Distributed Belief Averaging Using Sequential Observations

Yang Liu Ji Liu Tamer Başar Mingyan Liu

***Abstract*— This paper considers a distributed belief averaging problem with sequential observations in which a group of $n > 1$ agents in a network, each having sequentially arriving samples of its belief in an online manner, aim to reach a consensus at the average of their beliefs, by exchanging information only with their neighbors. The neighbor relationships among the n agents are described by a time-varying undirected graph whose vertices correspond to agents and whose edges depict neighbor relationships. A distributed algorithm is proposed to solve this problem over sequential observations with $O(\frac{1}{t})$ convergence rate. Extensions to the case of directed graphs are also detailed.**

I. Introduction

Over the past few decades, there has been considerable interest in developing algorithms for information distribution and computation among members of interactive agents via peer-to-peer interactions [1]–[3]. Recently, distributed computation and decision making problems of all types have arisen naturally, such as consensus problems [4], multi-agent coverage problems [5], rendezvous problems [6], multi-sensor localization [7], and multi-robot formation control [8]. These problems have found applications in different fields, including sensor networks [9], robotic teams [10], social networks [11], and electric power grids [12]. Distributed computation and control are promising for large-scale complex networks due to their fault tolerance and cost saving features, and their ability to accommodate various physical constraints such as limitations on sensing, computation, and communication.

One important problem in distributed control and computation is the consensus problem [4], [13]–[22]. In a typical consensus process, the agents in a given group all try to agree on some quantity by communicating what they know only to their neighboring agents. One particular type of consensus process, whose goal is to compute the average of the initial values of the quantity of interest at the agents, is called distributed averaging [23]. Solutions for such conventional distributed averaging problems have been elegantly developed, including linear iterations [9], [23], [24], gossiping [25], [26], and push-sum [27], which are also known as weighted gossip [28], ratio consensus [29], and double linear iterations [30].

In this paper, we extend the conventional distributed averaging problem setting to the case when each distributed agent has its local belief/measurement arriving sequentially. This is different from previous studies whereby it was assumed that each agent i at initial time $t = 1$ holds $x_i(1)$, which equals the true belief, and the subsequent averaging and communication processes are carried out entirely over $x_i(1)$, $i \in \{1, 2, \ldots, n\}$. By contrast, in this paper, we consider the case where each agent's true belief $\bar{x}_i$ is unknown and can only be measured through a series of local observations; these will be denoted by $x_i(t)$, $t \in \{1, 2, \ldots\}$ and are only available to agent i.

The problem considered here thus falls in the family of problems on distributed learning and control. However, a unique feature of the problem is the departure from the conventional distributed averaging process. This difference is elaborated below. Consider the motivating example of a sensor network where each sensor needs to take a sequence of local measurements in order to obtain an accurate estimate of a local mean (or belief as referred to earlier) due to environmental and instrumental uncertainties/noises; at the same time we have the goal of estimating the global mean through message exchanges among sensors. Note that several previous studies, e.g. [31], [32], also looked into distributed learning settings with noisy observations. The key differences between those studies and our work here are, among others, the following: (1) as in [31], uncertainties are often modeled as coming from external noise sources which are independent of the sample observations, which is not the case here; (2) many existing studies such as [32] are formulated within a distributed estimation framework, in which observability of the system is typically assumed, while we do not impose any such condition here.

The above description leads us to conclude that in our case there are multiple averaging processes going on simultaneously, one globally among sensors, and a local one at each sensor. This is the main difference between our problem and the classical literature on distributed averaging and consensus which only focuses on the global averaging process by assuming that the local mean is already available. In our setting, it would be desirable to embed the multiple underlying averaging processes into the same distributed updating procedure, by integrating the new measurements/samples into the single averaging process as they arrive; this is the goal of the present study. To this end, we introduce an algorithm referred to as *distributed belief averaging using sequential observations*, and formally establish its convergence and determine its convergence rate.

The main contribution of this paper is two-fold. First, we introduce a distributed belief averaging algorithm with sequential observations and establish its polynomial conver-

This research was supported in part by AFOSR MURI Grant FA 9550-10-1-0573, in part by Office of Naval Research (ONR) MURI Grant N00014-16-1-2710, and in part by NSF under grant CCF 11-11342. Y. Liu is with Harvard University (yangl@seas.harvard.edu). J. Liu and T. Başar are with the Coordinated Science Laboratory, University of Illinois at Urbana-Champaign ({jiliu, basar1}@illinois.edu). M. Liu is with University of Michigan (mingyan@eecs.umich.edu).

gence to consensus when the underlying graphs are undirected. Second, we extend the results to establish asymptotic convergence for the case of directed graphs.

