcal{N}_j$ is non-empty. At the next time-index all the agents which have already chosen o, keep their choice with positive probability, while their neighbors will choose o with positive probability. Since the communication network is assumed connected, every agent will be able to choose o with positive probability in at most $n-1$ steps, therefore agreement can be reached. Arguing similarly for any other initial states and other agents, it follows that the probability of reaching an agreement state from $Z(0)$ is positive. Since the agreement states are absorbing it follows that $Z(0)$ is a transient state. Therefore, the probability for the Markov chain $Z(k)$ to be in a transient state converges asymptotically to zero, while the probability to be in one of the agreement states converges asymptotically to one, i.e.

$$\lim_{k \to \infty} Pr(Z(k) = z) = 0, \ \forall \ z \notin \mathcal{A},$$

and

$$\lim_{k \to \infty} \sum_{z \in \mathcal{A}} Pr(Z(k) = z) = 1,$$

which is equivalent to

$$\lim_{k\to\infty}\sum_{o\in S} Pr(X_i(k)=o,\forall i)=1.$$

In addition, due to the geometric decay toward zero of the probability $Pr(Z(k)\notin\mathcal{A})$, by the Borel-Catelli Lemma it follows that $Z(k)$ converges to one of the agreement states not only in probability but also with probability one and the result follows. ■

C. Numerical example

In what follows we consider an example where a group of eight agents ($n=8$) have to choose between two opinions, i.e. $S=\{1,2\}$. We assume that the agents communication network is given by an undirected circular graph as in Figure 2, assumed fixed for all time-slots.

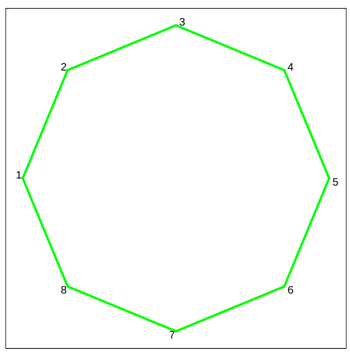

Fig. 2. Undirected circular graph with eight nodes

We assume that the agents use the scheme described by Corollary (6.2) for updating their states, i.e. the coefficients w_{ij} are constant. In particular we choose $w_{ii}=0.7778$ and $w_{i,i-1}=w_{i,i+1}=0.1111$ and choose as initial values $X_i(0)=1$ for $i=1\dots4$ and $X_i(0)=2$ for $i=5\dots8$ with probability one. Figure 3 presents an execution of our agreement algorithm which indeed shows that the agents agree on the same opinion. The different colors that appear indicates different agents.

Next we numerically analyze the evolution of the vector of distances $\mathbf{d}(k)=(d(X_i(k),X_j(k)))$, $\forall\ i\neq j$. First we see that under our assumption the entries of matrix $[\mathbf{W}(k)]_{\bar{i},\bar{j}}=w_{ip}w_{jq}$, where $\bar{i}$ and $\bar{j}$ correspond to the pairs of agents (i,j) and (p,q), respectively, and where w_{ij} define the probability mass function of the random variables $\theta_i(k)$ as described in Corollary 6.2. We

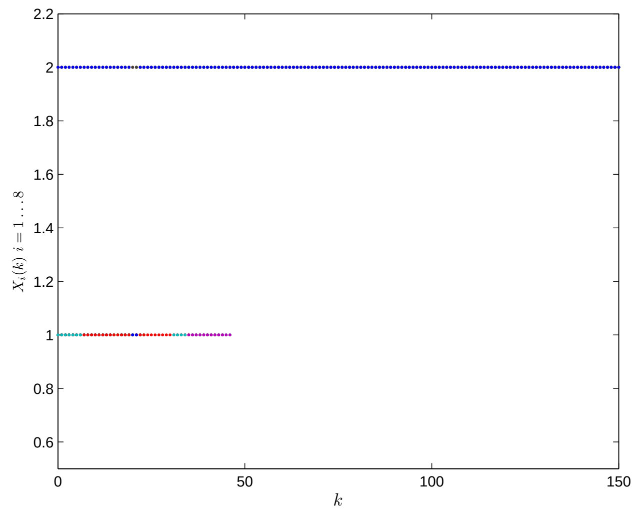

Fig. 3. Execution of the agreement algorithm

consider the linear system

$$\tilde{\mathbf{d}}(k+1) = \mathbf{W}(k)\tilde{\mathbf{d}}(k), \;\; \tilde{\mathbf{d}}(0) = \mathbf{d}(0).$$

By (17) of Lemma 4.1, we have that $\tilde{\mathbf{d}}(k)$ is an upper bound of $\mathbf{d}(k)$. Figure 4 presents the evolution of $\|\tilde{\mathbf{d}}(k)\|_\infty$ with time. It is worth mentioning that since ψ defined in (40) satisfies the definition of a convex structure with equality, it can be easily argued that (17) holds with equality and therefore the upper bound $\tilde{\mathbf{d}}(k)$ is in fact $\mathbf{d}(k)$.

