The Asymptotic Consensus Problem on Convex Metric Spaces

Ion Matei and John S. Baras, *Life Fellow, IEEE*

***Abstract*—A consensus problem consists of a group of dynamic agents who seek to agree upon certain quantities of interest. The agents exchange information according to a communication network modeled as a directed time-varying graph and evolve in a convex metric space; a metric space endowed with a convex structure. In this paper we generalize the asymptotic consensus problem to convex metric spaces. Under weak connectivity assumptions, we show that if at each iteration an agent updates its state by choosing a point from a particular subset of the generalized convex hull generated by the agent's current state and the states of its neighbors, then agreement is achieved asymptotically. In addition, we present several examples of convex metric spaces and their corresponding agreement algorithms.**

***Index Terms*—Agreement, convex metric spaces, distributed algorithms, time varying graphs.**

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 [3], Tsitsikils *et al.* [31], [32] on asynchronous agreement problems and distributed computing. A theoretical framework for solving consensus problems was introduced by Olfati-Saber and Murray in [19] and [20], while Jadbabaie *et al.* studied alignment problems [9] for reaching an agreement. Relevant extensions of the consensus problem were done by Ren and Beard [24], by Moreau [15] or, more recently, by Nedic and Ozdaglar [17], [18].

Typically agents are connected via a network that changes with time due to link failures, packet drops, node failures, etc. Such variations in topology can happen randomly which motivates the investigation of consensus problems under a stochastic framework. Hatano and Mesbahi consider in [7] 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 [21], 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 [25], while in [26], 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 [11] and [12].

A convex metric space is a metric space endowed with a convex structure. In this paper we 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 generalized convex hull of a finite set of points. Tsitsiklis showed in [31] that, under some minimal connectivity assumptions on the communication network, if an agent updates its value by choosing a point (in $\mathbb{R}^n$) 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. The work in [15] already emphasized the central role played by convexity for proving convergence to consensus. We show that the same idea applies beyond the vector space of real numbers, and that the convex structure provides a systematic way to derive agreement algorithms. Generalizing the convex property to non-Euclidean spaces allows for dropping a number of smoothness assumptions on the dynamics. For example, unlike [15], no continuity on the maps generating new states is required. Extensions of the results shown in [15] and presented in [33], where the authors consider convexity in metric spaces more general than the standard Euclidean spaces,[1] and use orthogonal projections on geodesic segments to prove convergence, since they are able to easily characterize contractions of a line segment.

Manuscript received April 2, 2010; revised January 8, 2013, September 13, 2013, April 6, 2014, and August 5, 2014; accepted October 8, 2014. Date of publication October 14, 2014; date of current version March 20, 2015. This work was supported in part by the National Institute of Standards and Technology (NIST) via the ARRA Measurement Science and Engineering Fellowship Program Grant 70NANB10H026, through the University of Maryland, by the U.S. Air Force Office of Scientific Research MURI Grant FA9550-09-1-0538, by NSF Grant CNS-1035655, and by NIST Grant 70NANB11H148. Recommended by Associate Editor A. Papachristodoulou.

I. Matei is with the System Sciences Laboratory, Palo Alto Research Center (PARC), Palo Alto, CA 94304 USA (e-mail: ion.matei@parc.com).

J. S. Baras is with the Institute for Systems Research, Electrical and Computer Engineering Department, University of Maryland, College Park, MD 20742 USA (e-mail: baras@umd.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TAC.2014.2362988

[1]The authors consider CAT(0) metric spaces; spaces on which any two points x, y in a geodesic triangle (union of three geodesic segments) together with their comparison points $\bar{x}, \bar{y}$ in the comparison triangle (a triangle on $\mathbb{R}^2$ for which the distances between corners are the same as the distances between the corners of the geodesic triangle) satisfy $d(x,y) \leq \|\bar{x} - \bar{y}\|_2$.

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, possibly overlapping convex hulls of two finite sets of points. These properties constitute the base for proving convergence of the agreement algorithm. *Second*, we provide an upper bound on the (infinity) norm of the vector of distances between the values of the agents. We show that the agents asymptotically reach agreement, by showing that this upper bound asymptotically converges to zero. *Third*, we emphasize the relevance of our framework, by providing several examples of convex metric spaces and their corresponding agreement algorithms. An initial version of the concepts presented here were introduced in [10], where due to space limitations almost all proofs were omitted. The current paper refines the respective concepts, includes all the proofs and several new examples of convex metric spaces. In [13] and [14] we extended our work to the case where the communication between agents is based on a randomized gossip algorithm. As a result, the technical approach used for the convergence analysis is completely different, having to deal with the underlying stochastic framework of the problem; namely, stochastic calculus was used to analysis the dynamical properties of a set of stochastic differential equations driven by Poisson counters.

The paper is organized as follows. Section II introduces the main concepts related to 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 V we present several examples of convex metric spaces and their corresponding agreement algorithms.

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 without self loops, for which every edge corresponds to a *non-zero*, *non-diagonal* entry of W. We denote by $\mathbb{1}_{\{A\}}$ the indicator function of an event A. Given some set $\mathcal{X}$ 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 presents 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 [29], [30].

Definition 2.1 [30, pp. 142]: 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

$$d\left(u, \psi(x,y,\lambda)\right) \leq \lambda d(u,x) + (1-\lambda)d(u,y)$$

for all $x, y, u \in \mathcal{X}$ and $\lambda \in [0,1]$.

Definition 2.2 [30, pp.142]: The metric space $(\mathcal{X}, d)$ together with the convex structure ψ is called a *convex metric space*, and is denoted henceforth by the triplet $(\mathcal{X}, d, \psi)$.

Example 2.1: The most common convex metric space is $\mathbb{R}^n$ together with the Euclidean distance and convex structure given by the standard convex combination operation. Indeed, for any $x, y, z \in \mathbb{R}^n$ and $\lambda \in [0,1]$, it follows that $\|z - (\lambda x + (1-\lambda)y)\| = \|\lambda(z-x) + (1-\lambda)(z-y)\| \leq \lambda\|z-x\| + (1-\lambda)\|z-y\|$, where the last inequality followed from the convexity of the norm operator.

Example 2.2 [30]: Let I be the unit interval [0,1] and $\mathcal{X}$ be the family of closed intervals, i.e., $\mathcal{X} = \{[a,b] | 0 \leq a \leq b \leq 1\}$. 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 $\mathcal{X}$ by the Hausdorff distance, i.e.,

$$d(I_i, I_j) = \max\left\{|a_i - a_j|, |b_i - b_j|\right\}.$$

Then $(\mathcal{X}, d, \psi)$ is a convex metric space.

More examples can be found in [29] and [30] and the references therein. In Section V, we introduce additional examples of convex metric spaces.

Definition 2.3 [30, pp. 144]: A convex metric space $(\mathcal{X}, d, \psi)$ is said to have *Property* (C) if every bounded decreasing net of nonempty closed convex subsets of $\mathcal{X}$ has a nonempty intersection.

Fortunately, convex matric spaces satisfying *Property* (C) are not that rare. Indeed, by Smulian's Theorem [5, p. 443], every weakly compact convex subset of a Banach space has *Property* (C).

The following definition introduces the notion of a convex set in a convex metric space.

Definition 2.4 [30, p. 143]: Let $(\mathcal{X}, d, \psi)$ 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 $\forall \lambda \in [0,1]$.

Let $\mathcal{P}(\mathcal{X})$ be the set of all subsets of $\mathcal{X}$. We define the set valued mapping $\tilde{\psi} : \mathcal{P}(\mathcal{X}) \to \mathcal{P}(\mathcal{X})$ as

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

where A is an arbitrary subset of $\mathcal{X}$.

In Proposition 1 [30, p. 143] 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 [29, p. 11]: Let $(\mathcal{X}, d, \psi)$ be a convex metric space. 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 $co(A)$.

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 [29] that the set sequence $\{A_m\}_{m\geq 0}$ is increasing and $\limsup_{m\to\infty} A_m$ exists, and $\limsup_{m\to\infty} A_m = \liminf_{m\to\infty} A_m = \lim_{m\to\infty} A_m = \bigcup_{m=0}^{\infty} A_m$.

Proposition 2.1 [29, pp. 12]: Let $(\mathcal{X}, d, \psi)$ be a convex metric space. The convex hull of a set $A \subset \mathcal{X}$ is given by

$$co(A) = \lim_{m\to\infty} A_m = \bigcup_{m=0}^{\infty} A_m.$$

It follows immediately from above that if $A_{m+1} = A_m$ for some m, then $co(A) = A_m$.

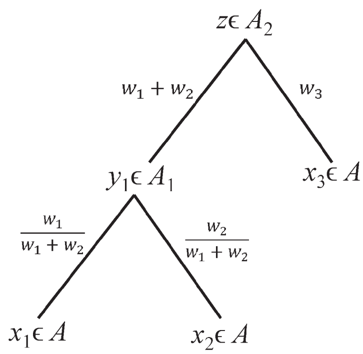

Fig. 1. Decomposition of a point $z \in A_3$ with $A = \{x_1, x_2, x_3\} \in \mathbb{R}^n$.

B. On the Convex Hull of a Finite Set

For a positive integer n, let $A = \{x_1, \dots, x_n\}$ be a finite set in a convex metric space $(\mathcal{X}, d, \psi)$, with convex hull $co(A)$ and let z belong to $co(A)$. By Proposition 2.1 it follows that there exists a positive integer m such that $z \in A_m$. But since $A_m = \tilde{\psi}(A_{m-1})$ it follows that there exist two *parents* $z_1, z_2 \in A_{m-1}$ and $\lambda_{(1,2)} \in [0,1]$ such that $z = \psi(z_1, z_2, \lambda_{(1,2)})$. Similarly, there exist z_3, z_4 (parents of z_1), and z_5, z_6 (parents of z_2) belonging to 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 parents, we obtain a *tree representation* of z, which has as leaves the elements of the set A. Using a graph theory terminology, z can be viewed as the root of a weighted binary tree with leaves belonging to the set A. Each node α (except the leaves) has two parents α_1 and α_2, and are related through the operator ψ in the sense $\alpha = \psi(\alpha_1, \alpha_2, \lambda)$ for some $\lambda \in [0,1]$. The weights of the edges connecting α with α_1 and α_2 are given by λ and $1 - \lambda$ respectively.

