51st IEEE Conference on Decision and Control
December 10-13, 2012. Maui, Hawaii, USA

A Randomized Gossip Consensus Algorithm on Convex Metric Spaces

Ion Matei, Christoforos Somarakis, John S. Baras

***Abstract*— A consensus problem consists of a group of dynamic agents who seek to agree upon certain quantities of interest. This problem can be generalized in the context of convex metric spaces that extend the standard notion of convexity. In this paper we introduce and analyze a randomized gossip algorithm for solving the generalized consensus problem on convex metric spaces. We study the convergence properties of the algorithm using stochastic differential equations theory. We show that the dynamics of the distances between the states of the agents can be upper bounded by the dynamics of a stochastic differential equation driven by Poisson counters. In addition, we introduce instances of the generalized consensus algorithm for several examples of convex metric spaces.**

I. Introduction

A particular distributed algorithm is the *consensus* (or agreement) algorithm, where a group of dynamic agents 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. Different aspects of the consensus problem were addressed by Olfati-Saber and Murray [14], Jadbabaie et al. [6], Ren and Beard [17], Moreau [12] or, more recently, by Nedic and Ozdaglar [13].

The random behavior of communication networks motivated the investigation of consensus algorithms under a stochastic framework [5], [9], [16], [18], [19]. In addition to network variability, nodes in sensor networks operate under limited computational, communication, and energy resources. These constraints have motivated the design of gossip algorithms, in which a node communicates with a randomly chosen neighbor. Studies of randomized gossip consensus algorithms can be found in [2], [20]. In particular, consensus based gossip algorithms have been extensively used in the analysis and study of the performance of wireless networks, with random failures [15].

Ion Matei is with the Institute for Research in Electronics and Applied Physics at the University of Maryland, College Park (email: imatei@umd.edu); Christoforos Somarakis and John S. Baras are with the Institute for Systems Research at the University of Maryland, College Park (emails: csomarak@umd.edu, baras@umd.edu).

This material is based in part upon work supported by the NIST ARRA Measurement Science and Engineering Fellowship Program award 70NANB10H026, through the University of Maryland, and in part upon work supported by the U.S. Army Research Office MURI award grant W911NF-08-1-0238, by the National Science Foundation (NSF) award grant CNS-1018346 and by the U.S. AFOSR MURI award grant FA9550-09-1-0538.

In this paper, we introduce and analyse a generalized randomized gossip algorithm for achieving consensus. The algorithm acts on *convex metric spaces*. These are abstract metric spaces endowed with a *convex structure*. We show that under the given algorithm, the agents' states converge to consensus with probability one and in the r^{th} mean sense. Additionally, for a particular network topology we investigate in more depth the rate of convergence of the first and second moments of the distances between the agents' states and we present instances of the generalized gossip algorithm for three convex metric spaces. The results of this paper complement our previous results in [7], [8]. This paper is a brief version of a more comprehensive technical report [10].

The paper is organized as follows. Section II introduces the main concepts related to convex metric spaces. Section III formulates the problem and states our main results. Sections IV and V give the proof of our main results, together with pertinent preliminary results. In Section VI we present an in-depth analysis of the rate of convergence to consensus (in the first and second moments), for a particular network topology. Section VII shows instances of the generalized consensus algorithm for three convex metric spaces, defined on the sets of real numbers, compact intervals and discrete random variables, respectively.

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 without self loops, for which every edge corresponds to a *non-zero*, *off-diagonal* entry of W. We denote by $\chi_{\{A\}}$ the indicator function of the event A.

II. Convex Metric Spaces

In this section we introduce a set of definitions and basic results about convex metric spaces. Additional information about the following definitions and results can be found in [21],[22].

Definition 2.1 ([22], 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

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

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

Definition 2.2 ([22], 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 $(\mathcal{X}, d, \psi)$.
In Section VII we present three examples of convex metric spaces to which the generalized gossip algorithm is applied.

Definition 2.3 ([22], 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.
The class of such convex metric spaces is rather large since by Smulian's Theorem ([3], page 443), every weakly compact convex subset of a Banach space has *Property* (C).

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

Definition 2.4 ([22], pp. 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, pp. 143 of [22] 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 ([21], pp. 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 $co(A)$.

An alternative 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 [21] that the set sequence $\{A_m\}_{m\geq 0}$ is increasing 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 ([21], 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 $co(A) = A_m$.

III. Problem formulation and main results

Let $(\mathcal{X}, d, \psi)$ be a convex metric space. We consider a set of n agents indexed by i, with states denoted by $x_i(t)$ taking values on $\mathcal{X}$, where t represents the continuous time.