The rest of the paper is organized as follows. The problem is formulated and stated in Section II. An algorithm solving the problem, along with a result on its convergence rate for undirected graphs, are presented in Section III, and the extension to directed graphs is discussed in the same section. The analysis of the algorithm and the proof of the main result are given in Section IV, and Section V concludes the paper.

II. Problem Formulation

Consider a network consisting of $n > 1$ agents. For ease of presentation, we label the agents 1 through n. The agents are not aware of such a global labeling, but each agent is able to identify its "neighbors". The neighbor relationships among the n agents are described by a time-varying n-vertex undirected graph $\mathbb{N}(t)$, called the *neighbor graph*, in which the n vertices represent the n agents and the edges indicate the neighbor relationships. In other words, a pair of agents i and j are neighbors at time t if and only if (i, j) is an edge in $\mathbb{N}(t)$. We use $\mathcal{N} = \{1, 2, \ldots, n\}$ to denote the set of agents (or vertices) and $\mathcal{N}_i(t)$ to denote the set of neighbors of agent i at time t. To be more precise,

$$\mathcal{N}_i(t) = \{j \in \mathcal{N}, j \neq i : (i, j) \text{ is an edge in } \mathbb{N}(t)\} \ , \ i \in \mathcal{N} \ .$$

We assume that time is discrete in that t takes values in $\{1, 2, \ldots\}$. Each agent i receives a real-valued scalar[1] $x_i(t)$ at each time t. We assume that the samples $\{x_i(t)\}_{t=1}^{\infty}$ form an i.i.d. process, and the sequence is generated according to a random variable X_i with distribution $f_{X_i}(\cdot)$. For simplicity, we assume that the support set for X_i is bounded, i.e., there exists a constant M such that $|X_i(\omega)| \leq M$ for all $i \in \mathcal{N}$ and ω, where ω indicates an arbitrary sample realization. Note that the agents' observations do not need to be identical, i.e., the $f_{X_i}(\cdot)$, $i \in \mathcal{N}$, do not need to be structurally the same. We use $\bar{x}_i$ to denote the expectation of agent i's local observations, i.e.,

$$\bar{x}_i = E[X_i] \ ,$$

and call $\bar{x}_i$ the *local belief* of agent i.

At each time step t, each agent i can exchange information only with its current neighbors $j \in \mathcal{N}_i(t)$. Thus, only local information is available to each agent, i.e., each agent i, only knows its own samples, the information received from its current neighbors in $\mathcal{N}_i(t)$, and nothing more, while the global connectivity patterns remain unknown to any agent in the network.

Suppose that each agent has control over a real-valued variable $y_i(t)$ which it can update from time to time. Let

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} \bar{x}_i \ , \tag{1}$$

which represents the *average belief* within the network. Our goal is to devise an algorithm for each agent i, under the local information constraint, which ensures that

$$\lim_{t \to \infty} y_i(t) = \bar{x} \ , \quad \forall i \ .$$

Moreover, we would like to characterize the convergence rate of such an algorithm.

III. Distributed Belief Averaging with Sequential Observations

In this section, we first propose a distributed algorithm for a time-varying undirected neighbor graph and establish its convergence under an appropriate connectivity assumption. A variant of the algorithm for directed neighbor graphs will be given subsequently.

A. Undirected Graphs

At initial time $t = 1$, each agent i sets its $y_i(1) = x_i(t)$, where $x_i(t)$ is as introduced earlier. For each time $t > 1$, each agent i first broadcasts its current $y_i(t)$ to all its current neighbors in $\mathcal{N}_i(t)$. At the same time, each agent i receives $y_j(t)$ from all its neighbors $j \in \mathcal{N}_i(t)$. Then, each agent i updates its variable by setting

$$\begin{aligned} y_i(t+1) \quad = \quad & w_{ii}(t) y_i(t) + \sum_{j \in \mathcal{N}_i(t)} w_{ij}(t) y_j(t) \\ & + z_i(t+1) - z_i(t) \ , \end{aligned} \tag{2}$$

where

$$z_i(t) = \frac{1}{t} \sum_{\tau=1}^{t} x_i(\tau) \ ,$$

and $w_{ij}(t)$ are real-valued weights such that the matrix $W(t) = [w_{ij}(t)]$ is a symmetric doubly stochastic matrix[2] with positive diagonal entries. Such $w_{ij}(t)$ weights can be designed in a distributed manner using the well-known Metropolis weights [9]. Specifically, $w_{ij}(t) = 0$ if $i \neq j$ and (i, j) is not an edge in $\mathbb{N}(t)$, and

$$w_{ij}(t) = \frac{1}{1 + \max\{n_i(t), n_j(t)\}} \ , \qquad (i, j) \in \mathbb{N}(t) \ ,$$

$$w_{ii}(t) = 1 - \sum_{j=1}^{n} w_{ij}(t) \ .$$

where $n_i(t)$ is the number of neighbors of agent i at time t, or equivalently, the degree of vertex i in $\mathbb{N}(t)$. Thus, the updating of each $y_i(t)$ only requires the information from agent i's current neighbors.