Remark 2.1: We would like to point out that a point z in $co(A)$ *does not necessarily have a unique tree representation*. Indeed, assuming that $A = \{x_1, x_2, x_3\} \in \mathbb{R}^n$, any point z in $co(A)$ can be written as $z = w_1 x_1 + w_2 x_2 + w_3 x_3$, where $w_i \geq 0$, $i \in \{1,2,3\}$ and $\sum_{i=1}^{3} w_i = 1$. If in particular we choose $A = \{1, 5, 2\}$ and $z = 3$, we can find an infinite number of tree representations of the type shown in Fig. 1, with weights given by $w_1 = 1/2 - w_3/4$, $w_2 = 1/2 - 3w_3/4$ and $w_3 \in [0, 2/3]$, such that $w_1 + 5w_2 + 2w_3 = 3$.

From the above discussion, we note that for any point $z \in co(A)$ there exists 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*, i.e., a full binary tree in which all leaves are at the same depth. That is because not all points in A are necessarily at the same level in the tree representation.

Consider now *a particular* tree decomposition of a point $z \in co(A)$ and let m be the height of the respective tree. For *this* tree decomposition, 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_lth path from the root z to the node x_i, for $l = 1, \dots, n_i$. We formally describe the paths from the root z to x_i as the set

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

where $\{y_{i_l j}\}_{j=0}^{m_{i_l}}$ is the set of points forming the i_lth 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 $(y_{i_l,j-1}, y_{i_l,j})$.

Definition 2.6: We define the *weight* of the point x_i in the *creation* of point z as the quantity

$$\mathcal{W}_{x_i}^z \triangleq \sum_{l=1}^{n_i} \prod_{j=1}^{m_{i_l}} \lambda_{i_l,j}.$$

It can be checked that the weights corresponding to the points $\{x_1, \dots, x_n\}$ sum up to one, i.e.,

$$\sum_{i=1}^{n} \mathcal{W}_{x_i}^z = 1.$$

Remark 2.2: We would like to emphasize that the weights corresponding to the points in A depend on a particular tree representation. A different tree representation implies a different set of weights. The weights definition was inspired by the convex hull definition in $\mathbb{R}^n$. Indeed, in the case of the example presented in Fig. 1, it can be noticed that $\mathcal{W}_{x_i}^z = w_i$, $i \in \{1,2,3\}$.

Definition 2.7: Given a small enough positive scalar $\varepsilon < 1$, we define the following subset of $co(A)$ consisting of all points in $co(A)$ that have at least one tree representation whose weights corresponding to the points in A are lower bounded by ε, i.e.,

$$co_\varepsilon(A) \triangleq \{z \in co(A) \mid \text{there exists at least one tree representation of } z \text{ such that } \mathcal{W}_{x_i}^z \geq \varepsilon,\ \forall x_i \in A\}.$$

Remark 2.3: By a *small enough* ε, we understand a value such that the inequality $\mathcal{W}(P_{z,x_i}) \geq \varepsilon$ is satisfied for all i. Obviously, for n agents ε needs to satisfy

$$\varepsilon \leq \frac{1}{n}$$

but usually we would want to choose a value much smaller then $1/n$ since this implies a richer set $co_\varepsilon(A)$.

Remark 2.4: We can iteratively generate points for which we can make sure that they belong to $co_\varepsilon(A)$, with $A = \{x_1, \dots, x_n\}$. Given a set of positive scalars $\{\lambda_1, \dots, \lambda_{n-1}\} \in (0,1)$, consider the iteration

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

We note that y_n is guaranteed to belong to the interior of $co(A)$. In addition, if we impose the condition

$$\varepsilon^{\frac{1}{n-1}} \leq \lambda_i \leq \frac{1-(n-1)\varepsilon}{1-(n-2)\varepsilon},\ i = 1, \dots, n-1$$

and ε respects the inequality

$$\varepsilon^{\frac{1}{n-1}} \leq \frac{1-(n-1)\varepsilon}{1-(n-2)\varepsilon} \tag{1}$$

then $y_n \in co_\varepsilon(A)$. We should note that for any $n \geq 2$ we can find a small enough value of ε such that inequality (1) is satisfied.

Example 2.3: To give an intuition on the meaning of the set $co_\varepsilon(A)$, we make a parallel with $\mathbb{R}^n$. Given that $A = \{x_1, \ldots, x_p\} \subset \mathbb{R}^n$, for some positive integer p, the set $co_\varepsilon(A)$ can be formally deduced to be equal to

$$co(A)_\varepsilon = \left\{ z \in \mathbb{R}^n \mid z = \sum_{i=1}^{p} w_i x_i,\ \sum_{i=1}^{p} w_i = 1,\ w_i \geq \varepsilon \right\}$$

which basically states that $co_\varepsilon(A)$ is a subset of the relative interior of the convex hull of A. Indeed, let $\mathcal{K}_\varepsilon(A) \stackrel{\Delta}{=} \{y \in \mathbb{R}^n | z = \sum_{i=1}^{p} w_i x_i, \sum_{i=1}^{p} w_i = 1, w_i \geq \varepsilon\}$. From Definition 2.7, z admits at least one tree representation with weights $\mathcal{W}_{x_i}^z$ lower bounded by ε. In addition, from the definition of the convex structure for $\mathcal{X} = \mathbb{R}^n$ (see Example 2.1), the point z can be written as $z = \sum_{i=1}^{p} \mathcal{W}_{x_i}^z x_i$, with $\sum_{i=1}^{p} \mathcal{W}_{x_i}^z = 1$ and $\mathcal{W}_{x_i}^z \geq \varepsilon$. Hence, $z \in \mathcal{K}_\varepsilon(A)$ and therefore $co_\varepsilon(A) \subset \mathcal{K}_\varepsilon(A)$. Let now z be a point in $\mathcal{K}_\varepsilon(A)$. Then, there exits $w_i \geq \varepsilon$ with $\sum_{i=1}^{p} w_i = 1$, such that $z = \sum_{i=1}^{p} w_i x_i$. Equivalently, the point z can be iteratively written as $y_{i+1} = \lambda_i y_i + (1 - \lambda_i) x_{i+1}$ for $i = 1, \ldots, p-1$, with $y_1 = x_1$, where $z = y_p$ and $\lambda_i = (w_1 + \ldots + w_i)/(w_1 + \ldots + w_{i+1})$. But this implies that z admits a tree decomposition (similar to the one depicted in Fig. 1), with $\mathcal{W}_{x_i}^z = w_i$, and therefore $z \in co_\varepsilon(A)$. Thus, $\mathcal{K}_\varepsilon(A) = co_\varepsilon(A)$.

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

Proposition 2.2: 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 $\varepsilon < 1$ be a positive scalar small enough.

(a) If $x \in co(X)$ and $y \in co(Y)$, then there exists a set of $n_x \times n_y$ non-negative scalars λ_{ij}, with $\sum_{i=1}^{n_x} \sum_{j=1}^{n_y} \lambda_{ij} = 1$, such that

$$d(x, y) \leq \sum_{i=1}^{n_x} \sum_{j=1}^{n_y} \lambda_{ij} d(x_i, y_j). \tag{2}$$

(b) If $x \in co_\varepsilon(X)$, $y \in co_\varepsilon(Y)$, then here exists a set of $n_x \times n_y$ non-negative scalars λ_{ij}, with $\sum_{i=1}^{n_x} \sum_{j=1}^{n_y} \lambda_{ij} = 1$, such that

$$\lambda_i \geq \varepsilon \text{ and } \lambda_{ij} \geq \varepsilon^2,\ \forall i, j. \tag{3}$$

(c) If the assumptions of (b) hold and in addition $X \cap Y \neq \emptyset$, then

$$\sum_{i=1}^{n_x} \sum_{j=1}^{n_y} \lambda_{ij} \mathbb{1}_{\{d(x_i, y_j) \neq 0\}} \leq 1 - \varepsilon^2 \tag{4}$$

where the coefficients λ_{ij} are the same as in (3).

Proof:

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

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

Recursively decomposing z_1, z_2 and their *parents*, and applying the convex structure definition, it can be easily argued that there exists a set of n_x positive scalars λ_i, summing up to one, such that

$$d(x, y) \leq \sum_{i=1}^{n_x} \lambda_i d(x_i, y). \tag{5}$$

In fact from Definition 2.6, it turns out that the scalars λ_i correspond to the weights $\mathcal{W}_{x_i}^x$, associated to this particular tree decomposition. We know repeat the above argument and upper-bound each of the terms $d(x_i, y)$ that appear in the inequality (5). We obtain that for each i, there exist n_y scalars $\mu_j \geq 0$ with $\sum_{j=1}^{n_y} \mu_j = 1$, so that

$$d(x_i, y) \leq \sum_{j=1}^{n_y} \mu_j d(x_i, y_j),\ \forall i. \tag{6}$$

Combining (5) and (6) 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 $\lambda_{ij} = \lambda_i \mu_j \geq 0$ and $\sum_{i=1}^{n_x} \sum_{j=1}^{n_y} \lambda_{ij} = 1$.

(b) Since $x \in co_\varepsilon(X)$ and $y \in co_\varepsilon(Y)$, the points x and y admit tree decompositions, where the weights on the edges are denoted by $\mathcal{W}_{x_i}^x$ and $\mathcal{W}_{y_i}^y$, respectively. In addition, the weights satisfy the inequalities $\mathcal{W}_{x_i}^x \geq \varepsilon$ and $\mu_j = \mathcal{W}_{y_j}^y \geq \varepsilon$, for all i and j. But since $\mathcal{W}_{x_i}^x$ and $\mathcal{W}_{y_j}^y$ play the role of λ_i and μ_j, respectively, it follows that $\lambda_i \geq \varepsilon$ and $\mu_j \geq \varepsilon$, and consequently $\lambda_{ij} \geq \varepsilon^2$.