A. Communication model

The communication among agents is modeled by a directed graph $G = (V, E)$, where $V = \{1, 2, \ldots, n\}$ is the set of agents, and $E = \{(j, i) \mid j \text{ can send information to } i\}$ is the set of edges. In addition, we denote by $\mathcal{N}_i$ the inward neighborhood of agent i, i.e.,

$$\mathcal{N}_i \triangleq \{j \mid (j, i) \in E\},$$

where by assumption node i does not belong to the set $\mathcal{N}_i$. We make the following connectivity assumption.

Assumption 3.1: The graph $G = (V, E)$ is strongly connected.

B. Randomized gossip algorithm

We assume that the agents can be in two modes: *sleep* mode and *update* mode. Let $N_i(t)$ be a Poisson counter associated to agent i. In the sleep mode, the agents maintain their states unchanged. An agent i exits the sleep mode and enters the update mode, when the associated counter $N_i(t)$ increments its value. Let t_i be a time-instant at which the Poisson counter $N_i(t)$ increments its value. Then at t_i, agent i picks agent j with probability $p_{i,j}$, where $j \in \mathcal{N}_i$ and updates its state according to the rule

$$x_i(t_i^+) = \psi(x_i(t_i), x_j(t_i), \lambda_i), \tag{1}$$

where $\lambda_i \in [0,\ 1)$, ψ is the convex structure and $\sum_{j\in\mathcal{N}_i} p_{i,j} = 1$. By $x_i(t_i^+)$ we understand the value of $x_i(t)$ immediately after the instant update at time t_i, which can be also written as

$$x_i(t_i^+) = \lim_{t\to t_i,\ t>t_i} x_i(t),$$

which implies that $x_i(t)$ is a right-continuous function of t. After agent i updates its state according to the above rule, it immediately returns to the sleep mode, until the next increase in value of the counter $N_i(t)$.

Assumption 3.2: The Poisson counters $N_i(t)$ are independent and with rate μ_i, for all i.

A similar form of the above algorithm (the Poisson counters are assumed to have the same rates) was extensively studied in [2], in the case where $\mathcal{X} = \mathbb{R}$.

We first note that since the agents update their state at random times, the distances between agents are random processes. Let $d(x_i(t), x_j(t))$ be the distance between the states of agents i and j, at time t. The following theorems state our main convergence results.

Theorem 3.1: Under Assumptions 3.1 and 3.2 and under the randomized gossip algorithm, the agents converge to consensus in r^{th} mean, that is

$$\lim_{t\to\infty} E\left[d(x_i(t), x_j(t))^r\right] = 0, \forall\ (i, j),\ i \neq j.$$

Theorem 3.2: Under Assumptions 3.1 and 3.2 and under the randomized gossip algorithm, the agents converge to consensus with probability one, that is

$$Pr\left(\lim_{t\to\infty} \max_{i,j} d(x_i(t), x_j(t)) = 0\right) = 1.$$

The above results show that the distances between the agents' states converge to zero. The following Corollary shows that in fact, for convex metric spaces satisfying *Property* (C), the states of the agents converge to some point in the convex metric space.

Corollary 3.1: Under Assumptions 3.1 and 3.2 and under the randomized gossip algorithm operating on convex metric spaces satisfying *Property* (C), for any sample path ω of state processes, there exits $x^* \in \mathcal{X}$ (that depends on ω and the initial conditions $x_i(0)$) such that

$$\lim_{t\to\infty} d(x_i(t, \omega), x^*(\omega)) = 0, \text{ for all } i.$$

In other words, the states of the agents converge to some point of the convex metric space with probability one.

IV. Preliminary results

In this section we construct the stochastic dynamics of the vector of distances between agents. Let t_i be a time-instant at which counter $N_i(t)$ increments its value. Then according to the gossip algorithm, at time t_i^+ the distance between agents i and j is given by

$$d(x_i(t_i^+), x_j(t_i^+)) = d(\psi(x_i(t_i), x_l(t_i), \lambda_i), x_j(t_i)), \quad (2)$$

with probability $p_{i,l}$.

Let $\theta_i(t)$ be an independent and identically distributed (i.i.d.) random process, such that $Pr(\theta_i(t) = l) = p_{i,l}$ for all $l \in \mathcal{N}_i$ and for all t. It follows that (2) can be equivalently written as

$$d(x_i(t_i^+), x_j(t_i^+)) = \sum_{l\in\mathcal{N}_i} \chi_{\{\theta_i(t_i)=l\}} d(\psi(x_i(t_i), x_l(t_i), \lambda_i), x_j(t_i)), \quad (3)$$

where $\chi_{\{\cdot\}}$ denotes the indicator function. Using the inequality property of the convex structure introduced in Definition 2.1, we further get

$$d(x_i(t_i^+), x_j(t_i^+)) \leq \lambda_i d(x_i(t_i), x_j(t_i)) + +(1-\lambda_i) \sum_{l\in\mathcal{N}_i} \chi_{\{\theta_i(t_i)=l\}} d(x_l(t_i), x_j(t_i)). \quad (4)$$

