Consensus Formation in a Switched Markovian Dynamical System

Kevin Topley, Vikram Krishnamurthy, George Yin

Abstract- We address the problem of distributively obtaining average-consensus among a connected network of sensors that each respectively track, by linear stochastic approximation, the stationary distribution of an ergodic Markov chain with slowly switching regimes. A hyper-parameter modeled as a Markov process on a slow time-scale modulates the regime of each observed Markov chain, thus at any given time the hyper-parameter determines what stationary distribution will be estimated by each sensor. If the Markov chains share a common stationary distribution conditional on the regime, it is shown the sequence of sensor state-values weakly-converge to an average-consensus under the distributed linear consensusfilter for all network communication graphs. Conversely, if the Markov chains have unique stationary distributions in each regime, then the average-consensus can be achieved only when sensors communicate state-values at a frequency that is on the same time-scale as the frequency at which they observe the fast Markov chain. In this scenario, unlike a static consensus filter, the state-value communication graph need not be connected for an average-consensus to be reached, however this is true only when the communication graph of observation data satisfies a specific connectivity condition. Simulations illustrate our conclusions and observation model.

1. INTRODUCTION

We consider consensus formation in an ad-hoc network of coupled sensors where each sensor individually tracks by an adaptive stochastic (SA) approximation the stationary distribution of a set of Markov chains with time-varying regime. By sharing information the sensors supplement each others incomplete knowledge of the entire set of observed parameters, our focus is on how distributive linear averaging can ensure all sensors eventually reach a common empirically-based estimate of the average of all observed stationary distributions as they vary in time, we refer to the desired estimate as an average-consensus.

Initially we will assume the sensor estimates are updated according to a distributed linear consensus filter similar to that proposed in [7], although in contrast to [7] we model the observed parameter as a finite-state Markov chain with piece-wise constant stationary distributions, also known as slowly switching regimes. We show the model of [7] implies the frequency of communication occurs on a *slower* timescale than that of sensor observation, and in this case the average-consensus is generally unattainable. As an alternative we consider the same algorithm when the frequency

V. Krishnamurthy is with the Department of Electrical Engineering, University of British Columbia, Vancouver, BC V6T 1Z4, Canada (vikramk@ece.ubc.ca).

G. Yin is with the Department of Mathematics, Wayne State University, Detroit, MI 48202 (gyin@math.wayne.edu).

of communication occurs on the *same* time-scale as sensor observations, and in this case we derive connectivity conditions on the network communication graph required to attain average-consensus.

Analogous to [7], each sensor i = 1, ..., n observes the state $X^i \in \mathbf{R}^{S \times 1}$ of a respective S-state Markov chain X^i . By a linear SA adaptive algorithm with fixed step-size $0 < \mu \ll 1$ each sensor forms their local estimate of the stationary distribution π^i associated with X^i ,

$$s_{k+1}^i = s_k^i - \mu(X_k^i - s_k^i) , \ s_0^i = X_0^i , \ k = 0, 1, \dots$$
 (1.1)

The sensor estimates $\{s^1, \ldots, s^n\}$ will, as the number of observations approaches infinity, then take the form of an empirical distribution, thus the sensors maintain "type" data [2].

Distributed communication of type data to a single fusion center is considered in, for instance, [11]. Due to the presumption of a fusion base node, the algorithms for processing communicated data proposed in works such as [11] are in general quite different from the distributed algorithms presented in [6], [12], [3]. The latter works assume no base node, rather it is assumed individual sensors are connected by a limited number and arbitrarily placed set of coupling links. This is identical to the ad-hoc network design that we consider. For consensus it is required that, in some sense, the union of all coupling links implies a path between any two sensors, but we emphasize there is no fusion center assumed, as this might imply consensus trivially.