(c) If $X \cap Y \neq \emptyset$ then there exists at least one pair (i, j) so that $d(x_i, y_j) = 0$. But since from part (b) $\lambda_{ij} \geq \varepsilon^2$, the inequality (4) follows. ■

III. Problem Formulation and Statement of the Main Result

We consider a convex metric space $(\mathcal{X}, d, \psi)$ and a set of n agents indexed by i which have states taking values in $\mathcal{X}$. Denoting by k the time index, the agents exchange information based on a communication network modeled by a *directed*, *time varying* graph $G(k) = (V, E(k))$, where V is the finite set of vertices (the agents) and $E(k)$ is the set of edges. An edge (communication link) $e_{ij}(k) \in E(k)$ exists if node i receives information from node j. Each agent has an initial value in $\mathcal{X}$ and, at each subsequent time-slot, it is adjusting its value based on the observations about the values of its 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

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

Similar to the communication models used in [2], [16], [32], we impose weak assumptions on the connectivity of the communication graph $G(k)$. Basically these assumptions consist of having the communication graph connected *infinitely often* and having *bounded intercommunication intervals* between neighboring nodes.

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

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

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 its 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 ... 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

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

Let $\mathcal{N}_i(k)$ denote the communication neighborhood of agent i, which contains all nodes sending information to i at time k, i.e., $\mathcal{N}_i(k) = \{j | e_{ij}(k) \in E(k)\} \cup \{i\}$, 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), and by $A(k) \triangleq \{x_i(k), i = 1, \ldots, n\}$ the set of all states of the agents.

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

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

$$x_i(k+1) \in co_\varepsilon\left(A_i(k)\right),\ \forall\, i$$

then they asymptotically reach consensus, i.e.,

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

The above theorem gives sufficient conditions so that the distances between the state of the agents converge to zero. The next corollary, introduces an additional condition so that the agents converge to the same value as well. This additional condition is similar to the completeness property of metric spaces.

Corollary 3.1: Let Assumptions 3.1 and 3.2 hold for $G(k)$ and let $\varepsilon < 1$ be a positive scalar sufficiently small. If the agents act on a convex metric space satisfying *Property* (C) and update their states according to the scheme

$$x_i(k+1) \in co_\varepsilon\left(A_i(k)\right),\ \forall\, i$$

then there exists $x^* \in \mathcal{X}$ such that

$$\lim_{k\to\infty} d\left(x_i(k), x^*\right) = 0,\ \forall\, i.$$

We will give the proofs for both Theorem 3.1 and Corollary 3.1 in the subsequent section.

Remark 3.1: A procedure for generating points that are guaranteed to belong to $co_\varepsilon(A_i(k))$ is described in Remark 2.4. The idea of picking $x_i(k+1)$ from $co_\varepsilon(A_i(k))$ rather than $co(A_i(k))$ is in the same spirit as the assumption imposed on the non-zero consensus weights in [2], [16], and [31], i.e., they are assumed lower bounded by a positive, sub-unitary scalar. Setting $x_i(k+1) \in co(A_i(k))$ 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, $x_i(k+1) \in co(A_i(k))$, $i = 1,2$ for all $k \geq 0$. It can be easily argued that

$$d\left(x_1(k+1), x_2(k+1)\right) \leq \left(\lambda(k)\left(1-\mu(k)\right) + \mu(k)\left(1-\lambda(k)\right)\right) d\left(x_1\left(k, x_2(k)\right)\right). \quad (7)$$

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 (7) 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 for 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 for the proof of Theorem 3.1 and Corollary 3.1. In what follows, in addition to indices referring to agents, we introduce indices that refer to pairs of agents. To differentiate between them, we depict the later with a bar on top.

Lemma 4.1: Given a small enough positive scalar $\varepsilon < 1$, assume that agents update their states according to the scheme $x_i(k+1) \in co_\varepsilon(A_i(k))$, 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 $N = n(n-1)/2$. Then, there exists a time varying $N \times N$ matrix $\mathbf{W}(k)$, such that

$$\mathbf{d}(k+1) \leq \mathbf{W}(k)\mathbf{d}(k),\ \mathbf{d}(0) = \mathbf{d}_0 \tag{8}$$

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

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

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

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

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

where η is the same as in part (a).

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

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

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

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

$$d\left(x_i(k+1), x_j(k+1)\right) \leq \sum_{p\in\mathcal{N}_i(k), q\in\mathcal{N}_j(k)} w_{pq}^{ij}(k) d\left(x_p(k), x_q(k)\right),\ i \neq j \tag{9}$$

for some $w_{pq}^{ij}(k) \geq 0$ with $\sum_{p\in\mathcal{N}_i(k), q\in\mathcal{N}_j(k)} w_{pq}^{ij}(k) = 1$. 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 (8) follows. We continue with proving the properties of matrix $\mathbf{W}(k)$.

(a) Since all $w_{pq}^{ij}(k) \geq 0$ 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 (b) of Proposition 2.2, there exists $\eta \triangleq \varepsilon^2$ 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 (9), and therefore $\mathbf{W}(k)$ has positive diagonal entries.

(b) Follows from part (c) of Proposition 2.2, with $\eta = \varepsilon^2$.

(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 (9). 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 (9) 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. ■

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 that 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

$$\bar{E}(k) = \{((i,j),(p,q))\,|(i,p) \in E(k), (j,q) \in E(k),\ i \neq j,\ p \neq q\}. \tag{10}$$

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 (strongly) connected, where

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

(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 the latest) at time $k + \bar{B} - 1$ for any $k \geq 0$.

Proof: It is not difficult to observe that similar to (10), $\bar{E}_\infty$ is given by

$$\bar{E}_\infty = \{((i,j),(p,q)) \mid (i,p) \in E_\infty, (j,p) \in E_\infty, p \neq q, i \neq j\}. \tag{11}$$

(a) Showing that $(\bar{V}, \bar{E}_\infty)$ is (strongly) connected is equivalent to showing that for any (i,j), all other pairs (p,q) have paths connecting to it. Let $\mathcal{N}_{i,\infty}$ be the inward neighborhood of node i in the graph (V, E_∞), i.e., $\mathcal{N}_{i,\infty} = \{j|(j,i) \in E_\infty\}$. We can also refer to $\mathcal{N}_{i,\infty}$ as the set of nodes connected to i in one hop. The nodes connected to nodes i in t hops are given by the set $\mathcal{N}_{i,\infty}^t = \bigcup_{j\in\mathcal{N}_{i,\infty}^{t-1}}$, where $\mathcal{N}_{i,\infty}^1 = \mathcal{N}_{i,\infty}$. Since (V, E_∞) is assumed strongly connected, $\mathcal{N}_{i,\infty}^{n-1} = \{1,2,\ldots,n\}$, for all i.

The one hop inward neighborhood of an arbitrary node (i,j) in the graph $(\bar{V}, \bar{E}_\infty)$ is given by $\bar{\mathcal{N}}_{(i,j),\infty} = \mathcal{N}_{i,\infty} \times \mathcal{N}_{j,\infty}$, where $\mathcal{N}_{i,\infty} \times \mathcal{N}_{j,\infty} = \{(p,q)|p \in \mathcal{N}_{i,\infty}, q \in \mathcal{N}_{j,\infty}, p \neq q\}$. The nodes connected to (i,j) in t hops are given by $\bar{\mathcal{N}}_{(i,j),\infty}^t = \bigcup_{(p,q)\in\bar{\mathcal{N}}_{(i,j),\infty}^{t-1}} \mathcal{N}_{(p,q),\infty}$. But $\bar{\mathcal{N}}_{(i,j),\infty}^t$ can be equivalently written as $\bar{\mathcal{N}}_{(i,j),\infty}^t = \mathcal{N}_{i,\infty}^t \times \mathcal{N}_{j,\infty}^t$. Therefore, in $n-1$ hops (i,j) is connected to all the nodes $\{(p,q)|p,q \in \{1,2,\ldots,n\}, p \neq q\}$, which corresponds to all nodes in the graph $(\bar{V}, \bar{E}_\infty)$. But since (i,j) was chosen arbitrarily, it follows that $(\bar{V}, \bar{E}_\infty)$ is strongly 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 $k \geq 0$

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

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 $k\geq s$. From the properties of $\mathbf{W}(k)$, it follows that $\Phi(k,s)$ is a non-negative matrix with positive diagonal entries and $\|\Phi(k,s)\|_\infty\leq 1$ for any $k\geq s$. The following result, whose proof can be found in the Appendix section, introduces a property of the entries of $\Phi(k,s)$ used in the proof of the main result.

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 exists a row index $\bar{i}^*$ such that

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

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 Proposition 4.1.

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

$$\left[\Phi\left(s+(N-1)\bar{B}-1,s\right)\right]_{ij}\geq\eta^{(N-1)\bar{B}}\ \forall s,i,j$$

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 Proposition 4.1.

Proof: By Proposition 4.1 and Lemma 4.1 all the assumptions of Lemma 2, [16] are satisfied, from which the result follows. ■

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

A. Proof of Theorem 3.1

From Lemma 4.1, there exists a non-negative matrix $\mathbf{W}(k)$ such 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:

$$\|\mathbf{d}(k+1)\|_\infty\leq\|\mathbf{d}(k)\|_\infty\,,\ \text{for } k\geq 0. \tag{12}$$

Let $\bar{B}_0\triangleq(N-1)\bar{B}-1$, where $\bar{B}$ is the positive integer from part (b) of Proposition 4.1. In the following we show that the sums of the elements of each of the rows 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

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

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

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

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

$$\begin{aligned}\sum_{\bar{j}=1}^{N}\left[\Phi(s+2\bar{B}_0,s)\right]_{\bar{i}\bar{j}}&\leq\sum_{\bar{j}=1,\bar{j}\neq\bar{j}^*}^{N}\left[\Phi(s+2\bar{B}_0,s+\bar{B}_0)\right]_{\bar{i}\bar{j}}\\&\quad+\left[\Phi(s+2\bar{B}_0,s+\bar{B}_0)\right]_{\bar{i}\bar{j}^*}(1-\eta^{\bar{B}_0})\\&=\sum_{\bar{j}=1}^{N}\left[\Phi(s+2\bar{B}_0,s+\bar{B}_0)\right]_{\bar{i}\bar{j}}-\left[\Phi(s+2\bar{B}_0,s+\bar{B}_0)\right]_{\bar{i}\bar{j}^*}\eta^{\bar{B}_0}.\end{aligned}$$

By Corollary 4.1, it follows that:

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

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

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

Therefore

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

It follows that:

$$\|\mathbf{d}(t_k)\|_\infty\leq\left(1-\eta^{2\bar{B}_0+1}\right)^k\|\mathbf{d}(0)\|_\infty\,,\ \forall\, k\geq 0 \tag{13}$$

where $t_k=2k\bar{B}_0$ which shows that the subsequence $\{\|\mathbf{d}(t_k)\|_\infty\}_{k\geq 0}$ asymptotically converges to zero. Combined with inequality (12), we further obtain that the sequence $\{\|\mathbf{d}(k)\|_\infty\}_{k\geq 0}$ asymptotically converges to zero. Therefore the agents asymptotically reach consensus.