Assuming that t_j is a time-instant at which the Poisson counter $N_j(t)$ increments its value, in a similar manner as above we get that

$$d(x_i(t_j^+), x_j(t_j^+)) \leq \lambda_j d(x_i(t_j), x_j(t_j)) + +(1-\lambda_j) \sum_{l\in\mathcal{N}_j} \chi_{\{\theta_j(t_j)=l\}} d(x_l(t_j), x_i(t_j)). \quad (5)$$

Consider now the scalars $\eta_{i,j}(t)$ which follow the same dynamics as the distance between agents i and j, but with equality, that is,

$$\eta_{i,j}(t_i^+) = \lambda_i \eta_{i,j}(t_i) + (1-\lambda_i) \sum_{l\in\mathcal{N}_i} \chi_{\{\theta_i(t_i)=l\}} \eta_{j,l}(t_i), \quad (6)$$

and

$$\eta_{i,j}(t_j^+) = \lambda_j \eta_{i,j}(t_j) + (1-\lambda_j) \sum_{l\in\mathcal{N}_j} \chi_{\{\theta_j(t_j)=l\}} \eta_{i,l}(t_j), \quad (7)$$

with $\eta_{i,j}(0) = d(x_i(0), x_j(0))$.

Remark 4.1: Note that the index pair of η refers to the distance between two agents i and j. As a consequence $\eta_{i,j}$ and $\eta_{j,i}$ will be considered the same objects, and counted only once.

Proposition 4.1 ([10]): The following inequalities are satisfied with probability one:

$$\eta_{i,j}(t) \geq 0, \quad (8)$$

$$\eta_{i,j}(t) \leq \max_{i,j} \eta_{i,j}(0), \quad (9)$$

$$d(x_i(t), x_j(t)) \leq \eta_{i,j}(t), \quad (10)$$

for all $i \neq j$ and $t \geq 0$.

We now construct the stochastic differential equation satisfied by $\eta_{i,j}(t)$. From equations (6) and (7) we note that $\eta_{i,j}(t)$ at time t_i and t_j satisfies the solution of a stochastic differential equation driven by Poisson counters. Namely, we have

$$d\eta_{i,j}(t) = \left[-(1-\lambda_i)\eta_{i,j}(t) + (1-\lambda_i)\sum_{l\in\mathcal{N}_i} \chi_{\{\theta_i(t)=l\}} \eta_{j,l}(t)\right] dN_i(t) + \left[-(1-\lambda_j)\eta_{i,j}(t) + (1-\lambda_j)\sum_{m\in\mathcal{N}_j} \chi_{\{\theta_j(t)=m\}} \eta_{i,m}(t)\right] dN_j(t). \quad (11)$$

Let us now define the $\bar{n}$ dimensional vector $\boldsymbol{\eta} = (\eta_{i,j})$, where $\bar{n} = \frac{n(n-1)}{2}$ (since (i,j) and (j,i) correspond to the same distance variable). Equation (11) can be compactly written as

$$d\boldsymbol{\eta}(t) = \sum_{(i,j),i\neq j} \Phi_{i,j}(\theta_i(t))\boldsymbol{\eta}(t) dN_i(t) + \sum_{(i,j),i\neq j} \Psi_{i,j}(\theta_j(t))\boldsymbol{\eta}(t) dN_j(t). \quad (12)$$

where the $\bar{n}\times\bar{n}$ dimensional matrices $\Phi_{i,j}(\theta_i(t))$ and $\Psi_{i,j}(\theta_j(t))$ are defined as:

$$\Phi_{i,j}(\theta_i(t)) = \begin{cases} -(1-\lambda_i) & \text{at entry } [(i,j)(i,j)] \\ (1-\lambda_i)\chi_{\{\theta_i(t)=l\}} & \text{at entries } [(i,j)(l,j)], \\ & \quad l\in\mathcal{N}_i,\ l\neq j, l\neq i, \\ 0 & \text{all other entries ,} \end{cases} \quad (13)$$

and

$$\Psi_{i,j}(\theta_j(t)) = \begin{cases} -(1-\lambda_j) & \text{at entry } [(i,j)(i,j)] \\ (1-\lambda_j)\chi_{\{\theta_j(t)=m\}} & \text{at entries } [(i,j)(m,i)], \\ & \quad m\in\mathcal{N}_j, m\neq j, m\neq i, \\ 0 & \text{all other entries .} \end{cases} \quad (14)$$