It is worth emphasizing that even though $z_i(t+1)$ and $z_i(t)$ are used in the updating of $y_i(t)$, an agent does not have to store all its received samples. Instead, each agent i only needs to keep track of $z_i(t)$ since

$$z_i(t+1) = \frac{t z_i(t) + x_i(t+1)}{t+1} \ .$$

[1] The results in this paper can be straightforwardly extended to the vector-valued case.

[2] A square matrix is called a *stochastic matrix* if all its entries are nonnegative and its row sums all equal to 1. A stochastic matrix is called a *doubly stochastic matrix* if also its column sums all equal to 1

In vector form, the n equations in (2) can be written as

$$y(t+1) = W(t)y(t) + z(t+1) - z(t) \ . \tag{3}$$

where $y(t)$ and $z(t)$ are the column vectors obtained by stacking up $y_i(t)$s and $z_i(t)$s, respectively.

B. Convergence

To state the convergence result, we need the following concepts.

An undirected graph $\mathbb{G}$ is *connected* if there is a path between each pair of distinct vertices in $\mathbb{G}$. By the *union* of a finite sequence of undirected graphs, $\mathbb{G}_1, \mathbb{G}_2, \ldots, \mathbb{G}_p$, each with the vertex set $\mathcal{V}$, is meant the undirected graph $\mathbb{G}$ with vertex set $\mathcal{V}$ and edge set equaling the union of the edge sets of all of the graphs in the sequence. We say that such a finite sequence is *jointly connected* if the union of its members is a connected graph. We say that an infinite sequence of undirected graphs $\mathbb{G}_1, \mathbb{G}_2, \ldots$ is *repeatedly jointly connected* if there is a positive integer r such that for each $k \geq 0$, the finite sequence $\mathbb{G}_{rk+1}, \mathbb{G}_{rk+2}, \ldots, \mathbb{G}_{r(k+1)}$ is jointly connected.

Theorem 1: Suppose that all n agents adhere to the update rule (2) and that the sequence of neighbor graphs $\mathbb{N}(1), \mathbb{N}(2), \mathbb{N}(3), \ldots$ is repeatedly jointly connected. Then,

$$\lim_{t\to\infty} y_i(t) = \bar{x} \ , \qquad \forall i \ ,$$

with a convergence rate at the order of $O(\frac{1}{t})$ with high probability (w.h.p.).[3]

The proof of the theorem will be given in Section IV.

The basic intuition behind the proof is as follows. First, it is straightforward to verify that

$$\lim_{t\to\infty} \frac{1}{n}\sum_{i=1}^{n} z_i(t) = \frac{1}{n}\sum_{i=1}^{n} \bar{x}_i \ . \tag{4}$$

Since $y(1) = x(1) = z(1)$, from (3), it follows that $\mathbf{1}^\top y(2) = \mathbf{1}^\top z(2)$, where $\mathbf{1}$ denotes the vector whose entries all equal to 1 and $\mathbf{1}^\top$ denotes its transpose. By induction, it follows that $\mathbf{1}^\top y(t) = \mathbf{1}^\top z(t)$ for any t, which implies that

$$\sum_{i=1}^{n} y_i(t) = \sum_{i=1}^{n} z_i(t) \ , \qquad \forall t \ .$$

Ignoring the small perturbation terms $z(t+1)-z(t)$ (later we will show that this term converges polynomially fast to the zero vector $\mathbf{0}$), it is known that $y(t+1) = W(t)y(t)$ leads all $y_i(t)$ to the same value [24]. Note that, from (4) and (1), $\sum_{i=1}^{n} z_i(t)$ will converge to $n\bar{x}$, and therefore we could expect that each $y_i(t)$ will converge to $\bar{x}$ since $\mathbf{1}^\top y(t) = \mathbf{1}^\top z(t)$.