We next analyze the distance between the initial points and the consensus point(s). Since ψ respects the definition of a convex structure with equality, we have that

$$d(X_i(k+1), X_l(0)) = \sum_{j \in \mathcal{N}_i} w_{ij} d(X_j(k), X_l(0)),$$

which is basically a consensus algorithm. Since the consensus matrix is doubly stochastic we know that

$$\lim_{k \to \infty} d(X_i(k), X_l(0)) = \frac{1}{n} \sum_{j=1}^{n} d(X_j(0), X_l(0))$$

Figure 5 presents the evolution of the distance between $X_i(k)$ and $X_1(0)$ for $i = 1 \ldots n$. Considering our choice for initial values and the fact that $n = 8$ it is not difficult to see that

$$\frac{1}{n} \sum_{j=1}^{n} d(X_j(0), X_l(0)) = \frac{1}{2}.$$

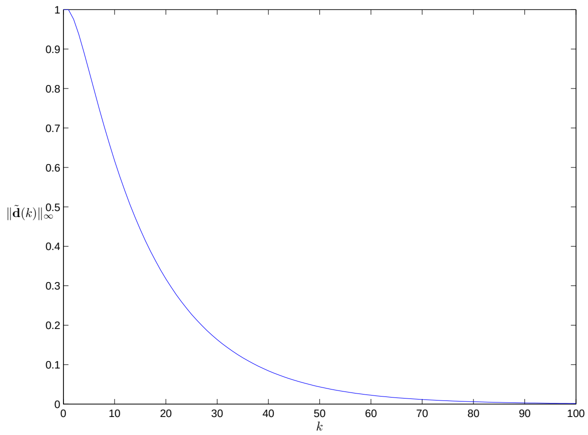

Fig. 4. Evolution of $\|\tilde{\mathbf{d}}(k)\|_\infty$ with time

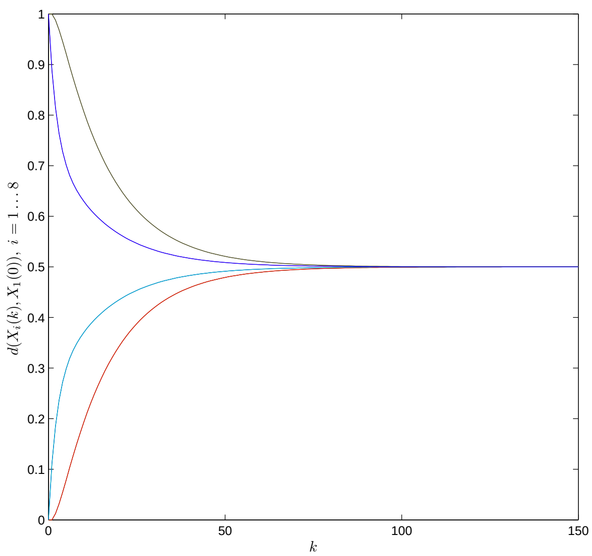

Fig. 5. The distances between $X_i(k)$ and $X_1(0)$ for $i = 1\ldots8$

which is also what Figure 5 shows.

VII. Conclusions

In this report we emphasized the importance of the convexity concept and in particular the importance of the convex hull notion for reaching consensus. We did this by generalizing the asymptotic consensus problem to the case of convex metric space. For a group of agents taking values in a convex metric space, we introduced an iterative algorithm which ensures asymptotic convergence to agreement under some minimal assumptions for the communication graph. As application, we provided an iterative algorithm which guarantees convergence to consensus of opinion.

References

[1] V. Borkar and P. Varaya, ”Asymptotic agreement in distributed estimation”, *IEEE Trans. Autom. Control*, vol. AC-27, no. 3, pp. 650-655, Jun 1982.

[2] V.D. Blondel, J.M. Hendrickx, A. Olshevsky, and J.N. Tsitsiklis, “Convergence in multiagent coordination, consensus, and flocking,” *Proceedings of IEEE CDC*, 2005.

[3] J.A. Fax, ”Optimal and cooperative control of vehicles formations”, Ph.D. dissertation, Control Dynamical Syst., California Inst. Technol., Pasadena, CA, 2001.

[4] J.A. Fax and R.M. Murray, ”Information flow and cooperative control of vehicles formations”, *IEEE Trans. Automatic Control*, vol. 49, no. 9, pp. 1456-1476, Sept. 2004.