The dynamics of the first moment of the vector $\boldsymbol{\eta}(t)$ is given by

$$\frac{d}{dt}E\{\boldsymbol{\eta}(t)\} = \sum_{(i,j),i\neq j} E\left\{\Phi_{i,j}(\theta_i(t))\boldsymbol{\eta}(t)\mu_i + \Psi_{i,j}(\theta_j(t))\boldsymbol{\eta}(t)\mu_j\right\}. \quad (15)$$

Using the independence of the random processes $\theta_i(t)$, we can further write

$$\frac{d}{dt}E\{\boldsymbol{\eta}(t)\} = \boldsymbol{W}E\{\boldsymbol{\eta}(t)\}, \quad (16)$$

where $\boldsymbol{W}$ is a $\bar{n}\times\bar{n}$ dimensional matrix whose entries are given by

$$[\boldsymbol{W}]_{(i,j),(l,m)} = \begin{cases} -(1-\lambda_i)\mu_i - (1-\lambda_j)\mu_j & l=i \text{ and } m=j \\ (1-\lambda_i)\mu_i p_{i,l} & l\in\mathcal{N}_i,\ m=j,\ l\neq j, \\ (1-\lambda_j)\mu_j p_{j,m} & l=i,\ m\in\mathcal{N}_j,\ m\neq i, \\ 0 & \text{otherwise.} \end{cases} \quad (17)$$

The following Lemma studies the properties of the matrix $\boldsymbol{W}$, introduced above.

Lemma 4.1 ([10]): Let $\boldsymbol{W}$ be the $\bar{n}\times\bar{n}$ dimensional matrix defined in (17). Under Assumption 3.1, the following properties hold:

(a) Let $\bar{G}$ be the directed graph (without self loops) corresponding to the matrix $\boldsymbol{W}$, that is, a link from (l,m) to (i,j) exists in $\bar{G}$ if $[\boldsymbol{W}]_{(i,j),(l,m)} > 0$. Then $\bar{G}$ is strongly connected.

(b) The row sums of the matrix $\boldsymbol{W}$ are non-positive, i.e.,

$$\sum_{(l,m),l\neq m} [\boldsymbol{W}]_{(i,j)(l,m)} \leq 0,\ \forall (i,j), i\neq j.$$

(c) There exits at least one row (i^*,j^*) of $\boldsymbol{W}$ whose sum is negative, that is,

$$\sum_{(l,m),l\neq m} [\boldsymbol{W}]_{(i^*,j^*)(l,m)} < 0.$$

Let $\xi_i \triangleq (1-\lambda_i)\mu_i$ and $\xi_j \triangleq (1-\lambda_j)\mu_j$ and consider now the matrix $\boldsymbol{Q} \triangleq I+\epsilon \boldsymbol{W}$, where I is the identity matrix and ϵ is a positive scalar satisfying $0<\epsilon<\frac{1}{\max_{i,j}(\xi_i+\xi_j)}$.

The following Corollary follows from the previous Lemma and describes the properties of the matrix $\boldsymbol{Q}$.

Corollary 4.1 ([10]): Matrix $\boldsymbol{Q}$ has the following properties:

(a) The directed graph (without self loops) corresponding to matrix $\boldsymbol{Q}$ (that is, a link from (l,m) to (i,j) exists if $[\boldsymbol{Q}]_{(i,j),(l,m)}>0$) is strongly connected.

(b) $\boldsymbol{Q}$ is a non-negative matrix with positive diagonal elements.

(c) The rows of $\boldsymbol{Q}$ sum up to a positive value not larger than one, that is,

$$\sum_{(l,m),l\neq m} [\boldsymbol{Q}]_{(i,j)(l,m)} \leq 1,\ \forall (i,j).$$

(d) There exits at least one row (i^*,j^*) of $\boldsymbol{Q}$ which sums up to a positive value strictly smaller than one, that is,

$$\sum_{(l,m),l\neq m} [\boldsymbol{Q}]_{(i^*,j^*)(l,m)} < 1.$$

Remark 4.2: The above Corollary says that the matrix $\boldsymbol{Q}$ is an *irreducible*, *substochastic* matrix. In addition, choosing $\gamma \geq \max_{i,j}\frac{1}{[\boldsymbol{Q}]_{(i,j)(i,j)}}$, it follows that we can find a non-negative, irreducible matrix $\tilde{\boldsymbol{Q}}$ such that $\gamma\boldsymbol{Q} = I+\tilde{\boldsymbol{Q}}$. Using a result on converting non-negativity and irreducibility to positivity ([11], page 672), we get that $(I+\tilde{\boldsymbol{Q}})^{\bar{n}-1}=\gamma^{\bar{n}-1}\boldsymbol{Q}^{\bar{n}-1}>0$, and therefore $\boldsymbol{Q}$ is a *primitive* matrix. The existence of γ is guaranteed by the fact that $\boldsymbol{Q}$ has positive diagonal entries.