B. Proof of Corollary 3.1

The main idea of the proof consists of showing that the set $co(A(k))$, where $A(k)=\{x_i(k),i=1,\ldots,n\}$, converges to a set containing one point.

We first note that since $A_i(k)\subseteq A(k)$ it can be easily argued that $co(A_i(k))\subseteq co(A(k))$, for all i and k. Also, since $co_\varepsilon(A_i(k))\subseteq co(A_i(k))$ it follows that $co_\varepsilon(A_i(k))\subseteq co(A(k))$ and consequently $x_i(k+1)\in co(A(k))$. Therefore, we have that $co(A(k+1)) \subseteq co(A(k))$ for all k and from the theory of limits of sequences of sets, it follows that:

$$\liminf co\,(A(k)) = \limsup co\,(A(k)) = \lim co\,(A(k)) = A_\infty$$

where $A_\infty = \bigcap_{k\geq 0} co(A(k))$. We denote the diameter of the set $A(k)$ by

$$\delta\,(A(k)) = \sup\,\{d(x,y) \mid x,y \in A(k)\}$$

and by Proposition 2 of [29], we have that

$$\delta\,(co\,(A(k))) = \delta\,(A(k))\,.$$

From Theorem 3.1 we have that

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

and consequently

$$\lim_{k\to\infty} \delta\,(A(k)) = \lim_{k\to\infty} \delta\,(co\,(A(k))) = 0$$

which also means that

$$\delta(A_\infty) = 0.$$

But since the convex metric space on which the algorithm operates satisfies *Property* (C), and the sets $A(k)$ are bounded (they have bounded diameter) and closed (by construction), it follows that the set A_∞ is non-empty. That is, A_∞ contains only one point, say $x^* \in \mathcal{X}$, or $A_\infty = co(x^*)$, or

$$\lim_{k\to\infty} co\,(A(k)) = co(x^*).$$

But since $x_i(k+1) \in co_\varepsilon(A_i(k)) \subseteq co(A(k))$ for all i,k it follows that:

$$\lim_{k\to\infty} d\,(x_i(k), x^*) = 0,\ \forall\, i$$

i.e., the states of the agents converge to the same point $x^* \in \mathcal{X}$.

V. Examples of Convex Metric Spaces and Their Consensus Algorithms

In this section, we introduce a number of convex metric spaces on which we define consensus algorithms based on their convex structure mappings. The following examples show how the generalized notion of convexity provides the means to design agreement algorithms that go beyond the standard linear consensus algorithms studied in the literature. Numerical simulations of the agreement algorithms described in what follows are shown at the end of this section. As in the previous sections, we assume that a group of N agents, indexed by i takes values in the set $\mathcal{X}$.

A. Consensus on the Set of Positive Real Vectors

The first example concerns a subset of the vector space of real numbers, namely the cone defined by vectors with positive entries. It was inspired by a consensus algorithm defined on a loglinear system introduced in [33]. The particular metric space in this example is a CAT(0) (see the the first section for the definition), and interestingly it is also a convex metric space. Let $\mathcal{X}$ be the set of positive real vectors, that is, $\mathcal{X} = \{\mathbf{x} \in \mathbb{R}^n | x_l > 0, l = 1\ldots n\}$, where the subscript l refers to the lth entry of vector $\mathbf{x}$. We define the metric $d(\mathbf{x},\mathbf{y}) = \|\ln(\mathbf{x}/\mathbf{y})\|$, where $\|\cdot\|$ denotes the standard Euclidean norm, $\mathbf{x}/\mathbf{y} \triangleq (x_l/y_l)$ and $\ln(\mathbf{x}) \triangleq (\ln(x_l))$. In addition, we consider the mapping $\psi : \mathcal{X}\times\mathcal{X}\times[0,1] \to \mathcal{X}$ defined by

$$\psi(\mathbf{x},\mathbf{y},\lambda) = \mathbf{x}^\lambda \mathbf{y}^{1-\lambda} \tag{14}$$

where $\mathbf{x}^\lambda \triangleq (x_l^\lambda)$ and $\mathbf{xy} = (x_l y_l)$.

Proposition 5.1: The mapping ψ defined in (14) is a convex structure on $(\mathcal{X}, d)$.

Proof: Let $\mathbf{z}$ be a point in $\mathcal{X}$. Using the definition of the mapping ψ together with the four operations defined above, we obtain

$$\ln\left(\frac{\mathbf{z}}{\psi(\mathbf{x},\mathbf{y},\lambda)}\right) = \left(\ln\left(\frac{z_l}{(x_l^\lambda y_l^{1-\lambda})}\right)\right) = \left(\ln(z_l) - \ln\left((x_l^\lambda y_l^{1-\lambda})\right)\right)$$
$$= \left(\lambda \ln\left(\frac{z_l}{x_l}\right) - (1-\lambda)\ln\left(\frac{z_l}{y_l}\right)\right) = \lambda \ln\left(\frac{\mathbf{z}}{\mathbf{x}}\right) + (1-\lambda)\ln\left(\frac{\mathbf{z}}{\mathbf{y}}\right)$$

and the result follows from the convexity property of the standard Euclidean norm. ■

Using Remark 2.4, we obtain the following agreement algorithm:

$$\mathbf{x}_i(k+1) = \prod_{j\in\mathcal{N}_i(k)} \mathbf{x}_j^{w_{ij}(k)} \tag{15}$$

where $\sum_{j\in\mathcal{N}_i(k)} w_{ij}(k) = 1$ for all i,k. Assuming that the weights w_{ij} are uniformly lower bounded, Theorem 3.1 ensures that by following iteration (15), the agents converge to the same value. The above agreement algorithm is rather an academic example since agreement can be obtained using the standard convex combination operator as a convex structure. It is however an interesting nonlinear agreement algorithm based on the generalized notion of convexity.

B. Consensus on the Set of Discrete Random Variables

Let s be a positive integer, let $S = \{1,2,\ldots,s\}$ be a finite set and let $(\Omega,\mathcal{F},\mathcal{P})$ be a probability space. We consider the set $\mathcal{X}$ to be the space of discrete measurable functions (random variables) on $(\Omega,\mathcal{F},\mathcal{P})$ with values in S. We choose as metric the expected value of the *discrete metric*, that is

$$d(X,Y) = E\,[\rho(X,Y)] \tag{16}$$

where $\rho : \mathbb{R}\times\mathbb{R} \to \{0,1\}$ is the discrete metric given by

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

and where the expectation operator is defined with respect to the probability space $(\Omega,\mathcal{F},\mathcal{P})$. Since for all $X,Y,Z \in \mathcal{X}$, the mapping d satisfies the following properties:

(a) $d(X,Y) = 0$ if and only if $X = Y$ with probability one;
(b) $d(X,Z) + d(Y,Z) \geq d(X,Y)$;
(c) $d(X,Y) = d(Y,X)$,
(d) $d(X,Y) \geq 0$;

it is indeed a metric on $\mathcal{X}$. Note that an equivalent definition for d is $d(X,Y) = E[\mathbb{1}_{\{X\neq Y\}}] = Pr(X \neq Y)$, where $\mathbb{1}_{\{X\neq Y\}}$ is the indicator function of the event $\{X \neq Y\}$. We now introduce the convex structure on $\mathcal{X}$. Let $\theta \in \{1,2\}$ be an independent random variable defined on the probability space $(\Omega, \mathcal{F}, \mathcal{P})$, with probability mass function $Pr(\theta = 1) = \lambda$ and $Pr(\theta = 2) = 1 - \lambda$, where $\lambda \in [0,1]$. We define the mapping $\psi : \mathcal{X} \times \mathcal{X} \times [0,1] \to \mathcal{X}$ given by

$$\psi(X_1, X_2, \lambda) = \mathbb{1}_{\{\theta=1\}} X_1 + \mathbb{1}_{\{\theta=2\}} X_2,\ \forall X_1, X_2 \in \mathcal{X}, \lambda \in [0,1]. \tag{17}$$

The following proposition shows that the mapping ψ is in fact a convex structure, and therefore $(\mathcal{X}, d, \psi)$ is a convex metric space.

Proposition 5.2: The mapping ψ is a convex structure on $(\mathcal{X}, d)$.

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

$$\begin{aligned} d\left(U, \psi(X_1, X_2, \lambda)\right) &= E\left[\rho\left(U, \psi(X_1, X_2, \lambda)\right)\right] \\ &= E\left[E\left[\rho\left(U, \psi(X_1, X_2, \lambda)\right) | U, X_1, X_2\right]\right] \\ &= E\left[E\left[\rho\left(U, \mathbb{1}_{\{\theta=1\}} X_1 + \mathbb{1}_{\{\theta=2\}} X_2\right)\right] | U, X_1, X_2\right] \\ &= E\left[\lambda \rho(U, X_1) + (1-\lambda)\rho(U, X_2)\right] \\ &= \lambda d(U, X_1) + (1-\lambda) d(U, X_2). \end{aligned}$$

■

We showed so far that the set of discrete random variables together with the metric (16) and the mapping (17) form a convex metric space. This would be enough to derive an agreement algorithm. As it turn out, we can give more insight into the convex sets defined on this particular convex metric space. The next theorem (whose proof can be found in the Appendix section) characterizes the convex hull of a finite set in $\mathcal{X}$.

Theorem 5.1: Let n be a positive integer and let $A = \{X_1, \ldots, X_n\}$ be a set of points in $\mathcal{X}$. Consider an independent random variable θ defined on the probability space $(\Omega, \mathcal{F}, \mathcal{P})$ and taking values in the finite set $\{1, \ldots, n\}$, with probability measure given by $Pr(\omega : \theta(\omega) = i) = w_i$, for some non-negative scalars w_i, where $\sum_{i=1}^N w_i = 1$. Then the convex hull of the set A is given by

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

The next corollary characterizes the interior of the convex hull of a set A of $\mathcal{X}$.