Remark 1: The above convergence result holds for any sequence of doubly stochastic matrices $\{W(t)\}$ satisfying the properties that each matrix has positive diagonal entries and all nonzero entries are uniformly bounded below by some positive number. Note that the matrices whose entries are the Metropolis weights satisfy these properties. The convergence result also holds for any static connected graph, which can be regarded as a special case of repeatedly jointly connected time-varying graphs. □

[3]Throughout the paper, when we say w.h.p. (with high probability), it implies that the probability goes to 1 when t goes to infinity.

C. Directed Graphs

The algorithm just described requires that the communication between any pair of neighboring agents is bidirectional. Such a requirement may not always be satisfied in applications. For example, different agents may have differing transmission radii. In this subsection, we propose another algorithm to deal with the case when the neighbor graph is directed, i.e., the communication between agents can be unidirectional. The algorithm makes use of the idea of the push-sum protocol [27] which solves the conventional distributed averaging problem for directed neighbor graphs.

Since the communication may be unidirectional, we need to modify the definition of "neighbors" in this section. Specifically, we say that an agent j is a *neighbor* of agent i at time t if agent i can send information to agent j at time t. In this case, we also say that agent i is an *observer* of agent j at time t. The neighbor (and observer) relationships among the n agents are described by a time-varying n-vertex directed graph $\mathbb{N}(t)$ in which the directions of directed edges represent the directions of information flow. Thus, agent j is a neighbor of agent i (and thus agent i is an observer of agent j) at time t if and only if (i,j) is a directed edge in $\mathbb{N}(t)$. We use $o_i(t)$ to denote the number of observers of agent i at time t, or equivalently, the out-degree of vertex i in $\mathbb{N}(t)$.

We propose the following algorithm for directed neighbor graphs. Each agent i has control over two real-valued variables $y_i(t)$ and $v_i(t)$ which are initialized as $y_i(1) = x_i(1)$ and $v_i(1) = 1$ respectively. At each time $t > 1$, each agent i sends the weighted current values $\frac{y_i(t)}{1+o_i(t)}$ and $\frac{v_i(t)}{1+o_i(t)}$ to all its current neighbors and updates its variables according to the rules

$$\begin{aligned} w_i(t+1) &= \frac{y_i(t)}{1+o_i(t)} + \sum_{j\in\mathcal{N}_i(t)} \frac{y_j(t)}{1+o_j(t)} \ , \\ v_i(t+1) &= \frac{v_i(t)}{1+o_i(t)} + \sum_{j\in\mathcal{N}_i(t)} \frac{v_j(t)}{1+o_j(t)} \ , \\ y_i(t+1) &= w_i(t+1) + z_i(t+1) - z_i(t) \ . \end{aligned} \tag{5}$$

The limiting behavior of the above algorithm also depends on the connectivity of the neighbor graph. A directed graph $\mathbb{G}$ is *strongly connected* if there is a directed path between each pair of distinct vertices in $\mathbb{G}$. We say that a finite sequence of directed graphs with the same vertex set is *jointly strongly connected* if the union of its members is a strongly connected graph. We say that an infinite sequence of directed graphs $\mathbb{G}_1, \mathbb{G}_2, \ldots$ is *repeatedly jointly connected* if there is a positive integer r such that for each $k \geq 0$, the finite sequence $\mathbb{G}_{rk+1}, \mathbb{G}_{rk+2}, \ldots, \mathbb{G}_{r(k+1)}$ is jointly strongly connected.

Proposition 1: Suppose that all n agents adhere to the update rule (5) and that the sequence of neighbor graphs $\mathbb{N}(1), \mathbb{N}(2), \mathbb{N}(3), \ldots$ is repeatedly jointly strongly connected. Then, w.h.p.,

$$\lim_{t\to\infty} \frac{w_i(t)}{v_i(t)} = \bar{x} \ , \qquad \forall i \ . \tag{6}$$

Proof: In the proof of Theorem 1 (given in the next section), it will be shown that $z_i(t+1) - z_i(t)$ will converge to 0 w.h.p.. Then, by Lemma 1(b) in [33], it follows that w.h.p.,

$$\lim_{t\to\infty} \left| \frac{w_i(t)}{v_i(t)} - \frac{1}{n}\sum_{j=1}^{n} y_j(t) \right| = 0$$

and the vector $y(t)$ will converge to a consensus vector as $t \to \infty$. It is straightforward to verify that $\mathbf{1}^\top y(t) = \mathbf{1}^\top z(t)$ for all time t. From (4), it follows that (6) holds w.h.p.. ■

The convergence rate of the algorithm is a subject of future research.

IV. Analysis

In this section, we provide a proof of the main result in Section III. Our approach will appeal to the stability properties of discrete-time linear consensus processes.