We have the following result on the spectral radius of $\boldsymbol{Q}$, denoted by $\rho(\boldsymbol{Q})$.

Lemma 4.2 ([10]): The spectral radius of the matrix $\boldsymbol{Q}$ is smaller than one, that is,

$$\rho(\boldsymbol{Q}) < 1.$$

V. Proof of the main results

In this section we prove our main results introduced in Section III.

A. Proof of Theorem 3.1

We first show that the vector $\boldsymbol{\eta}(t)$ converges to zero in mean. By Lemma 4.2 we have that the spectral radius of $\boldsymbol{Q}$ is smaller than one, that is

$$\rho(\boldsymbol{Q}) < 1,$$

where $\rho(\boldsymbol{Q}) = \max_{\bar{i}}|\lambda_{\bar{i},\boldsymbol{Q}}|$, with $\lambda_{\bar{i},\boldsymbol{Q}}$, $\bar{i}=1,\ldots,\bar{n}$ being the eigenvalues of $\boldsymbol{Q}$. This also means that

$$Re(\lambda_{\bar{i},\boldsymbol{Q}}) < 1,\ \forall \bar{i}. \tag{18}$$

But since $\boldsymbol{W}=\frac{1}{\epsilon}(\boldsymbol{Q}-I)$, it follows that the real part of the eigenvalues of $\boldsymbol{W}$ are given by

$$Re(\lambda_{\bar{i},\boldsymbol{W}}) = \frac{1}{\epsilon}\left(Re(\lambda_{\bar{i},\boldsymbol{Q}})-1\right) < 0,\ \forall \bar{i},$$

where the last inequality follows from (18). Therefore, the dynamics

$$\frac{d}{dt}E\{\boldsymbol{\eta}(t)\} = \boldsymbol{W}E\{\boldsymbol{\eta}(t)\}$$

is asymptotically stable, and hence $\boldsymbol{\eta}(t)$ converges in mean to zero.

We now show that $\boldsymbol{\eta}(t)$ converges in r^{th} mean, for any $r\geq 1$. We showed above that $\eta_{i,j}(t)$ converges in mean to zero, for any $i\neq j$. But this also implies that $\eta_{i,j}(t)$ converges to zero in probability (Theorem 3, page 310, [4]), and therefore, for any $\delta>0$

$$\lim_{t\to\infty} Pr(\eta_{i,j}(t)>\delta) = 0. \tag{19}$$

Using the indicator function, the quantity $\eta_{i,j}(t)$ can be expressed as

$$\eta_{i,j}(t) = \eta_{i,j}(t)\chi_{\{\eta_{i,j}(t)\leq\delta\}} + \eta_{i,j}(t)\chi_{\{\eta_{i,j}(t)>\delta\}},$$

for any $\delta>0$. Using (9) of Proposition 4.1 we can further write

$$\eta_{i,j}(t)^r \leq \delta^r\chi_{\{\eta_{i,j}(t)\leq\delta\}} + \left(\max_{i,j}\eta_{i,j}(0)\right)^r\chi_{\{\eta_{i,j}(t)>\delta\}},$$

where to obtain the previous inequality we used the fact that $\chi_{\{\eta_{i,j}(t)\leq\delta\}}\chi_{\{\eta_{i,j}(t)>\delta\}}=0$. Using the expectation operator, we obtain

$$E\{\eta_{i,j}(t)^r\} \leq \delta^r \Pr(\eta_{i,j}(t)\leq\delta) + \left(\max_{i,j}\eta_{i,j}(0)\right)^r Pr(\eta_{i,j}(t)>\delta).$$

Taking t to infinity results in

$$\limsup_{t\to\infty} E\{\eta_{i,j}(t)^r\} \leq \delta^r,\ \forall\delta>0,$$

and since δ can be made arbitrarily small, we have that

$$\lim_{t\to\infty} E\{\eta_{i,j}(t)^r\} = 0,\ \forall r\geq 1.$$

Using (10) of Proposition 4.1, the result follows.

B. Proof of Theorem 3.2

In the following we show that $\boldsymbol{\eta}(t)$ converges to zero almost surely. Equations (6) and (7) show that with probability one $\eta_{i,j}(t)$ is non-negative and that for any $t_2\leq t_1$, with probability one $\eta_{i,j}(t_2)$ belongs to the convex hull generated by $\{\eta_{l,m}(t_1) \mid \text{for all pairs } (l,m)\}$. But this also implies that with probability one

$$\max_{i,j}\eta_{i,j}(t_2) \leq \max_{i,j}\eta_{i,j}(t_1). \tag{20}$$

Hence for any sample path of the random process $\boldsymbol{\eta}(t)$, the sequence $\{\max_{i,j}\eta_{i,j}(t)\}_{t\geq 0}$ is monotone decreasing and lower bounded. Using the monotone convergence theorem, we have that for any sample path ω, there exits $\tilde{\eta}(\omega)$ such that