Corollary 5.1: Let n be a positive integer and let $A = \{X_1, \ldots, X_n\}$ be a set of points in $\mathcal{X}$. Consider an independent random variable θ taking values in the finite set $\{1, \ldots, n\}$, with probability measure given by $Pr(\omega : \theta(\omega) = i) = w_i$, for some non-negative scalars w_i, where $\sum_{i=1}^N w_i = 1$. Then

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

Proof: Follows immediately from Definition 2.7 and Theorem 5.1. ■

Using the above corollary, the agreement algorithm on the set of discrete random variables can be written as

$$X_i(k+1) = \sum_{j \in \mathcal{N}_i(k)} \mathbb{1}_{\{\theta_i(k)=j\}} X_j(k) \tag{18}$$

where each $\theta_i(k)$ is an independent random variable taking values in the set $\mathcal{N}_i(k)$, with probability mass function $Pr(\theta_i(k) = j) = w_{ij}(k)$ for all $j \in \mathcal{N}_i(k)$. In addition, $\sum_{j\in\mathcal{N}_i(k)} w_{ij}(k) = 1$ and $w_{ij} \geq \varepsilon$ for all i and k.

Note that above algorithm ensures that $X_i(k+1) \in co_\varepsilon(A_i(k))$ for all i and k, and therefore, by virtue of Theorem 3.1 we have that

$$\lim_{k\to\infty} Pr\left(X_i(k) \neq X_j(k)\right) = \lim_{k\to\infty} d\left(X_i(k), X_j(k)\right) = 0,\ \forall i, j. \tag{19}$$

Remark 5.1: Using a probability theory terminology, (19) shows that the random sequences $\{X_i(k)\}_{k\geq 0}$ and $\{X_j(k)\}_{k\geq 0}$ converge to the same value in probability, for all $i \neq j$. Indeed, defining the set $B_k(\epsilon) \triangleq \{\omega : \max_{i\neq j} |X_i(k) - X_j(k)| > \epsilon\}$, we have

$$\begin{aligned} Pr\left(B_k(\epsilon)\right) &= Pr\left(\bigcup_{i\neq j} \{\omega : |X_i(k) - X_j(k)| > \epsilon\}\right) \\ &\leq \sum_{i\neq j} Pr\left(|X_i(k) - X_j(k)| > \epsilon\right) \\ &\leq \sum_{i\neq j} Pr\left(X_i(k) \neq X_j(k)\right) \end{aligned} \tag{20}$$

and convergence in probability follows. It turns out, that convergence is achieved in the almost sure sense, as well. This follows from the Borel-Cantelli Lemma [6], by recalling that $d(X_i(k), X_j(k)) = Pr(X_i(k) \neq X_j(k))$ converge at least geometrically to zero for all i, j, and therefore

$$\sum_{k\geq 0} Pr\left(B_k(\epsilon)\right) < \infty.$$

The above algorithm can be applied to problems, where the agents' goal is to agree on a value in a discrete set. Examples of such problems can be found in social networks, where individuals try to agree on the fit of potential new hires or the merits of politicians. The agreement algorithm (18) may appear to be connected with the randomized gossip algorithms studied in [4] and [28]. There are however some important differences between (18) and the respective algorithms. In our case, the randomization comes from applying the convex structure operator, rather than from picking agents at random or from the random clocks associated to the agents. Also, the gossip algorithms are defined on the set of real numbers, rather than discrete sets. Using similar notations as in the case of algorithm (18), a synchronous version of a randomized gossip algorithm (where at each time instant the agents pick randomly one of its neighbors and exchange information) can be expressed as

$$X_i(k+1) = \sum_{j\in\mathcal{N}_i(k)} \mathbb{1}_{\{\theta_i(k)=j\}} \left[\lambda_{ij} X_i(k) + (1-\lambda_{ij}) X_j(k)\right].$$

In this case, $X_i(k)$ take real values, $\theta_i(k)$ is an independent random variable taking values in the set $\mathcal{N}_i(k)$, with probability mass function $Pr(\theta_i(k) = j) = w_{ij}(k)$, where w_{ij} is the probability of agent i to choose agent j at time k. In the case $S = \{0, 1\}$, algorithm (18) is more in the spirit of the binary influence model algorithm, studied in [1].

C. Consensus on the Collection of Convex, Compact Sets

In this subsection we present a consensus algorithm based on Minkowski sums between convex, compact sets in $\mathbb{R}^n$. For this example we draw inspiration from the analysis of linear dynamics driven by compact, convex sets studied in [22] and [23]. Let $\mathcal{X} = \mathrm{ComConv}(\mathbb{R}^n)$ denote the collection of convex, compact sets in $\mathbb{R}^n$. Given two sets $X \in \mathbb{R}^n$ and $Y \in \mathbb{R}^n$, the Minkowski sum between the two sets is given by $X \oplus Y \triangleq \{x + y : x \in X, y \in Y\}$. The multiplication of a set X by a scalar is defined as $\lambda X \triangleq \{\lambda x : x \in X\}$. It can be easily checked that for any $X, Y \in \mathcal{X}$, we have that $\lambda X \oplus (1-\lambda)Y \in \mathcal{X}$ and $\lambda X \oplus (1-\lambda)X = X$ for all $\lambda \in [0, 1]$. It is well-known that $\mathcal{X} = \mathrm{ComConv}(\mathbb{R}^n)$ endowed with the Hausdorff distance is a complete metric space [27], where the Hausdorff distance is defined as

$$H(L, X, Y) = \min_{\alpha \geq 0}\{\alpha \mid X \subseteq Y \oplus \alpha L, Y \subseteq X \oplus \alpha L\} \quad (21)$$

with $L \in \mathcal{X}$ a symmetric, non-empty set containing the origin. Let us now define the mapping $\psi(X, Y, \lambda) = \lambda X \oplus (1-\lambda)Y$, where $X, Y \in \mathcal{X}$ and $\lambda \in [0, 1]$. Using the above observations it should be clear that any set produced by the mapping ψ belongs to $\mathcal{X}$. The following proposition shows that $\psi(X, Y, \lambda)$ is indeed a convex structure:

Proposition 5.3: The mapping ψ is a convex structure on $\mathcal{X}$, with respect to the Hausdorff distance.

Proof: All we have to show is that the following inequality holds:

$$H(L, U, \psi(X, Y, \lambda)) \leq \lambda H(L, U, X) + (1-\lambda)H(L, U, X) \quad (22)$$

for all $U, X, Y \in \mathcal{X}$ and $\lambda \in [0, 1]$. To simplify the proof, we use the fact that the Hausdorff distance can also be represented in terms of the support function of a closed, convex set. Given that the support function at a point $z \in \mathbb{R}^n$ is given by

$$s(X, z) = \sup_x \{z'x \mid x \in X\}$$

the Hausdorff distance between two closed and convex sets X and Y can be equivalently expressed as

$$H(L, X, Y) = \|s(X, \cdot) - s(Y, \cdot)\|_\infty \quad (23)$$

where $\|\cdot\|$ is the uniform norm on the unit sphere, that is, $\|f\|_\infty = \sup_z\{f(z) | z'z \leq 1\}$. Therefore, we have that

$$H(L, U, \psi(X, Y, \lambda)) = \|s(U, \cdot) - s(\psi(X, Y, \lambda), \cdot)\|_\infty. \quad (24)$$

Observing that the support function of a set $\lambda X \oplus (1-\lambda)Y$ can be expressed as

$$\begin{aligned} &s(\lambda X \oplus (1-\lambda)Y, z) \\ &\quad = \sup_{x,y}\{\lambda z'x + (1-\lambda)z'y \mid y, x \in X, y \in Y\} \\ &\quad = \lambda s(X, z) + (1-\lambda)s(Y, z) \end{aligned}$$

we have that (24) can be further expressed as

$$\begin{aligned} &H(L, U, \psi(X, Y, \lambda)) \\ &\quad = \|s(\lambda U \oplus (1-\lambda)U, \cdot) - s(\lambda X \oplus (1-\lambda)Y, \cdot)\|_\infty \\ &\quad = \|\lambda[s(U, \cdot) - s(X, \cdot)] + (1-\lambda)[s(U, \cdot) - s(Y, \cdot)]\|_\infty \\ &\quad \leq \lambda\|s(U, \cdot) - s(X, \cdot)\|_\infty + (1-\lambda)\|s(U, \cdot) - s(Y, \cdot)\|_\infty \\ &\quad = \lambda H(L, U, X) + (1-\lambda)H(L, U, Y) \end{aligned}$$

where the last equality followed from (23), and the result follows. ■

The next result, whose proof can be found in the Appendix section, characterizes the convex hull of a finite collection of compact, convex sets in terms of the Minkowski sum operator.

Proposition 5.4: Let n be a positive integer and let $A = \{X_1, \ldots, X_n\}$ be a finite collection of sets in $\mathrm{ComConv}(\mathbb{R}^n)$. Then the convex hull of the set A is given by

$$co(A) = \left\{Z \in \mathcal{X} \mid Z = \bigoplus_{i=1}^n w_i X_i,\ \forall w_i \geq 0,\ \sum_{i=1}^n w_i = 1\right\}.$$

A description of an arbitrarily accurate approximation of $co(A)$ that it is guaranteed to belong to its interior is given in the next corollary.

Corollary 5.2: Let n be a positive integer and let $A = \{X_1, \ldots, X_n\}$ be a set of points in $\mathrm{ComConv}(\mathbb{R}^n)$. Then

$$co_\varepsilon(A) = \left\{Z \in \mathcal{X} \mid Z = \bigoplus_{i=1}^n w_i X_i,\ \forall w_i \geq \varepsilon,\ \sum_{i=1}^n w_i = 1\right\}.$$

Proof: Follows immediately from Definition 2.7 and Proposition 5.4. ■

Using the above corollary, the agreement algorithm on the collection of compact, convex sets is given by

$$X_i(k+1) = \bigoplus_{j \in \mathcal{N}_i(k)} w_{ij}(k) X_j(k) \quad (25)$$

where $w_{ij}(k)$ are positive scalars summing up to one, and lower bounded by ε.

D. Numerical Examples

In this subsection, we present instances of executions of three of the algorithms described in the previous section. The graph that constrains the communication between agents has sixty nodes and its structure is shown in Fig. 2.