We begin with the idea of a certain semi-norm which was introduced in [34]. Let $||\cdot||$ be the induced infinity norm on $\mathbb{R}^{m\times n}$. For $M \in \mathbb{R}^{m\times n}$, define

$$|M|_\infty = \min_{c\in\mathbb{R}^{1\times n}} ||M - \mathbf{1}c|| \ .$$

It has been shown in [34] that $|\cdot|_\infty$ is a *semi-norm*, namely that it is positively homogeneous and satisfies the triangle inequality. Moreover, this particular semi-norm is *sub-multiplicative* in the following sense.

Lemma 1: (Lemma 1 in [22]) Suppose that $\mathcal{S}$ is a subset of $\mathbb{R}^{n\times n}$ such that $M\mathbf{1} = \mathbf{1}$ for all $M \in \mathcal{S}$. Then,

$$|Mx|_\infty \le |M|_\infty |x|_\infty \ , \qquad |MN|_\infty \le |M|_\infty |N|_\infty \ ,$$

for any $M, N \in \mathcal{S}$ and $x \in \mathbb{R}^n$.

For column vectors and nonnegative matrices, more can be said.

Lemma 2: (Lemmas 2 and 3 in [22]) Let x be a vector in $\mathbb{R}^n$. Then,

$$|x|_\infty = \frac{1}{2}\left(\max_i x_i - \min_j x_j\right) \ .$$

Let $A \in \mathbb{R}^{n\times n}$ be a nonnegative matrix. Then,

$$|A|_\infty = \frac{1}{2}\max_{i,j}\sum_{k=1}^{n} |a_{ik} - a_{jk}| \ .$$

It is worth noting that $|A|_\infty$ is the same as the well-known coefficient of ergodicity [35] if A is a stochastic matrix. Thus, $|A|_\infty \le 1$ if A is stochastic.

We now characterize stability properties of linear consensus processes using the semi-norm just introduced. We first introduce the notion of internal stability, which has been studied in [22].

Consider a discrete-time linear consensus process modeled by a linear recursion equation of the form

$$x(k+1) = S(k)x(k) \ , \qquad x(k_0) = x_0 \ , \tag{7}$$

where $x(k)$ is a vector in $\mathbb{R}^n$ and $S(k)$ is an $n\times n$ stochastic matrix. It is easy to verify that the equilibria of (7) include points of the form $a\mathbf{1}$. We say that (7) is *uniformly consensus stable* if there exists a finite positive constant γ such that for any k_0 and x_0, the corresponding solution satisfies

$$|x(k)|_\infty \le \gamma |x_0|_\infty \ , \qquad k \ge k_0 \ .$$

This definition is different from the standard notion of uniform stability of linear systems [36] since it uses the semi-norm instead of any norm. The same is true for all the definitions of stability in this section. Consensus stability of the system is equivalent to $x(t)$ converging to a consensus vector (i.e., all the entries of $x(t)$ have the same value). Note that any trajectory of (7) is bounded for all k since each $S(k)$ is stochastic. Thus, the system (7), without input or external disturbance, is uniformly consensus stable for any sequence $\{S(k)\}$.

We say that the system described by (7) is *uniformly exponentially consensus stable* if there exist a finite positive constant γ and a constant $0 \le \lambda < 1$ such that for any k_0 and x_0, the corresponding solution satisfies

$$|x(k)|_\infty \le \gamma\lambda^{k-k_0}|x_0|_\infty \ , \qquad k \ge k_0 \ .$$

Uniform exponential consensus stability implies uniform consensus stability and imposes an additional requirement that all solutions of (7) approach a consensus vector exponentially fast.

Let $\Phi(k,j)$ be the *discrete-time state transition matrix* of $S(k)$, i.e.,

$$\Phi(k,j) = \begin{cases} S(k-1)S(k-2)\cdots S(j) & \text{if } \ k > j \ , \\ I & \text{if } \ k = j \ . \end{cases}$$

It is easy to verify that $\Phi(k,j)$ is a stochastic matrix for any $k \ge j$. The following result characterizes uniform exponential consensus stability.

Lemma 3: (Theorem 1 in [22]) The discrete-time linear recursion equation (7) is uniformly exponentially consensus stable if and only if there exist a finite positive constant γ and a constant $0 \le \lambda < 1$ such that

$$|\Phi(k,j)|_\infty < \gamma\lambda^{k-j} \ ,$$

for all k, j such that $k \ge j$.

Exponential consensus stability can be characterized by graph connectivity. Toward this end, we need the following concept. The *graph of a nonnegative symmetric matrix* $M \in \mathbb{R}^{n\times n}$, denoted by $\gamma(M)$, is an undirected graph on n vertices with an edge between vertex i and vertex j if and only if $m_{ji} \ne 0$ (and thus $m_{ij} \ne 0$).