A. Consensus Algorithm

As an averaging algorithm we consider the one-hop distributed linear consensus-algorithm proposed in [7], that is at every iteration $k \in \{0, 1, 2, ...,\}$ each sensor *i* computes the state-value $s_{k+1}^i \in \mathbf{R}^{S \times 1}$ by taking an element-wise weighted average of their previous state-value s_k^i , their current observed value X_k^i , and the state and observed values received from all sensors that are adjacent to sensor *i*. In combination with (1.1) this algorithm has been expressed in [7] as a discrete-time iteration with constant step-size $0 < \mu \ll 1$,

$$s_{k+1} = s_k - \mu(\mathcal{L}^v + \mathcal{D}^o)s_k + \mu \mathcal{W}^o X_k \tag{1.2}$$

where we define the network communication weight matrices, $\mathcal{D}^o = diag(\mathcal{W}^o 1\!\!1)$, $\mathcal{L}^v = \mathcal{D}^v - \mathcal{W}^v$, and $\mathcal{D}^v = diag(\mathcal{W}^v 1\!\!1)$. The elements of the matrices $\{\mathcal{W}^v, \mathcal{W}^o\}$ are the linear weights attributed to each transmission of a statevalue (s_k^i) or observation values (X_k^i) , respectively.

Each $\mathcal{W}_{ij}^v = w_{ij}^v I$, i.e., a scalar multiple of the $S \times S$ identity matrix such that w_{ij}^v are non-zero only when a

communication link exists between the respective sensors iand j. The collection of all communication links comprises an edge set \mathcal{E} of the network communication graph $\mathcal{G} = \{n, \mathcal{E}, \mathcal{W}^v\}$. We initially assume this graph is connected, undirected, and fixed, thus (1.3) possesses the following constraints,

1) $\mathcal{W}_{ij}^v = 0 \Leftrightarrow (i,j) \notin \mathcal{E},$

- 2) $(i, j) \in \mathcal{E} \Leftrightarrow (j, i) \in \mathcal{E},$
- 3) $\forall i, \exists j \text{ such that } (i,j) \in \mathcal{E}$

The same constraints apply to W^o , although we note the diagonals of W^o are assumed non-zero, whereas by (1.2) the diagonal elements of W^v are irrelevant and are set equal to zero for convenience.

B. Background

Although much research has explored the properties of the "static" consensus algorithm

$$s_{k+1}^{i} = s_{k}^{i} + \mu \sum_{j=1}^{n} \mathcal{W}_{ij}^{v}(s_{k}^{j} - s_{k}^{i}), \quad i = 1, \dots, n,$$
 (1.3)

under time-varying or stochastic graphs (for instance see [4], [10], [5]) significantly less research has considered distributed averaging across nodes that *track* an external parameter while simultaneously seeking consensus. An essential example of such research is [7], which considers the continuous-time limit of (1.3) when each node shares and is also linearly updated by a continuous signal $\gamma_t \in \mathbf{R}$ of bounded rate, observed in i.i.d. Gaussian noise. All nodes are then shown to remain within a bounded distance ε of the observed parameter at all times. This framework has subsequently been expanded in a number of respects, such as with implementation of local Kalman-filtering, time-varying communication graphs, as well as dynamic tracking algorithms for mobile sensors, to name a few [9], [8].

The preliminary difference between [7] and the current model is that we now replace the communally observed continuous signal γ by a set of S-state Markov chains $\{X^1, \ldots, X^n\}$, each of which is privately observed by a corresponding sensor s^i , $i = 1, \ldots, n$. Secondly, we consider the Markov chains $\{X^i\}$ are dependent on a parameter θ , where the sequence $\{\theta_k\}$ evolves according to a Markov chain whose transitions take place infrequently and with state space represented by $\mathcal{M}_{\theta} = \{\theta^1, \ldots, \theta^m\}$.

As the essential results of this paper, we show that provided the transition probability matrix of θ_k is "near" identity, or specifically that θ evolves on a time-scale of the same order $O(\mu)$ as that of the sensor consensus-tracking algorithm (1.2), then under a suitably weighted communication graph the continuous-time linear interpolation of the state-values s_k^i will converge weakly to a stochastically switching consensus of the average observed stationary distribution $\bar{\pi}(\theta) = \frac{1}{n} \sum_{i=1}^{n} \pi^i(\theta)$. In addition, the sequence of tracking errors between each sensor and the average-consensus, when properly scaled, is shown to converge weakly to the solution of a switching diffusion. These results constitute an altogether distinct set of network consensus dynamics than yet reported in the literature.

The paper is organized as follows: in $\S2$ we present the main convergence theorems, \$3 investigates some ramifications of these. Simulations are provided in Section \$4 to illustrate our results, and Section \$5 provides a summary.