Each of the considered algorithms are implemented using the approach described in Remark 2.4, that is, by repeatedly

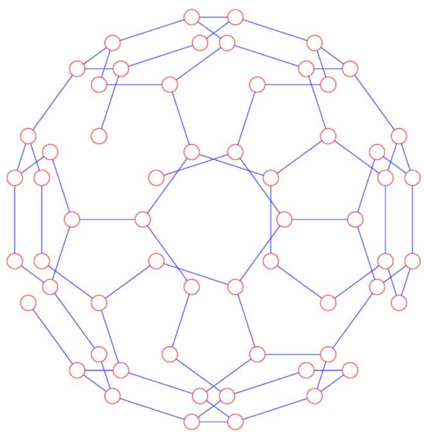

Fig. 2. Network of sixty agents.

applying the convex structure operator on the neighbors of the agents. We show numerical simulations for two values of the parameter λ used in the convex structure, namely $\lambda = 0.3$ and $\lambda = 0.7$. For each algorithm we show the evolution of the states and of the distances between agents, for the two values of λ.

The Fig. 3(a)–(c) present numerical simulations of the agreement algorithm in the case of the set of vectors with positive entries. The initial values of the agents are uniformly distributed in the interval [0, 9], and are the same for the two values of λ.

Simulation results of the agreement algorithm applied on a discrete set of numbers are depicted in Fig. 4(a)–(c), in which the initial conditions are uniformly chosen from the set $\{0, 1, \ldots, 9\}$. Note that since the distance on this space is defined as an expectation, the convergence speed of the distances between agents does not necessarily reflect the convergence speed of particular realizations of the algorithm as shown in Fig. 4(a) and (b).

Finally, Fig. 5(a)–(c) show the behavior of the generalized randomized gossip algorithm applied on the collection of compact, convex sets.

The initial values of the states are polytope approximations of circles with radiuses uniformly chosen from the interval [0.8, 4.8] and number of edges uniformly picked from the set $\{3, \ldots, 7\}$. The figures shows how the shape of the sets corresponding to the agents change as the agents interact. For executing the numerical simulation of the agreement algorithm on compact, convex sets, we used the Multi-Parametric toolbox [8] that provides efficient numerical algorithms for computing Minkowski sums of convex sets. In all three examples the agents converge to the same value and the distances between the states of the agents converge to zero, as well. Note also the for the smaller value of λ the rate of convergence was higher compared to the larger value, suggestion that a selfish attitude from the part of the agents does not favor rapid convergence.

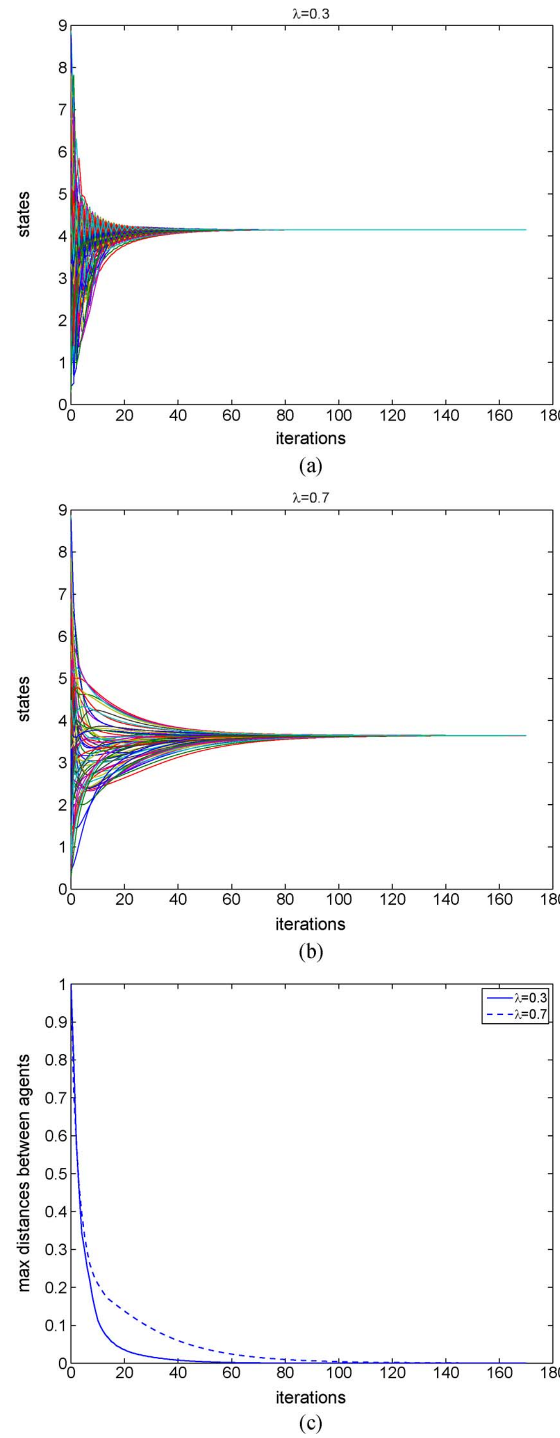

Fig. 3. Agreement Algorithm on the set of Positive Real Vectors: (a) states evolution for $\lambda = 0.3$; (b) states evolution for $\lambda = 0.7$; (c) (upper bounds on the) maximum of the distances between the states of the agents.

VI. Conclusion

In this paper, 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 spaces. 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. In addition, we gave several examples of convex metric spaces and their corresponding agreement algorithms. They show that apparently very different algorithms belong to the same class of algorithms; algorithms defined on convex metric spaces.

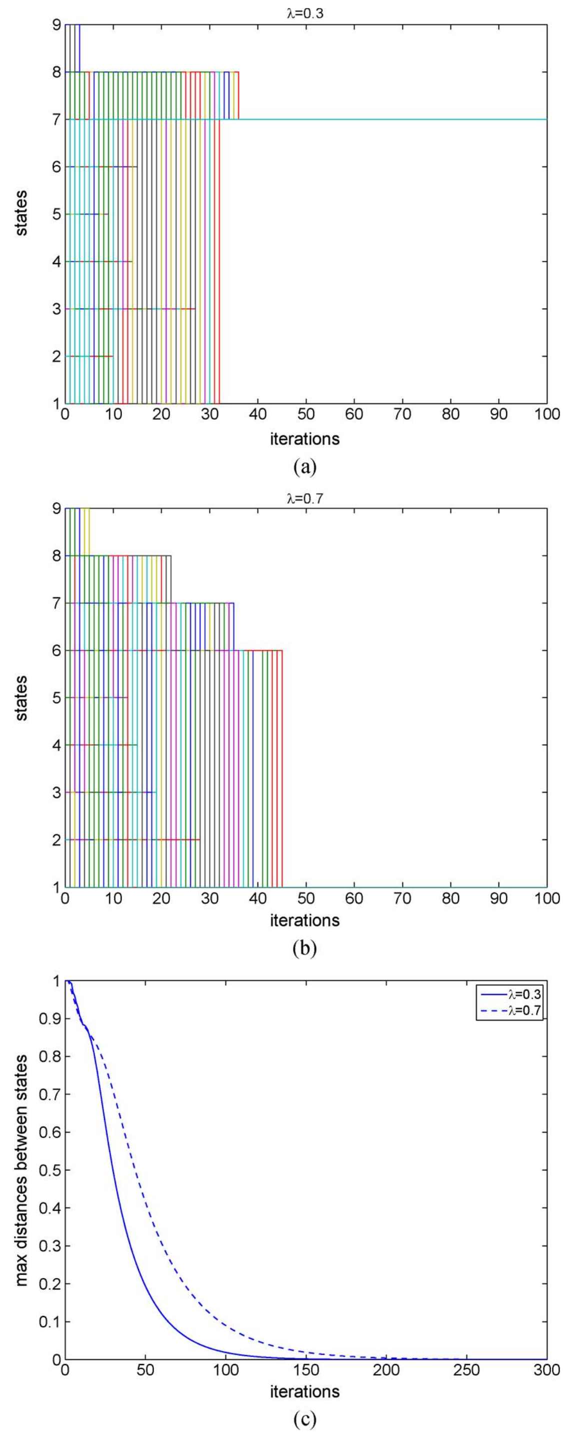

Fig. 4. Agreement algorithm on Discrete Finite Sets: (a) states evolution for $\lambda = 0.3$; (b) states evolution for $\lambda = 0.7$; (c) (upper bounds on the) maximum of the distances between the states of the agents.

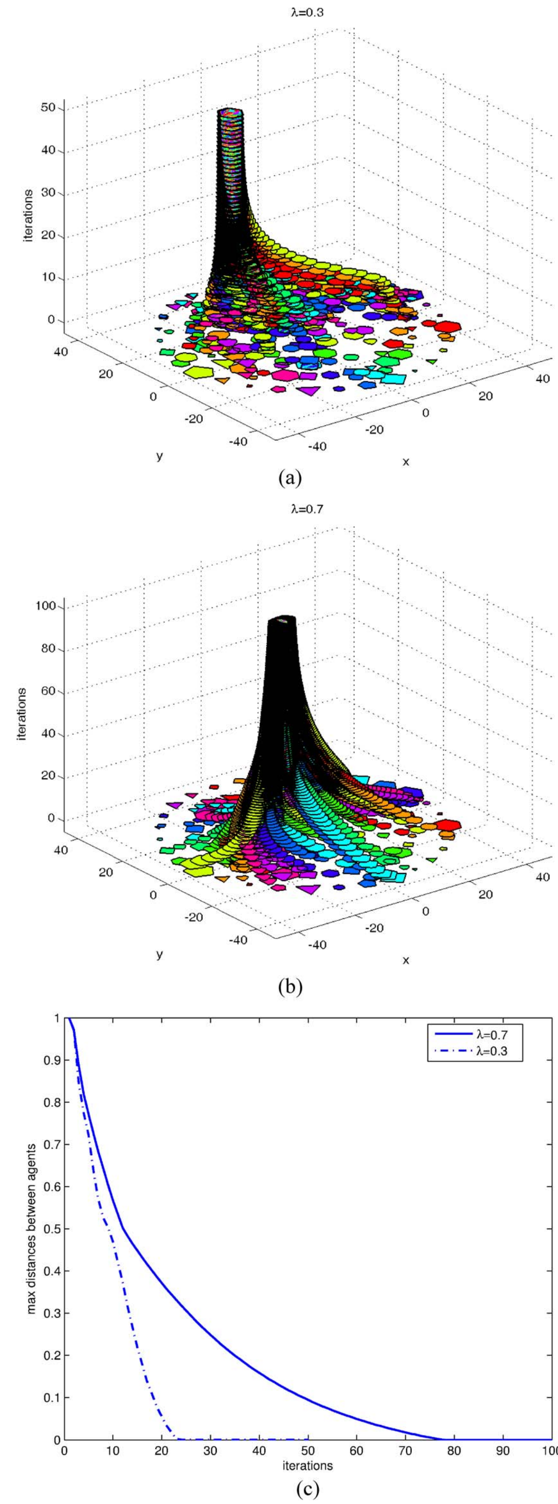

Fig. 5. Agreement Algorithm on Compact, Convex Sets: (a) states evolution for $\lambda = 0.3$; (b) states evolution for $\lambda = 0.7$; (c) (upper bounds on the) maximum of the distances between the states of the agents.