Lemma 4: Let $\mathcal{F}$ denote a compact subset of the set of all $n \times n$ symmetric stochastic matrices with positive diagonal entries. Suppose that $F(1), F(2), \ldots$ is an infinite sequence of matrices in $\mathcal{F}$. Then, the discrete-time linear recursion equation $x(k+1) = F(k)x(k)$, $k \geq 1$, is uniformly exponentially consensus stable if and only if the sequence of graphs $\gamma(F(1)), \gamma(F(2)), \gamma(F(3)) \ldots$ is repeatedly jointly connected.

This lemma is a direct consequence of Theorem 4 in [22].

Now we turn to input-output stability of discrete-time linear consensus processes. Toward this end, we rewrite the equation (7) in an input-output form as follows:

$$x(k+1) = S(k)x(k) + B(k)u(k) \ , \tag{8}$$

$$y(k) = C(k)x(k) \ . \tag{9}$$

We are interested in the case when $B(k)$ and $C(k)$ are stochastic matrices for all k. The input-output behavior of the system (8)-(9) is specified by the unit-pulse response

$$G(k,j) = C(k)\Phi(k,j+1)B(j) \ , \qquad k \geq j+1 \ ,$$

which is also a stochastic matrix.

We say that the system defined by (8)-(9) is *uniformly bounded-input, bounded-output consensus stable* if there exists a finite constant η such that for any k_0 and any input signal $u(k)$ the corresponding zero-state response satisfies

$$\sup_{k \geq k_0} |y(k)|_\infty \leq \eta \sup_{k \geq k_0} |u(k)|_\infty \ .$$

It is worth noting that $y(t)$ may not be bounded even though the system is uniformly bounded-input, bounded-output consensus stable, which is different from the standard notion of input-output stability of linear systems [36].

The following result establishes the connection between uniform bounded-input, bounded-output stability, a property of the zero-state response, and uniform exponential stability, a property of the zero-input response.

Proposition 2: (Theorem 2 in [37]) Suppose that (7) is uniformly exponentially consensus stable. Then, the system (8)-(9) is uniformly bounded-input, bounded-output consensus stable.

Now we are in a position to prove Theorem 1.

Proof of Theorem 1: The system (3) can be viewed as a linear consensus system with input $u(t) = z(t+1) - z(t)$ and $B(t) = C(t) = I$ where I is the identity matrix which is also a stochastic matrix. Then, the output

$$y(t) = \Phi(t,1)y(1) + \hat{y}(t) \ ,$$

where $\hat{y}(t)$ is the zero-state response and the first component on the right-hand side is the zero-input response. It follows that

$$\begin{aligned} |y(t)|_\infty &\leq |\Phi(t,1)y(1)|_\infty + |\hat{y}(t)|_\infty \\ &\leq |\Phi(t,1)y(1)|_\infty + \sup_{\tau \geq t} |\hat{y}(\tau)|_\infty \ . \end{aligned}$$

Since the sequence of neighbor graphs is repeatedly jointly connected, by Lemma 4, the system $y(t+1) = W(t)y(t)$ is uniformly exponentially consensus stable, and thus the system (3) is uniformly input-bounded, output-bounded consensus stable by Proposition 2. Since $|\Phi(t,1)y(1)|_\infty$ converges to 0 exponentially fast, to prove the theorem, it will be enough to show that $\sup_{\tau \geq t} |\hat{y}(\tau)|_\infty$ converges to 0 at the order of $O(\frac{1}{t})$ w.h.p.. Since $\sup_{\tau \geq t} |\hat{y}(\tau)|_\infty \leq \eta \sup_{\tau \geq t} |u(\tau)|_\infty$ for some constant η, and noting that by Lemma 2, $\sup_{\tau \geq t} |u(\tau)|_\infty \leq 2 \sup_{\tau \geq t} \max_i |u_i(\tau)|$, it will be sufficient to bound the convergence rate of $\sup_{\tau \geq t} \max_i |u_i(\tau)|$.