$$\lim_{t\to\infty}\max_{i,j}\eta_{i,j}(t,\omega) = \tilde{\eta}(\omega),$$

or similarly

$$Pr\left(\lim_{t\to\infty}\max_{i,j}\eta_{i,j}(t)=\tilde{\eta}\right)=1.$$

In the following we show that $\tilde{\eta}$ must be zero with probability one. We achieve this by showing that there exits a subsequence of $\{\max_{i,j}\eta_{i,j}(t)\}_{t\geq 0}$ that converges to zero with probability one.

In Theorem 3.1 we proved that $\boldsymbol{\eta}(t)$ converges to zero in the r^{th} mean. Therefore, for any pair (i,j) and (l,m) we have that $E\{\eta_{i,j}(t)\eta_{l,m}(t)\}$ converge to zero. Moreover, since

$$E\{\eta_{i,j}(t)\eta_{l,m}(t)\}\leq\max_{i,j}\eta_{i,j}(0)E\{\eta_{l,m}(t)\},$$

and since $E\{\eta_{l,m}(t)\}$ converges to zero exponentially, we have that $E\{\eta_{i,j}(t)\eta_{l,m}(t)\}$ converges to zero exponentially as well.

Let $\{t_k\}_{k\geq 0}$ be a time sequence such that $t_k=kh$, for some $h>0$. From above, it follows that $E\{\|\boldsymbol{\eta}(t_k)\|^2\}$ converges to zero geometrically. But this is enough to show that the sequence $\{\boldsymbol{\eta}(t_k)\}_{k\geq 0}$ converges to zero, with probability one by using the Borel-Cantelli lemma (Theorem 10, page 320, [4]). Therefore, $\tilde{\eta}$ must be zero. Using (10) of Proposition 4.1, the result follows.

VI. The rate of convergence of the generalized gossip consensus algorithm under complete and uniform connectivity

Under specific assumptions on the topology of the graph, on the parameters of the Poisson counters and on the convex structure, in the following we obtain more explicit results on the rate of convergence of the algorithm.

Assumption 6.1: The Poisson counters have the same rate, that is $\mu_i=\mu$ for all i. Additionally, the parameters used by the agents in the convex structure are equal, that is $\lambda_i=\lambda$, for all i. In the update mode, each agent i picks one of the rest $n-1$ agents uniformly, that is $\mathcal{N}_i=\mathcal{N}-\{i\}$ and $p_{i,j}=\frac{1}{n-1}$, for all $j\in\mathcal{N}_i$.

The following two Propositions give upper bounds on the rate of convergence for the first and second moments of the distance between agents, under Assumption 6.1.

Proposition 6.1 ([10]): Under Assumptions 3.1, 3.2 and 6.1, the first moment of the distances between agents' states, using the generalized gossip algorithm converges exponentially to zero, that is

$$E\{d(x_i(t),x_j(t))\}\leq c_1e^{\alpha_1 t},\ \text{for all pairs } (i,j),\qquad(21)$$

where $\alpha_1=-\frac{2(1-\lambda)\mu}{n-1}$ and c_1 is a positive scalar depending of the initial conditions.

Proposition 6.2 ([10]): Under Assumptions 3.1, 3.2 and 6.1, the second moment of the distances between agents' states, using the generalized gossip algorithm converges exponentially to zero, that is

$$E\{d(x_i(t),x_j(t))^2\}\leq c_2e^{\alpha_2 t},\ \text{for all pair } (i,j),$$

where $\alpha_2=-\mu\frac{2(1-\lambda^2)}{n-1}$ and c_2 is a positive scalar depending of the initial distances between agents.

Remark 6.1: As expected, the eigenvalues α_1 and α_2 approach zero as n approaches infinity, and therefore the rate of converges decreases. Interestingly, in both the first and the second moment analysis, we observe that the minimum values of α_1 and α_2 are attained for $\lambda=0$, that is when an awaken agent never picks its own value, but the value of a neighbor.

VII. The generalized gossip consensus algorithm for particular convex metric spaces

In this section we present several instances of the gossip algorithm for particular examples of convex metric spaces. We consider three cases for $\mathcal{X}$: the set of real numbers, the set of compact intervals and the set of discrete random variables. We endow each of these sets with a metric d and convex structure ψ in order to form convex metric spaces. We show the particular form the generalized gossip algorithm takes for these convex metric spaces.