APPENDIX A
PROOF OF LEMMA 4.2

Let $(i^*, j^*) \in E_\infty$ be a pair of agents. By Assumptions 3.1 and 3.2, there exists a positive integer $s' \in \{s, s+1, \ldots, s+\bar{B}-1\}$ such that agent j^* sends information to agent i^* at time s'. 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 elements of row $\bar{i}^*$ of the 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

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

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

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

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

$$\sum_{\bar{j}=1}^N [\Phi(s'+m,s)]_{\bar{i}^*\bar{j}} \leq 1-\eta^m, \ \forall m \geq 0. \tag{26}$$

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 (26), we obtain that for $s' = s+B-1$

$$\sum_{\bar{j}=1}^N [\Phi(s+B-1+m,s)]_{\bar{i}^*\bar{j}} \leq 1-\eta^m, \ \forall m \geq 0 \tag{27}$$

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

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

From (27) and (28), 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

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

and the result follows.

Appendix B
Proof of Theorem 5.1

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

$$co(A) = \lim_{m\to\infty} A_m = \bigcup_{m=1}^{\infty} A_m$$

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 $m \geq 1$. We define the set

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

The proof consists of 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 $co(A)$ 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 $\{\theta_1, \ldots, \theta_{n-1}\}$ be a set of independent random variables taking values in $\{1, 2\}$, defined on the probability space $(\Omega, \mathcal{F}, \mathcal{P})$, with probability mass functions given by

$$\begin{aligned}
&Pr(\theta_1 = 1) = \frac{w_1}{w_1+w_2}, \ Pr(\theta_1 = 2) = \frac{w_2}{w_1+w_2} \\
&Pr(\theta_2 = 1) = \frac{w_1+w_2}{w_1+w_2+w_3}, \ Pr(\theta_1 = 2) = \frac{w_3}{w_1+w_2+w_3} \\
&\qquad \vdots \\
&Pr(\theta_{n-1} = 1) = \frac{w_1+\ldots+w_{n-1}}{w_1+\ldots+w_n}, Pr(\theta_{n-1} = 2) = \frac{w_n}{w_1+\ldots+w_n}.
\end{aligned} \tag{29}$$

We also make the assumption that

$$\begin{aligned}
\omega_n &= \{\omega : \theta_{n-1}(\omega) = 2\} \\
\omega_{n-1} &= \{\omega : \theta_{n-1}(\omega) = 1, \theta_{n-2}(\omega) = 2\} \\
\omega_{n-2} &= \{\omega : \theta_{n-1}(\omega) = 1, \theta_{n-2}(\omega) = 1, \theta_{n-3}(\omega) = 2\} \\
&\ \vdots \\
\omega_2 &= \{\omega : \theta_{n-1}(\omega) = 1, \ldots, \theta_2(\omega) = 1, \theta_1(\omega) = 2\} \\
\omega_1 &= \{\omega : \theta_{n-1}(\omega) = 1, \ldots, \theta_2(\omega) = 1, \theta_1(\omega) = 1\}.
\end{aligned} \tag{30}$$

The above assumptions make sense, since for each row of (30), the measures of the right and left hand sides are equal. This follows from the independence of θ_is and from the definitions of their probability mass functions. Consider now the following iterative expression

$$Y_i = \mathbb{1}_{\{\theta_i=1\}} Y_{i-1} + \mathbb{1}_{\{\theta_i=2\}} X_{i+1}$$

for $i = 1, \ldots, n-1$, with $Y_0 = X_1$. By using the independence assumption on θ_is together with (29) and (30), it can be easily

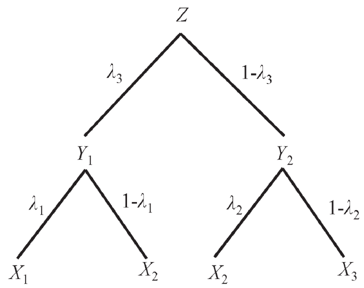

Fig. 6. Example of a tree representation of a point $Z \in A_2$ with $A = \{X_1, X_2, X_3\} \in \mathcal{X}$.

checked that $Y_{n-1} = Z$. But since $Y_i \in A_i$, $i = 1, \ldots, n-1$ it follows that $Z \in A_{n-1}$ or $Z \in co(A)$ which implies that $\mathcal{K}(A) \subset co(A)$.

We now begin the second part of the proof and show that any point in $co(A)$ belongs to $\mathcal{K}(A)$ as well. If $Z \in co(A)$, from Section II-B we have that there exists 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. To simplify the notations and make the argument more clear, we will assume particular values for m and n and a particular tree decomposition. However, as the reader will notice, the principle of the proof used in this particular example can be easily generalized for any value of m and n and any tree representation.

In the following we consider a point $Z \in A_2$, with a tree representation given in Fig. 6. We have that there exist the independent random variables $\theta_i \in \{1, 2\}$, for $i = 1$, 2, 3, defined on the probability space $(\Omega, \mathcal{F}, \mathcal{P})$ and with probability distributions $Pr(\theta_1 = 1) = \lambda_1$, $Pr(\theta_1 = 2) = 1 - \lambda_1$, $Pr(\theta_2 = 1) = \lambda_2$, $Pr(\theta_2 = 2) = 1 - \lambda_2$ and $Pr(\theta_3 = 1) = \lambda_3$, $Pr(\theta_3 = 2) = 1 - \lambda_3$, such that

$$Y_1 = \mathbb{1}_{\{\theta_1=1\}} X_1 + \mathbb{1}_{\{\theta_1=2\}} X_2$$
$$Y_2 = \mathbb{1}_{\{\theta_2=1\}} X_2 + \mathbb{1}_{\{\theta_2=2\}} X_3$$
$$Z = \mathbb{1}_{\{\theta_3=1\}} Y_1 + \mathbb{1}_{\{\theta_3=2\}} Y_2.$$

From above, we have that Z can be alternatively written as

$$Z = \mathbb{1}_{\{\theta_3=1\}} \mathbb{1}_{\{\theta_1=1\}} X_1 + \mathbb{1}_{\{\theta_3=1\}} \mathbb{1}_{\{\theta_1=2\}} X_2 + \mathbb{1}_{\{\theta_3=2\}} \mathbb{1}_{\{\theta_2=1\}} X_2 + \mathbb{1}_{\{\theta_3=2\}} \mathbb{1}_{\{\theta_2=2\}} X_3.$$

We now define the following sets:

$$\omega_1 = \{\omega : \theta_3(\omega) = 1, \theta_1(\omega) = 1\}$$
$$\omega_2 = \{\omega : \theta_3(\omega) = 1, \theta_1(\omega) = 2\} \cup \{\omega : \theta_3(\omega) = 2, \theta_2(\omega) = 2\}$$
$$\omega_3 = \{\omega : \theta_3(\omega) = 2, \theta_2(\omega) = 2\}$$

and define the random variable θ on the probability space $(\Omega, \mathcal{F}, \mathcal{P})$, taking values in {1, 2, 3} such that $\theta(\omega_1) = 1$, $\theta(\omega_2) = 2$, and $\theta(\omega_3) = 3$. The probability mass function of θ is given by

$$Pr(\theta = 1) = \lambda_1 \lambda_3 = \mathcal{W}^Z_{X_1}$$
$$Pr(\theta = 2) = (1 - \lambda_1)\lambda_3 + \lambda_2(1 - \lambda_3) = \mathcal{W}^Z_{X_2}$$
$$Pr(\theta = 3) = (1 - \lambda_2)(1 - \lambda_3) = \mathcal{W}^Z_{X_3}.$$

Therefore, from the independence of θ_1 θ_2 and θ_3, Z can be expressed as

$$Z = \mathbb{1}_{\{\theta=1\}} X_1 + \mathbb{1}_{\{\theta=2\}} X_2 + \mathbb{1}_{\{\theta=3\}} X_3$$

and hence $Z \in \mathcal{K}(A)$ with $w_1 = \mathcal{W}^Z_{X_1}$, $w_2 = \mathcal{W}^Z_{X_2}$ and $w_3 = \mathcal{W}^Z_{X_3}$. Noticing that the approach presented above can be easily generalized for any tree representation of a point Z, we have that $co(A) \subset \mathcal{K}(A)$, and the result follows.

APPENDIX C
PROOF OF PROPOSITION 5.4

The proof follows that same lines as the proof of the previous theorem. We define the set

$$\mathcal{K}(A) \triangleq \left\{ Z \in \mathrm{ComConv}(\mathbb{R}^n) \mid Z = \bigoplus_{i=1}^{n} x_i X_i, \; \forall w_i \geq 0, \; \sum_{i=1}^{n} w_i = 1 \right\}.$$

We can easily note that any element $Z \in \mathcal{K}(A)$ can be represented using the iteration

$$Y_i = \lambda_i Y_{i-1} \oplus (1 - \lambda_i) X_{i+1}$$

for $i = 1, \ldots, n-1$, with $Y_0 = X_1$. Recalling the definition of the convex hull from Proposition 2.1, it follows that $Y_i \in A_i$, $i = 1, \ldots, n-1$ and therefore $Z \in co(A)$ since $Z = Y_{n-1} \in A_{n-1}$.