2. DYNAMIC CONSENSUS FORMATION

Let X_k^i be an S-state Markov chain with a state space $\{e_1, \ldots, e_S\}$, where $\{e_i\}$ are orthogonal $S \times 1$ standard unit vectors. Each X_k^i is θ dependent for $\theta \in \mathcal{M}_{\theta} = \{\theta^1, \ldots, \theta^m\}$, that is the transition matrix of X_k^i conditioned on θ is given by $A^i(\theta) = (a_{ij}^i(\theta))$, where

$$a_{lj}^{i}(\theta) = P(X_{k+1}^{i} = e_{j}|X_{k}^{i} = e_{l}, \theta_{k} = \theta).$$
 (2.4)

In the following section we first assume the weights $\{W^o, W^v\}$ are fixed such that $H = \mathcal{L}^v + \mathcal{D}^o$ has only eigenvalues with positive real parts, denoted $H \succ_{re} 0$, thus providing bounded stability of the consensus algorithm (1.2) in the limit μ approaches zero [1]. Under this condition we describe how the state-values s_k of (1.2) will then weakly-converge as μ vanishes.

A. Asymptotic Properties of Stochastic Approximation Algorithms

To proceed, we pose the conditions below.

- (A) The following conditions hold.
 - For each θ ∈ M_θ, the transition matrix Aⁱ(θ) is irreducible and aperiodic with stationary measure πⁱ(θ).
 - 2) Parameterize the transition probability matrix of θ as

$$P^{\varepsilon} = I + \varepsilon Q_{1}$$

where ε is a small parameter satisfying $0 < \varepsilon \ll 1$ and Q is the generator of a continuous-time finitestate Markov chain.

The process θ_k is slow in the sense ε = O(μ). For simplicity, we take ε = μ henceforth.

(B)
$$H \succ_{re} 0$$
 where $H = (\mathcal{L}^v + \mathcal{D}^o)$.

Although the proof is omitted for brevity (see [13]), we shall show that it is feasible for the network to track changes of a linear combination of $\{\pi^1, \ldots, \pi^n\}$. In particular, as the step-size goes to zero, each sensor state value converges weakly to the solution of a stochastically switching ODE and possesses a scaled tracking error with a switching diffusion limit.

Theorem 2.1: Assume conditions in (A) - (B) hold and $\varepsilon = \mu$. Define the interpolated sequences of iterates

$$s_t^{\mu} = s_k, \ \ \theta_t^{\mu} = \theta_k \ \ \text{for} \ \ t \in [k\mu, (k+1)\mu).$$

Then as $\mu \to 0$, $(s^{\mu}(\cdot), \theta^{\mu}(\cdot))$ converge weakly to $(s(\cdot), \theta(\cdot))$ such that $\theta(\cdot)$ is a continuous-time Markov chain with generator Q and $s(\cdot)$ satisfies

$$\frac{ds_t}{dt} = -Hs_t + \mathcal{W}^o \pi(\theta_t), \ t \ge 0, \tag{2.5}$$

where

$$\pi(\theta) = [\pi^1(\theta), \dots, \pi^n(\theta)]'$$

We next study the associated tracking errors. Define

$$v_k = \frac{s_k - G\mathbf{E}(\pi(\theta_k))}{\sqrt{\mu}},\tag{2.6}$$

where $G = H^{-1} W^o$. Define the interpolated sequence of scaled tracking errors by

$$v_t^{\mu} = v_k, \ t \in [k\mu, (k+1)\mu).$$
 (2.7)

Theorem 2.2: Under the same conditions as Theorem 2.1, $v^{\mu}(\cdot)$ converges weakly to $v(\cdot)$, which is a solution of the switching diffusion

$$dv_t = -Hv_t dt + \mathcal{W}^o \Sigma^{1/2}(\theta_t) dw, \qquad (2.8)$$

where $w(\cdot)$ is a standard Brownian motion and for a fixed θ , $\Sigma(\theta)$ is a covariance matrix.

Theorem 2.1 and Theorem 2.2 both present a convergence result for small μ and large k such that μk remains bounded, we refer to this time-scale as O(1). From the theorems it is clear how the asymptotic sensor dynamics depend on both the variation in θ and the choice of edge weights $\{W^o, W^v\}$ of the communication graph. In the following section we focus on Theorem 2.1 regarding the sufficient or necessary weights for all sensors to reach the average-consensus $\bar{\pi}(\theta)$.