A. The set of real numbers

Let $\mathcal{X}=\mathbb{R}$ and consider as metric the standard Euclidean norm, that is $d(x,y)=\|x-y\|_2$, for any $x,y\in\mathbb{R}$. Consider the mapping $\psi(x,y,\lambda)=\lambda x+(1-\lambda)y$ for all $x,y\in\mathbb{R}$ and $\lambda\in[0,1]$. It can be easily shown that $(\mathbb{R},\|\cdot\|_2,\psi)$ is a convex metric space. For this particular convex metric space, the generalized randomized consensus algorithm takes the following form.

Note that this algorithm is exactly the randomized gossip algorithm for solving the consensus problem, which was studied in [2].

Algorithm 1: Randomized gossip algorithm on $\mathbb{R}$

Input: $x_i(0)$, λ_i, $p_{i,j}$
for *each counting instant t_i of N_i* **do**
  Agent i enters update mode and picks a neighbor j with probability $p_{i,j}$;
  Agent i updates its state according to

$$x_i(t_i^+)=\lambda_ix_i(t_i)+(1-\lambda_i)x_j(t_i);$$

  Agent i enters sleep mode;

B. The set of compact intervals

Let $\mathcal{X}$ be the family of closed intervals, that is $\mathcal{X}=\{[a,b]\mid -\infty<a\leq b<\infty\}$. For $x_i=[a_i,b_i]$, $x_j=[a_j,b_j]$ and $\lambda\in[0,1]$, we define a mapping ψ by $\psi(x_i,x_j,\lambda)=[\lambda a_i+(1-\lambda)a_j,\lambda b_i+(1-\lambda)b_j]$ and use as metric the Hausdorff distance given by $d(x_i,x_j)=\max\{|a_i-a_j|,|b_i-b_j|\}$. Then, as shown in [22], $(\mathcal{X},d,\psi)$ is a convex metric space. For this convex metric space, the randomized gossip consensus algorithm is given bellow.

C. The set of discrete random variables

In this section we apply our algorithm on a particular convex metric space that allows us the obtain a probabilistic algorithm for reaching consensus on discrete sets.

Let $S=\{s_1,s_2,\ldots,s_m\}$ be a finite and countable set of real numbers and let $(\Omega,\mathcal{F},\mathcal{P})$ be a probability space. We denote

Algorithm 2: Randomized gossip algorithm on a set of compact intervals

Input: $x_i(0)$, λ_i, $p_{i,j}$
for *each counting instant* t_i *of* N_i **do**
  Agent i enters update mode and picks a neighbor j with probability $p_{i,j}$;
  Agent i updates its state according to
  $$x_i(t_i^+) = [\lambda_i a_i(t_i) + (1-\lambda_i)a_j(t_i), \lambda_i b_i(t_i) + (1-\lambda_i)b_j(t_i)];$$
  Agent i enters sleep mode;

by $\mathcal{X}$ the space of discrete measurable functions (random variable) on $(\Omega, \mathcal{F}, \mathcal{P})$ with values in S.

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

$$d(X,Y) = E_{\mathcal{P}}[\rho(X,Y)], \tag{22}$$

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

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

and the expectation is taken with respect to the measure $\mathcal{P}$. It can be readily shown that the above mapping $d(\cdot,\cdot)$ is a metric on $\mathcal{X}$ and therefore $(\mathcal{X}, d)$ is a metric space.

Let $\gamma \in \{1,2\}$ be an independent random variable defined on the probability space $(\Omega, \mathcal{F}, \mathcal{P})$, with probability mass function $Pr(\gamma = 1) = \lambda$ and $Pr(\gamma = 2) = 1 - \lambda$, where $\lambda \in [0,1]$. We define the mapping $\psi : \mathcal{X} \times \mathcal{X} \times [0,1] \to \mathcal{X}$ given by

$$\psi(X_1, X_2, \lambda) = \mathbb{1}_{\{\gamma=1\}} X_1 + \mathbb{1}_{\{\gamma=2\}} X_2, \forall X_1, X_2 \in \mathcal{X}, \lambda \in [0,1]. \tag{23}$$

Proposition 7.1 ([8]): The mapping ψ is a convex structure on $\mathcal{X}$.

From the above proposition it follows that $(\mathcal{X}, d, \psi)$ is a *convex metric space*. For this particular convex metric space the randomized consensus algorithm is summarized in what follows.

Algorithm 3: Randomized gossip algorithm on countable, finite sets

Input: $x_i(0)$, λ_i, $p_{i,j}$
for *each counting instant* t_i *of* N_i **do**
  Agent i enters update mode and picks a neighbor j with probability $p_{i,j}$;
  Agent i updates its state according to
  $$x_i(t_i^+) = \begin{cases} x_i(t_i) & \text{with probability } \lambda_i \\ x_j(t_i) & \text{with probability } 1-\lambda_i \end{cases}$$
  Agent i enters sleep mode;

VIII. Conclusions

In this paper we analyzed the convergence properties of a generalized randomized gossip algorithm acting on convex metric spaces. We gave convergence results in almost sure and r^{th} mean sense for the distances between the states of the agents. Under specific assumptions on the communication topology, we computed explicitly estimates of the rate of convergence for the first and second moments of the distances between the agents. Additionally, we introduced instances of the generalized gossip algorithm for three particular convex metric spaces.