[5] A. Jadbabaie, J. Lin and A.S. Morse, ”Coordination of groups of mobile autonomous agents using nearest neighbor”, *IEEE Trans. Autom. Control*, vol. 48, no. 6, pp. 998-1001, Jun 2003.

[6] Y. Hatano and M. Mesbahi, ”Agreement over Random Networks”, *IEEE Trans. Autom. Control*, vol. 50, no. 11, pp. 18671872, 2005.

[7] I. Matei, N. Martins and J. Baras, ”Almost sure convergence to consensus in Markovian random graphs”, *the* 47^{th} *Conference on Decision and Control*, pp. 3535-3540, Cancun, Mexico, Dec. 9-11, 2008.

[8] I. Matei, N. Martins and J. Baras, ”Consensus Problems with Directed Markovian Communication Patterns,“ *Proceedings of the 2009 American Control Conference*, pp. 1298-1303, St. Louis, MO, USA June 10-12, 2009.

[9] L. Moreau, ”Stability of multi-agents systems with time-dependent communication links”, *IEEE Trans. Automatic Control*, vol. 50, no. 2, pp. 169-182, Feb. 2005.

[10] A. Nedic and A. Ozdalgar, ”Distributed Subgradient Methods for Multi-agent Optimization,“ something.

[11] A. Nedic, A. Ozdaglar, and P.A. Parrilo, ”Constrained Consensus and Optimization in Multi-Agent Networks,”, to appear in IEEE Transactions on Automatic Control, 2009.

[12] A. Nedic and A. Ozdaglar, ”Convergence Rate for Consensus with Delays,” to appear in Journal of Global Optimization, 2008.

[13] M. Porfiri and D.J. Stilwell, ”Consensus seeking over Random Directed Weighted Graphs”, *IEEE Trans. Autom. Control*, vol. 52, no. 9, Sept. 2007.

[14] W. Ren and R.W. Beard, ”Consensus seeking in multi-agents systems under dynamically changing interaction topologies”, IEEE Trans. Autom. Control, vol. 50, no. 5, pp. 655-661, May 2005.

[15] S. Roy, K. Herlugson and A. Saberi, “A Control-Theoretic Approach to Distributed Discrete-Valued Decision-Making in Networks of Sensing Agents,” *IEEE Tran. Mobile Computing*, vol. 5, no. 8, pp. 945-957, August 2006.

[16] R.O. Saber and R.M. Murray, ”Consensus protocols for networks of dynamic agents”, in *Proc. 2003 Am. Control Conf.*, 2003, pp. 951-956.

[17] R.O. Saber and R.M. Murray, ”Consensus problem in networks of agents with switching topology and time-delays”, *IEEE Trans. Automatic Control*, vol. 49, no. 9, pp. 1520-1533, Sept. 2004.

[18] R.O. Saber, ”Distributed Kalman Filter with Embedded Consensus Filters,”, *Proceedings of the 44th IEEE Conference on Decision and Control*, 2005.

[19] R.O. Saber, J.A. Fax and R.M. Murray, ”Consensus and Cooperation in Networked Multi-Agent Systems”, *Proceedings of the IEEE*, vol. 95, no. 1, pp. 215-233, Jan. 2007.

[20] E. Senata, ”Non-negative Matrices and Markov Chains,“ *Springer-Verlag*, 1973

[21] A. Tahbaz Salehi and A. Jadbabaie, ”Necessary and Sufficient Conditions for Consensus over random networks”, *IEEE Transactions on Automatic Control*, accepted, August 2007.

[22] A. Tahbaz Salehi and A. Jadbabaie, ”Consensus Over Ergodic Stationary Graph Processes,“ *IEEE Transactions on Automatic Control*, to appear

[23] J.N. Tsitsiklis, ”Problems in decentralized decision making and computation”, Ph.D. dissertation, Dept. Electr. Eng., Massachusetts Inst. Technol., Cambridge, MA, Nov. 1984.

[24] J.N. Tsitsiklis, D.P. Bertsekas and M. Athans, ”Distributed asynchronous deterministic and stochastic gradient optimization algorithms”, *IEEE Trans. Automatic Control*, vol. 31, no. 9, pp. 803-812, Sept. 1986.

[25] B.K. Sharma and C.L. Dewangan, “Fixed Point Theorem in Convex Metric Space,” *Novi Sad Journal of Mathematics*, vol. 25, no. 1, pp. 9-18, 1995.

[26] W. Takahashi, “A convexity in metric space and non-expansive mappings I,” *Kodai Math. Sem. Rep.*, 22(1970), 142-149.