Note that $u_i(t) = z_i(t+1) - z_i(t)$ for any $i \in \mathcal{N}$, and

$$\begin{aligned} & |z_i(t+1) - z_i(t)| \\ = & \left| \frac{\sum_{k=1}^{t+1} x_i(k)}{t+1} - \frac{\sum_{k=1}^{t} x_i(k)}{t} \right| \\ = & \frac{\left| t \cdot x_i(t+1) - \sum_{k=1}^{t} x_i(k) \right|}{t(t+1)} \\ \leq & \frac{|x_i(t+1)|}{t+1} + \frac{\left| \sum_{k=1}^{t} x_i(k) \right|}{t(t+1)} \ . \end{aligned}$$

Using the Chernoff bound, for any $\delta > 0$, with a probability of at least $1 - 2e^{-2\delta t}$, we have

$$\left| \frac{\sum_{k=1}^{t} x_i(k)}{t} - \bar{x}_i \right| \leq \delta \ .$$

Then,

$$\begin{aligned} & P\left(\sup_{\tau \geq t} \left| \frac{\sum_{k=1}^{\tau} x_i(k)}{\tau} - \bar{x}_i \right| \leq \delta \right) \\ = & P\left(\forall \tau \geq t, \ \left| \frac{\sum_{k=1}^{\tau} x_i(k)}{\tau} - \bar{x}_i \right| \leq \delta \right) \\ = & 1 - P\left(\exists \ \tau \geq t, \ \left| \frac{\sum_{k=1}^{\tau} x_i(k)}{\tau} - \bar{x}_i \right| > \delta \right) \\ \geq & 1 - \sum_{\tau \geq t} P\left(\left| \frac{\sum_{k=1}^{\tau} x_i(k)}{\tau} - \bar{x}_i \right| > \delta \right) \\ \geq & 1 - \sum_{\tau \geq t} 2e^{-2\delta\tau} \\ = & 1 - \frac{2e^{-2\delta t}}{1 - e^{-2\delta}} \ , \end{aligned}$$

which implies that w.h.p.,

$$\sup_{\tau \geq t} \left| \frac{\sum_{k=1}^{\tau} x_i(k)}{\tau} \right| \leq \bar{x}_i + \delta \ ,$$

and thus

$$\sup_{\tau \geq t} |z_i(\tau+1) - z_i(\tau)| \leq \frac{K + \bar{x}_i + \delta}{t+1} \ ,$$

which is decreasing uniformly in the order of $O(\frac{1}{t})$. This completes the proof. ■

V. Conclusion

In this paper, we have studied a distributed belief averaging problem with sequential observations, where in contrast to the conventional distributed averaging problem setting each agent's local belief is not available immediately and can only be learned through her own observations. A distributed algorithm has been proposed for solving this problem, and its polynomial convergence rate to network-wide consensus has been established analytically for time-varying and undirected graphs. We have also shown that it is possible to extend parts of the model and results to the case of directed graphs to establish asymptotic convergence.

For future work, we will study convergence rate in the case of directed neighbor graphs, and analyze the effect of quantized communication [38].

References

[1] J. N. Tsitsiklis. *Problems in Decentralized Decision Making and Computation*. PhD thesis, Department of Electrical Engineering and Computer Science, MIT, 1984.

[2] D. P. Bertsekas and J. N. Tsitsiklis. *Parallel and Distributed Computation: Numerical Methods*. Prentice Hall, 1989.

[3] N. A. Lynch. *Distributed Algorithms*. Morgan Kaufmann Publishers, 1997.

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

[5] J. Cortés, S. Martínez, T. Karataş, and F. Bullo. Coverage control for mobile sensing networks. *IEEE Transactions on Robotics and Automation*, 20(2):243–255, 2004.

[6] J. Lin, A. S. Morse, and B. D. O. Anderson. The multi-agent rendezvous problem. Part 1: the synchronous case. *SIAM Journal on Control and Optimization*, 46(6):2096–2119, 2007.

[7] L. Hu and D. Evans. Localization for mobile sensor networks. In *Proceedings of the 10th Annual International Conference on Mobile Computing and Networking*, pages 45–57, 2004.

[8] L. Krick, M. E. Broucke, and B. A. Francis. Stabilisation of infinitesimally rigid formations of multi-robot networks. *International Journal of Control*, 82(3):423–439, 2009.

[9] L. Xiao, S. Boyd, and S. Lall. A scheme for robust distributed sensor fusion based on average consensus. In *Proceedings of the 4th International Conference on Information Processing in Sensor Networks*, pages 63–70, 2005.

[10] F. Bullo, J. Cortés, and S. Martínez. *Distributed Control of Robotic Networks*. Princeton University Press, 2009.

[11] J. Liu, N. Hassanpour, S. Tatikonda, and A. S. Morse. Dynamic threshold models of collective action in social networks. In *Proceedings of the 51st IEEE Conference on Decision and Control*, pages 3991–3996, 2012.