We now show that an element $Z \in co(A)$ belongs to $\mathcal{K}(A)$ as well. If $Z \in co(A)$, from Section II-B we have that there exists 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, and weights in the interval [0, 1]. Recall that each node X in the tree can be expressed as $X = \lambda Y \oplus (1 - \lambda) Z$, where Y and Z are two nodes in the tree at one level lower than X and $\lambda \in [0, 1]$. Therefore, by recursively representing the top node Z in terms of lower level nodes, we find that Z is given by $Z = \bigoplus_{i=1}^{N} w_i X_i$, where w_i are non-negative scalars summing-up to one, collecting the contribution of all the weights of the tree. Consequently, $Z \in \mathcal{K}(A)$ and the result follows.

REFERENCES

[1] C. Asavathiratham, "The influence model: A tractable representation for the dynamics of networked Markov chains," Ph.D. dissertation, MIT, Cambridge, MA, Apr. 2001.

[2] V. D. Blondel, J. M. Hendrickx, A. Olshevsky, and J. N. Tsitsiklis, "Convergence in multiagent coordination, consensus, flocking," in *Proc. 44th IEEE Conf. Decision and Control*, Dec. 2005, pp. 2996–3000.

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

[4] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah, "Randomized gossip algorithms," *IEEE/ACM Trans. Netw.*, vol. 14, no. SI, pp. 2508–2530, 2006.

[5] G. Dunford and J. T. Schwartz, *Linear Operators Part 1. Interscience*. New York, 1958.

[6] G. R. Grimmett and D. R. Stirzaker, *Probability and Random Processes*, 2nd ed. New York: Oxford University Press, 1992.

[7] Y. Hatano and M. Mesbahi, "Agreement over random networks," *IEEE Trans. Autom. Control*, vol. 50, no. 11, pp. 1867–1872, Nov. 2005.
[8] M. Herceg, M. Kvasnica, C. N. Jones, and M. Morari, "Multi-Parametric Toolbox 3.0," in *Proc. European Control Conf.*, Zürich, Switzerland, Jul. 17–19, 2013, pp. 502–510.
[9] 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. 2004.
[10] I. Matei and J. S. Baras, "Geometry of a probabilistic consensus of opinion algorithm," in *Proc. American Control Conf.*, Jun. 28–30, 2011, pp. 2198–2203.
[11] I. Matei, N. Martins, and J. Baras, "Almost sure convergence to consensus in markovian random graphs," in *Proc. 47th IEEE Conf. Decision and Control*, Dec. 2008, pp. 3535–3540.
[12] I. Matei, N. Martins, and J. Baras, "Consensus problems with directed markovian communication patterns," in *Proc. 2009 IEEE American Control Conf.*, Jun. 2009, pp. 1298–1303.
[13] I. Matei, C. Somarakis, and J. S. Baras, "A randomized gossip consensus algorithm on convex metric spaces," in *Proc. Conf. Decision and Control*, Dec. 10–13, 2012, pp. 7425–7430.
[14] I. Matei, C. Somarakis, and J. S. Baras, "A generalized gossip algorithm on convex metric spaces," *IEEE Trans. Autom. Control*, to be published.
[15] L. Moreau, "Stability of multi-agents systems with time-dependent communication links," *IEEE Trans. Autom. Control*, vol. 50, no. 2, pp. 169–182, Feb. 2005.
[16] A. Nedić and A. Ozdaglar, "Distributed subgradient methods for multi-agent optimization," *IEEE Trans. Autom. Control*, vol. 54, no. 1, pp. 48–61, Jan. 2009.
[17] A. Nedić and A. Ozdaglar, "Convergence rate for consensus with delays," *J. Global Optimiz.*, vol. 47, no. 3, pp. 437–456, Jul. 2010.
[18] A. Nedić, A. Ozdaglar, and P. A. Parrilo, "Constrained consensus and optimization in multi-agent networks," *IEEE Trans. Autom. Control*, vol. 55, no. 4, pp. 922–938, Apr. 2010.
[19] R. Olfati-Saber and R. M. Murray, "Consensus protocols for networks of dynamic agents," in *Proc. 2003 IEEE American Control Conf.*, 2003, pp. 951–956.
[20] R. Olfati-Saber and R. M. Murray, "Consensus problem in networks of agents with switching topology and time-delays," *IEEE Trans. Autom. Control*, vol. 49, no. 9, pp. 1520–1533, Sep. 2004.
[21] M. Porfiri and D. J. Stilwell, "Consensus seeking over random directed weighted graphs," *IEEE Trans. Autom. Control*, vol. 52, no. 9, pp. 1767–1773, Sep. 2007.
[22] S. V. Raković, "Minkowski algebra and banach contraction principle in set invariance for linear discrete time systems," in *Proc. 46th IEEE Conf. Decision and Control*, 2007, pp. 2169–2174.
[23] S. V. Raković, I. Matei, and J. S. Baras, "Reachability analysis for linear discrete time set-dynamics driven by random convex compact sets," in *Proc. IEEE 51st Annu. Conf. Decision and Control (CDC)*, 2012, pp. 4751–4756.
[24] 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.
[25] A. Tahbaz Salehi and A. Jadbabaie, "Necessary and sufficient conditions for consensus over random networks," *IEEE Trans. Autom. Control*, vol. 53, no. 3, pp. 791–795, Mar. 2008.
[26] A. Tahbaz Salehi and A. Jadbabaie, "Consensus over ergodic stationary graph processes," *IEEE Trans. Autom. Control*, vol. 55, no. 1, pp. 225–230, Jan. 2010.
[27] R. Scheneider, *Convex Bodies: The Brunn-Minkowski Theory*. Cambridge, U.K.: Cambridge University Press, 1993.
[28] D. Shah, "Gossip algorithms," *Found. Trends Netw.*, vol. 3, no. 1, pp. 1–125, 2008.
[29] B. K. Sharma and C. L. Dewangan, "Fixed point theorem in convex metric space," *Novi Sad J. Math.*, vol. 25, no. 1, pp. 9–18, 1995.
[30] W. Takahashi, "A convexity in metric mpace and non-expansive mappings I," *Kodai Math. Sem. Rep.*, vol. 22, pp. 142–149, 1970.
[31] J. N. Tsitsiklis, "Problems in decentralized decision making and computation," Ph.D. dissertation, MIT, Cambridge, MA, Nov. 1984.
[32] J. N. Tsitsiklis, D. P. Bertsekas, and M. Athans, "Distributed asynchronous deterministic and stochastic gradient optimization algorithms," *IEEE Trans. Autom. Control*, vol. 31, no. 9, pp. 803–812, Sep. 1986.
[33] S. E. Tuna and R. Sepulchre, "Consensus under general convexity," in *Proc. 46th IEEE Conf. Decision and Control*, 2007, pp. 294–299.

Ion Matei received the B.S. degree in electrical engineering (with highest distinction) and the M.S. degree in electrical engineering from the Politehnica University of Bucharest, Bucharest, Romania, in 2002 and 2003, respectively. He also received the M.S. and Ph.D. degrees in electrical engineering from the University of Maryland, College Park, in 2009 and 2010 respectively.

He is currently a Research Scientist in the System Sciences Laboratory, Palo Alto Research Center (PARC), CA. Prior to joining PARC, he was a Research Scientist with the Institute for Systems Research, University of Maryland, College Park, and a Guest Researcher at the National Institute of Standards and Technology (NIST), Gaithersburg, MD. As part of the Systems Engineering Group at NIST, he contributed to the development of standards for modeling cyber-physical systems, with emphasis on representation of hybrid dynamical systems. His research interests revolve around modeling and analysis of complex dynamical systems. His early research focused on designing and analyzing distributed algorithms for optimization, estimation, and control for multi-agent systems. In his most recent work, he focuses on reliability and diagnosis studies for cyber-physical systems. He has published over 40 peer-reviewed articles and papers in leading journals and conferences in control, optimization, and signal processing.

John S. Baras (LF'13) received the Diploma in electrical and mechanical engineering (with highest distinction) from the National Technical University of Athens, Athens, Greece, in 1970, and the M.S. and Ph.D. degrees in applied mathematics from Harvard University, Cambridge, MA, in 1971 and 1973, respectively.

Since 1973, he has been with the Department of Electrical and Computer Engineering, University of Maryland at College Park, where he is currently a Professor, member of the Applied Mathematics, Statistics and Scientific Computation Program Faculty, and Affiliate Professor in the Fischell Department of Bioengineering and the Department of Mechanical Engineering. From 1985 to 1991, he was the Founding Director of the Institute for Systems Research (ISR) (one of the first six NSF Engineering Research Centers). In February 1990, he was appointed to the Lockheed Martin Chair in Systems Engineering. Since 1991, he has been the Director of the Maryland Center for Hybrid Networks (HYNET), which he co-founded. He has held visiting research scholar positions with Stanford, MIT, Harvard, the Institute National de Reserche en Informatique et en Automatique (INRIA), the University of California at Berkeley, Linkoping University and the Royal Institute of Technology (KTH) in Sweden. He was the initial architect and continuing innovator of the pioneering M.S. on Systems Engineering program of the ISR. His research interests include control, communication, and computing systems. He holds eight patents and has four more pending.

Dr. Baras has received numerous awards, including the 1980 George S. Axelby Award of the IEEE Control Systems Society; the 1978, 1983, and 1993 Alan Berman Research Publication Awards from the NRL; the 1991, 1994, and 2008 Outstanding Invention of the Year Awards from the University of Maryland; the 1998 Mancur Olson Research Achievement Award from the University of Maryland College Park; the 2002 and 2008 Best Paper Awards at the 23rd and 26th Army Science Conferences; the 2004 Best Paper Award at the Wireless Security Conference WISE04; the 2007 IEEE Communications Society Leonard G. Abraham Prize in the Field of Communication Systems; the 2008 IEEE Globecom Best Paper Award for wireless networks; and the 2009 Maryland Innovator of the Year Award. In November 2012, he was honored by the Awards for both the "Principal Investigator with Greatest Impact" and for the "Largest Selling Product" with Hughes Network Systems for HughesNet, over the last 25 years of operation of the Maryland Industrial Partnerships Program. These awards recognized Dr. Baras' pioneering invention, prototyping, demonstration, and help with commercialization of Internet protocols and services over satellites in 1994, which created a new industry serving tens of millions worldwide. In 2014, he was awarded the 2014 Tage Erlander Guest Professorship by the Swedish Research Council, and a three year (2014–2017) Hans Fischer Senior Fellowship by the Institute for Advanced Study of the Technical University of Munich. He is a Fellow of SIAM and a Foreign Member of the Royal Swedish Academy of Engineering Sciences (IVA).