3. RAMIFICATIONS

It is clear by (2.5) that each sensor in the network weakly-converges to an equilibrium, or steady-state, that is modulated by θ and can be represented as the product $\Lambda \pi(\theta)$, where

$$\Lambda = H^{-1} \mathcal{W}^o = (\mathcal{L}^v + \mathcal{D}^o)^{-1} \mathcal{W}^o.$$
(3.9)

If each observed Markov chain has an identical stationary distribution $\pi^*(\theta)$ conditioned on θ , then $\bar{\pi}(\theta) = \pi^*(\theta)$ and the above theorems imply weak-convergence to the *average*consensus $\bar{\pi}(\theta)$ by all sensor state-values, that is, $s_k^i \to \bar{\pi}(\theta)$ for i = 1, ..., n. This is in fact true regardless of the network communication graph \mathcal{G} , provided (B) holds.

Corollary 3.1: If $\pi^{i}(\theta) = \pi^{*}(\theta)$ for all i and $\theta \in \mathcal{M}$, then provided (A) - (B) and $\epsilon = \mu$, an average-consensus is reached for all communication graphs \mathcal{G} , including the null graph $\mathcal{G} = \{n, 0, 0\}$.

Proof. By the first assumption $\pi(\theta) = [\pi^*(\theta), \ldots, \pi^*(\theta)]'$ for some $\pi^*(\theta) \in \mathbf{R}^{S \times 1}$, thus $\Lambda \pi = \pi$ for all $\Lambda \mathbb{1} = \mathbb{1}$. We now show that all Λ satisfying (3.9) are row-stochastic, hence the asymptotic equilibrium will be the consensus π^* . Rearranging (3.9) yields the necessary condition

$$\mathcal{W}^o = (\mathcal{L}^v + D^o)\Lambda. \tag{3.10}$$

If $\Lambda 1 = 1$ then we have

$$\mathcal{W}^{o}\mathbb{1} = \mathcal{D}^{v}\mathbb{1} - \mathcal{W}^{v}\mathbb{1} + D^{o}\mathbb{1} , \qquad (3.11)$$

which is true by the definitions $\mathcal{D}^v = \text{diag}(\mathcal{W}^v 1)$ and $\mathcal{D}^o = \text{diag}(\mathcal{W}^o 1)$. If $\Lambda 1 \neq 1$ then, as the inverse of a matrix is unique, (3.9) could not hold. \Box

If the observed Markov chains do not have identical stationary distributions conditioned on θ , that is $\pi^i(\theta) \neq$

 $\pi^{j}(\theta)$ for all $i \neq j$ and $\theta \in \mathcal{M}$, then average-consensus is achieved only if $\Lambda = \frac{1}{n} \mathbb{1} \mathbb{1}'$. This is a much stronger constraint than $\Lambda \mathbb{1} = \mathbb{1}$, and in fact it can only be achieved in approximation if the edge set \mathcal{E} is not complete, that is there exists an ordered pair (i, j) such that $(i, j) \notin \mathcal{E}$, we refer to such a graph as "unsaturated".

WeC12.5

Corollary 3.2: Under the assumption each sensor will asymptotically observe a unique stationary distribution, consensus cannot be achieved by a linear consensus-filter if G is unsaturated.

Proof. Since any consensus requires Λ have identical rows, left-multiplication of Λ by $(\mathcal{L}^v + \mathcal{D}^o)$ results in $\mathcal{D}^o \Lambda$ by definition of \mathcal{L}^v . Since \mathcal{D}^o is diagonal, by (3.9) the matrix \mathcal{W}^o will have every element non-zero, thus implying a saturated communication graph. \Box

Contrary to the above result, by increasing the frequency of communication, consensus *can* be achieved. We inspect the consequences of this scenario in the following subsection.

A. Communication Step-Size

Although the observation update weights W^o must be scaled by the step-size $0 < \mu \ll 1$ in order for weakconvergence of sensor state-values, there is no apparent reason the update weights W^v also need to be scaled by μ . To see this we note,

- 1) communication of state-values consists implicitly of a communication of past observed values as scaled by $\mu(1-\mu)^l$, where *l* denotes the *l*th most recent observation since last communication.
- 2) if $(\mathcal{L}^v + \mathcal{D}^o) \succ_{re} 0$ then $(m\mathcal{L}^v + \mathcal{D}^o) \succ_{re} 0$ for all m > 0 only under the conditions $\{\mathcal{L}^v \succeq_{re} 0, \mathcal{D}^o \succ_{re} 0\}$. We assume the latter conditions.