References

[1] V. Borkar and P. Varaya. Asymptotic agreement in distributed estimation. *IEEE Trans. Autom. Control*, AC-27(3):650–655, Jun 1982.

[2] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah. Randomized gossip algorithms. *IEEE/ACM Trans. Netw.*, 14(SI):2508–2530, 2006.

[3] GN. Dunford and J.T. Schwartz. *Linear operators, Part 1.* Interscience, New York, 1958.

[4] G. Grimmet and D. Stirzaker. *Probability and Random Processes*. Oxford University Press, third edition edition, 2001.

[5] Y. Hatano and M. Mesbahi. Agreement over random networks. *IEEE Trans. Autom. Control*, 50(11):1867–1872, 2005.

[6] A. Jadbabaie, J. Lin, and A.S. Morse. Coordination of groups of mobile autonomous agents using nearest neighbor. *IEEE Trans. Autom. Control*, 48(6):998–1001, Jun 2004.

[7] I. Matei and J. Baras. The asymptotic consensus problem on convex metric spaces. *Proceedings of the 2nd IFAC Workshop on Distributed Estimation and Control in Networked Systems(NecSys 2010)*, pages 287–292, September 2010.

[8] I. Matei and J. Baras. The geometry of a probabilistic consensus of opinion algorithm. *Proceedings of the 2011 American Control Conference*, pages 2198–2203, June 2011.

[9] I. Matei, N. Martins, and J. Baras. Consensus problems with directed markovian communication patterns. *Proceedings of the 2009 IEEE American Control Conference*, pages 1298–1303, Jun 2009.

[10] I. Matei, C. Somarakis, and J.S. Baras. A randomized gossip consensus algorithm on convex metric spaces (http://hdl.handle.net/1903/12413). Technical report, University of Maryland, College Park, February 2012.

[11] C. D. Meyer. *Matrix Analysis and Applied Linear Algebra*. SIAM: Society for Industrial and Applied Mathematics, 2000.

[12] L. Moreau. Stability of multi-agents systems with time-dependent communication links. *IEEE Trans. Autom. Control*, 50(2):169–182, Feb 2005.

[13] A. Nedic, A. Ozdaglar, and P.A. Parrilo. Constrained consensus and optimization in multi-agent networks. *IEEE Trans. Autom. Control*, 55(4):922–938, Apr 2010.

[14] R. Olfati-Saber and R.M. Murray. Consensus problem in networks of agents with switching topology and time-delays. *IEEE Trans. Autom. Control*, 49(9):1520–1533, Sept 2004.

[15] Silvana Silva Pereira and Alba Pagès-Zamora. Consensus in correlated random wireless sensor networks. *IEEE Transactions on Signal Processing*, 59(12):6279–6284, 2011.

[16] M. Porfiri and D.J. Stilwell. Consensus seeking over random directed weighted graphs. *IEEE Trans. Autom. Control*, 52(9):1767–1773, Sept 2007.

[17] W. Ren and R.W. Beard. Consensus seeking in multi-agents systems under dynamically changing interaction topologies. *IEEE Trans. Autom. Control*, 50(5):655–661, May 2005.

[18] A. Tahbaz Salehi and A. Jadbabaie. Necessary and sufficient conditions for consensus over random networks. *IEEE Trans. Autom. Control*, 53(3):791–795, Apr 2008.

[19] A. Tahbaz Salehi and A. Jadbabaie. Consensus over ergodic stationary graph processes. *IEEE Trans. Autom. Control*, 55(1):225–230, Jan 2010.

[20] D. Shah. Gossip algorithms. *Foundations and Trends in Networking*, 3(1):1–125, 2008.

[21] B.K. Sharma and C.L. Dewangan. Fixed point theorem in convex metric space. *Novi Sad Journal of Mathematics*, 25(1):9–18, 1995.

[22] W. Takahashi. A convexity in metric mpace and non-expansive mappings i. *Kodai Math. Sem. Rep.*, 22:142–149, 1970.

[23] J.N. Tsitsiklis. Problems in decentralized decision making and computation. *Ph.D. dissertation, Dept. Electr. Eng.*, Nov 1984.

[24] J.N. Tsitsiklis, D.P. Bertsekas, and M. Athans. Distributed asynchronous deterministic and stochastic gradient optimization algorithms. *IEEE Trans. Autom. Control*, 31(9):803–812, Sept 1986.