[12] F. Dörfler, M. Chertkov, and F. Bullo. Synchronization in complex oscillator networks and smart grids. *Proceedings of the National Academy of Sciences*, 110(6):2005–2010, 2013.

[13] R. Olfati-Saber and R. M. Murray. Consensus problems in networks of agents with switching topology and time-delays. *IEEE Transactions on Automatic Control*, 49(9):1520–1533, 2004.

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

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

[16] A. Kashyap, T. Başar, and R. Srikant. Quantized consensus. *Automatica*, 43(7):1192–1203, 2007.

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

[18] M. Cao, A. S. Morse, and B. D. O. Anderson. Reaching a consensus in a dynamically changing environment: a graphical approach. *SIAM Journal on Control and Optimization*, 47(2):575–600, 2008.

[19] J. Liu, A. S. Morse, B. D. O. Anderson, and C. Yu. Contractions for consensus processes. In *Proceedings of the 50th IEEE Conference on Decision and Control*, pages 1974–1979, 2011.

[20] J. M. Hendrickx, A. Olshevsky, and J. N. Tsitsiklis. Distributed anonymous discrete function computation. *IEEE Transactions on Automatic Control*, 56(10):2276–2289, 2011.

[21] B. Touri and A. Nedić. Product of random stochastic matrices. *IEEE Transactions on Automatic Control*, 59(2):437–448, 2014.

[22] J. Liu, A. S. Morse, A. Nedić, and T. Başar. Internal stability of linear consensus processes. In *Proceedings of the 53rd IEEE Conference on Decision and Control*, pages 922–927, 2014.

[23] L. Xiao and S. Boyd. Fast linear iterations for distributed averaging. *Systems and Control Letters*, 53(1):65–78, 2004.

[24] A. Nedić, A. Olshevsky, A. Ozdaglar, and J. N. Tsitsiklis. On distributed averaging algorithms and quantization effects. *IEEE Transactions on Automatic Control*, 54(11):2506–2517, 2009.

[25] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah. Randomized gossip algorithms. *IEEE Transactions on Information Theory*, 52(6):2508–2530, 2006.

[26] J. Liu, S. Mou, A. S. Morse, B. D. O. Anderson, and C. Yu. Deterministic gossiping. *Proceedings of the IEEE*, 99(9):1505–1524, 2011.

[27] D. Kempe, A. Dobra, and J. Gehrke. Gossip-based computation of aggregate information. In *Proceedings of the 44th Annual IEEE Symposium on Foundations of Computer Science*, pages 482–491, 2003.

[28] F. Bénézit, V. Blondel, P. Thiran, J. N. Tsitsiklis, and M. Vetterli. Weighted gossip: distributed averaging using non-doubly stochastic matrices. In *Proceedings of the 2010 IEEE International Symposium on Information Theory*, pages 1753–1757, 2010.

[29] A. D. Domínguez-García, S. T. Cady, and C. N. Hadjicostis. Decentralized optimal dispatch of distributed energy resources. In *Proceedings of the 51st IEEE Conference on Decision and Control*, pages 3688–3693, 2012.

[30] J. Liu and A. S. Morse. Asynchronous distributed averaging using double linear iterations. In *Proceedings of the 2012 American Control Conference*, pages 6620–6625, 2012.

[31] M. Huang and J. H. Manton. Coordination and consensus of networked agents with noisy measurements: stochastic algorithms and asymptotic behavior. *SIAM Journal on Control and Optimization*, 48(1):134–161, 2009.

[32] S. Kar, J. M. Moura, and K. Ramanan. Distributed parameter estimation in sensor networks: nonlinear observation models and imperfect communication. *IEEE Transactions on Information Theory*, 58(6):3575–3605, 2012.

[33] A. Nedić and A. Olshevsky. Distributed optimization over time-varying directed graphs. *IEEE Transactions on Automatic Control*, 60(3):601–615, 2015.

[34] J. Liu, A. S. Morse, B. D. O. Anderson, and C. Yu. Contractions for consensus processes. In *Proceedings of the 50th IEEE Conference on Decision and Control*, pages 1974–1979, 2011.

[35] E. Seneta. *Non-negative Matrices and Markov Chains*. Sringer, 2006.

[36] W. J. Rugh. *Linear Systems Theory*. Prentice-Hall, 1996.

[37] J. Liu, T. Başar, and A. Nedić. Input-output stability of linear consensus processes. In *Proceedings of the 55th IEEE Conference on Decision and Control*, pages 6978–6983, 2016.

[38] M. El Chamie, J. Liu, and T. Başar. Design and analysis of distributed averaging with quantized communication. *IEEE Transactions on Automatic Control*, 61(12):3870–3884, 2016.