Our analysis until now has assumed W^v is scaled by μ , we now explain how this can be seen to imply a distributed linear consensus-filter for which the frequency of communication is much smaller than the frequency of observation.

1) Slow Communication: In $\S 2 - A$ we assert that as the step-size μ approaches zero if $\epsilon = \mu$ the continuous-time piece-wise constant interpolation of the sequence $\{s_k\}$,

$$s_t = s_k \quad t \in [k\mu, (k+1)\mu),$$
 (3.12)

will converge weakly to solutions of the switching ODE (2.5) where Q is the generator of a continuous-time Markov chain, which we emphasize is continuous on the time-scale O(1).

Taking μ to zero when $\epsilon = \mu$ implies the transition matrix P^{ϵ} is "near" identity on the time-scale $O(\mu)$. By (1.1) this implies the observed parameter θ evolves on the same time-scale as the adaptive SA algorithm, or equivalently the time-scale of sensor observation.

We now inspect the time-scale O(1), which is the timescale of reference for (2.5). On this time-scale the matrix P^{ϵ} is not near identity, and the parameter θ is seen to moderately vary. We once again assume the observation model (1.1), but now we explicitly model sensor communication to occur on the slow time-scale O(1) and will subsequently show that this is equivalent to (1.2). First consider the *n* sensors are independently updated as (1.1) by *n* Markov chains X^i . By (2.5) each sensor would then weakly-converge to the solutions of

$$\frac{ds_t}{dt} = -s_t + \pi(\theta_t), \ t \ge 0, \tag{3.13}$$

where t is measured on the time-scale O(1). Suppose that at every time $t^* \in T = \{t : t \in \mathbb{N}\}$, each sensor *i* communicates its observation value $X_{t^*}^i$, along with its current state-value $s_{t^*}^i$, through all out-going communication links; the receiving sensors then compute a weighted linear average of their own state-value and the elements of data $(s_{t^*}^i, X_{t^*}^i)$ they just received.

The continuous-time sensor dynamics are then expressed,

$$s_t = \begin{cases} e^{-At}(s_{\bar{t}^*} - A^{-1}\pi(\theta_t)) + A^{-1}\pi(\theta_t) , & \text{if } t \notin \mathbb{N} \\ (I - \mathcal{L} - \mathcal{D}_{-d}^o)s_t + \mathcal{W}_{-d}^o X_t , & \text{if } t \in \mathbb{N} \end{cases}$$

$$(3.14)$$

where $\bar{t}^* = max_{t^*} \{t^* < t\}, \mathcal{W}_{-d}^o$ equals \mathcal{W}^o but with zeros on the diagonal, and $A = \mathcal{D}^o - \mathcal{D}_{-d}^o$ is the diagonal matrix of each sensors individual observation update weight. We also denote $\mathcal{L}^v = \mathcal{L}$ for remainder of this section.

Defining $s_0 = X_0$ and $t_-^* = \lim t \nearrow t^*$, we then have for arbitrary t_-^* the expected sensor state-values,

$$\mathbf{E}(s_{t_{-}^{*}}) = W^{t^{*}-1}e^{-At^{*}}s_{0} + \sum_{l=0}^{t^{*}-1}W^{l}e^{-Al}$$

$$(I + e^{-A}(\mathcal{W}^{o} - I))\pi(\theta_{t^{*}})$$
(3.15)

where we define $W = I - \mathcal{L} + \mathcal{D}^o$ and note that, without prior knowledge, the expected value of X_t^i is $\pi^i(\theta_t)$.

For large t_{-}^{*} it is clear that under $\S 2 - A$ (B) the coefficient of X_0 vanishes whereas the coefficient of π converges to

$$(I - We^{-A})^{-1}(I + e^{-A}(\mathcal{W}^{o}_{-d} - I)) = \frac{e^{A} - I + \mathcal{W}^{o}_{-d}}{e^{A} - I + \mathcal{L} + \mathcal{D}^{o}_{-d}}$$
(3.16)

which we find is equal to the equilibrium (3.9) under the continuous model (1.2) when either A = ln(2) or equivalently when the rate of continuous-time is scaled $\Delta t_{disc} = ln(2)\Delta t_{cont}$.

We note that in general if sensor *i* has an individual observation update weight \mathcal{W}_{ii}^{o} in the continuous model (1.2), then the corresponding update weights of the sensors in the discrete model (3.14) is given by $A_{ii} = ln(\mathcal{W}_{ii}^{o} + 1)$. If $A_{ii} \neq A_{jj}$ then the sensors (i, j) might operate on time-scales that are uniquely scaled and thus would possess asynchronous times of communication to yield (3.14).

Specifically, we find that under the discrete model there exist an infinite subsequence of sensor state-values that, with the correct time factors, will converge in expectation to the sensor state-values under the continuous model. To illustrate this fact we simulate a distributed ad hoc network of 10 sensors operating under both (1.2) and (3.14). For clarity in Fig.1 we plot the *linear sum* of each sensors state-value, that is we plot the the scalar $11's^i$ rather than the $S \times 1$ vector s^i for each sensor *i*. A jump in θ occurs mid-way in our simulation, as signified by the vertical line.



Fig. 1. Comparison of Discrete (3.14) and Continuous (1.2) Models assuming slow communication. As t increases from the most recent transition time of θ , both models have identical equilibriums at the times t_{-}^{*} .

2) Consensus Properties under Slow Communication: It is evident from Fig.1 and proven by Corollary 3.1 that in either the discrete or continuous model, consensus cannot be reached when $A \neq 0$. However, when A = 0 no tracking occurs, which implies that linearly tracking a set of distinct parameters by the SA (1.1) is fundamentally opposed to distributed linear consensus formation. Since consensus can only be partially achieved in this setting, we refer to an arbitrary graph has having an inherent "consensus ability"; although we leave this term formally undefined for the moment, it is intuitive that graphs with many strategically placed communication links should have better consensus ability than graphs with only a few randomly placed links, and indeed it is only the saturated graph that can actually attain consensus in this framework.

We also note that the discrete model requires unscaled communication of observation data, thus the sensor iterates are stochastic in the sense that at each time t_{-}^{*} it is only in *expectation* that (3.14) yields the same equilibrium as (1.2). At any given time t_{-}^{*} the sensors under (3.14) will be sharing realizations of $\{X^{1}, \ldots, X^{n}\}$, not their stationary measures $\pi^{i}(\theta)$. Thus the iterates $s_{t_{-}^{*}}$ under the discrete model do not actually approach $\bar{\pi}(\theta)$, rather their average $\sum_{l=min(l^{*}>l_{\theta})}^{m} \frac{1}{m} s_{l_{-}^{*}}$ will approach (3.16) with probability 1 (w.p.1) as *m* increases, where l_{θ} denotes the most recent transition time of θ_{t} .

In Fig.1 we parameterized the variability of X as relatively small such that the stochastic property we are discussing cannot be noticed. Fig.2 displays a more accurate illustration.



Fig. 2. Comparison of Discrete (3.14) and Continuous (1.2) Models assuming slow communication but when X has large variation. As m increases the average $\sum_{l=min(l^*>l_{\theta})}^{m} \frac{1}{m} s_{l^*}$ of discrete iterates and s_{t^*} of (1.2) have identical equilibriums w.p.1.

The continuous algorithm results in no visible stochastic element, although both models possess the scaled diffusion (2.8). The stochastic element of (3.14) is the result of updating the sensors by the parameter X unscaled by μ . We refer to this stochastic element as "stochastic variation" and note that it is the essential difference between the two models (1.2) (3.14), since otherwise their equivalence depends only on scaling the rate at which the sensors of either algorithm measure time.

To avoid the undesirable stochastic variation present in (3.14) we can set $\mathcal{W}_{-d}^o = 0$, then an unscaled X is no longer used to update the sensor state-values and thus we need not take their expectation. When $\mathcal{W}_{-d}^o = 0$ the equilibrium (3.16) becomes

$$\frac{e^A - I}{e^A - I + \mathcal{L}} {.} {(3.17)}$$

The extent to which this reduces the sensor ability to form consensus is considered negligible, and not discussed here.

3) Fast Communication: We now inspect the sensor equilibrium as the frequency of communication in (3.14) increases. For some $m \in \mathbb{N}$ we take $t^* \in T = \{t : mt \in \mathbb{N}\}$ and consider again (3.14). In this case the equilibrium (3.16) becomes

$$\frac{e^{A/m} - I + \mathcal{W}^{o}_{-d}}{e^{A/m} - I + \mathcal{L} + \mathcal{D}^{o}_{-d}} .$$
(3.18)

If the communication weights W^o are not scaled, then an expectation must be taken to attain equivalence between the discrete and continuous models. As m increases the stochastic element described above will affect the sensors at more frequent time points and thus perpetuate a random dynamic behavior of the sensor iterates, pictured in Fig.3 below.



Fig. 3. A connected network under fast communication with unscaled communication weights of observation data. The consensus ability is not significantly improved as compared to the slow communication model pictured in Fig.3 – A.1-3-A.2. However, the stochastic variation present in Fig.3 – A.2) now perturbs the sensor iterates with greater frequency as *m* increases.

The sensor equilibrium remains given by nearly the same expression as it is under the slow communication models, that is as m approaches infinity (3.18) becomes

$$\lim_{m \to \infty} \frac{A/m + \mathcal{W}_{-d}^{\circ}}{A/m + \mathcal{L} + \mathcal{D}_{-d}^{\circ}} .$$
(3.19)

Although the above limit can be seen to aid in consensus formation, we will not explore this here. We also assert without proof that Cor.3.2 applies to (3.19) just as it did to (3.16), thus consensus can again be attained only in approximation, similar to the slow communication algorithms (1.2) (3.14).

Due to both the unwanted stochastic variation and the network inability to form consensus, we now scale the communication weights W^o by a factor of order $O(m^{-1})$. In this case, as m approaches infinity the sensor equilibrium (3.18) becomes

$$\lim_{m \to \infty} (m\mathcal{L} + \mathcal{D}^o)^{-1} \mathcal{W}^o .$$
 (3.20)

In contrast to Cor.3.2, the following lemma shows (3.20) implies that weak-convergence to the average-consensus may be attained by sensors operating under (3.14) when $t^* \in T = \{t : mt \in \mathbb{N}\}$ as the frequency of communication m approaches infinity.

Lemma 3.3: For any connected network graph with Laplacian \mathcal{L} , there exist left and right eigenvectors ω_r and ω_ℓ satisfying

$$\mathcal{L}\omega_r = 0 , \ \omega'_\ell \mathcal{L} = 0 , \ \omega'_\ell \omega_r = 1, \qquad (3.21)$$

such that

$$\lim_{m \to \infty} (m\mathcal{L} + \mathcal{D}^o)^{-1} \mathcal{W}^o = (\omega_\ell \mathcal{D}^o \omega_r)^{-1} \omega_r \omega'_\ell \mathcal{W}^o.$$
(3.22)

Proof. For a connected network the matrix \mathcal{L} is guaranteed by its definition to have eigenvectors $\{\omega_r, \omega_\ell\}$ satisfying (3.21), furthermore $\omega_r = c \mathbb{1}$ for any $c \in \mathbb{R}$.

Denoting the eigendecomposition of $\mathcal{L}(\mathcal{D}^o)^{-1}$ as UJU^{-1} we note the argument in (3.22) may be rearranged as follows,

$$(m\mathcal{L} + \mathcal{D}^o)^{-1}\mathcal{W}^o = (\mathcal{D}^o)^{-1}U(I + mJ)U^{-1}$$
. (3.23)

Since $\mathcal{L}(\mathcal{D}^o)^{-1}$ will have the eigenvalue pair $\{\mathcal{D}^o\omega_r, \omega_\ell\}$ satisfying the same conditions as (3.21) in regard to \mathcal{L} , the rearranged eigendecomposition $U^{-1}\mathcal{L}(\mathcal{D}^o)^{-1} = JU^{-1}$ implies that if the i^{th} row of U^{-1} equals ω'_ℓ then $J_{jj} = 0$ only when j = i. This in turn implies that only the i^{th} diagonal of the matrix (I + mJ) will remain bounded in the limit $m \to \infty$, and furthermore, since this element is unity, the inverse $(I + mJ)^{-1}$ will have this precise term as its only non-zero element.

By the same reasoning, rearranging the eigendecomposition as $\mathcal{L}(\mathcal{D}^o)^{-1}U = UJ$ implies the i^{th} column of U is $\mathcal{D}^o\omega_r$. Since $U^{-1}U = I$ we have $\omega'_{\ell}\mathcal{D}^o\omega_r = 1$ and thus (3.22) is obtained by multiplying \mathcal{W}^o on both sides of (3.23) and taking the limit. \Box

It is clear then under fast communication the sensors may achieve consensus if and only if the weight matrices $\{W^v, W^o\}$ are such that

$$(\omega_{\ell} \mathcal{D}^{o} \mathbf{1})^{-1} \mathbf{1} \omega_{\ell}^{\prime} \mathcal{W}^{o} = \frac{1}{n} \mathbf{1} \mathbf{1}^{\prime} . \qquad (3.24)$$

Due to this, we note that when communicating on the fast time-scale $O(\mu)$, sensor communication of observation data is not required to attain the average-consensus, provided \mathcal{L} is connected. To see this we may take the Laplacian \mathcal{L}_c of an arbitrarily connected graph and set $\mathcal{W}_{ii}^o = \frac{1}{n} (\omega_\ell)_i$ where ω_ℓ satisfies (3.21).

We also note that \mathcal{L}_c need not be balanced, which by [6] is in contradiction to the average-consensus requirements of the static algorithm (1.3). As shown in [6], only for a connected graph will the static algorithm yield an asymptotic consensus, and only if \mathcal{L} is balanced, $\mathbb{1}'\mathcal{L} = \mathcal{L}\mathbb{1}$, will (1.3) yield the average-consensus.

On the other hand, if we assume communication of scaled observation data, then by allowing two distinct communication graph edge sets \mathcal{E}^v and \mathcal{E}^o associated with \mathcal{W}^v and \mathcal{W}^o respectively, we can extend Lem.3.3 provided an extra connectivity condition holds in regard to the sensor communication graphs $\{\mathcal{E}^v, \mathcal{E}^o\}$. Specifically we can ensure an average-consensus even when \mathcal{L} is not connected.

The extra connectivity condition may be stated as follows,

• $\{\mathcal{E}^v, \mathcal{E}^o\}$ "jointly connect" a set of nodes if for any two nodes (i, j) there either exists a directed path $\{e_1^{ij}, \ldots, e_{p(i,j)}^{ij}\}$ of length $1 \leq p(i, j) \leq n$, that is of the form

$$e_1^{ij} \in \mathcal{E}^o \cup \mathcal{E}^v \ , \ e_l^{ij} \in \mathcal{E}^v \ \forall \ l \ge 2, \tag{3.25}$$

This definition does not require the Laplacian \mathcal{L} to have a connected edge set, thus under fast communication averageconsensus may be attained even when breaching the consensus requirement for the static algorithm (1.3). The rationale that necessitates (3.25) is intuitive; if any sensor does not have an incoming state-value communication path from a subset of the rest of the network, then that sensor must have an incoming observation communication link from every sensor in that subset because otherwise it receives no data from these sensors, either directly or indirectly, and so it cannot possibly achieve the average-consensus. A full proof is omitted as it is trivial and requires only tedious notation, in lieu we demonstrate this result in the following section.

To summarize then, previously on the slow time-scale we considered $W_{-d}^o = 0$ which implies zero sensor communication of observation data, this was deemed not to have a significant effect on the resulting network ability to form consensus. We now see that on the fast time-scale it does have a significant effect in terms of the network connectivity conditions required for average-consensus.

4. CONSENSUS UNDER JOINT CONNECTIVITY

We aim here to illustrate by simulation the last result of $\S3-A.3$, specifically that average-consensus can be achieved when sensors communicate state-value information through an unconnected edge set \mathcal{E}^v , but communicate observation data through an edge set \mathcal{E}^o such that $\{\mathcal{E}^v, \mathcal{E}^o\}$ jointly connect a group of n sensors.

For clarity we assume a network of only 4 sensors connected by $\{\mathcal{E}^v, \mathcal{E}^o\}$ as pictured in Fig.4. We note that although \mathcal{E}^v is not connected, both graphs $\{\mathcal{E}^v, \mathcal{E}^o\}$ together jointly connect every sensor.

Using unit weights (3.24) will hold, and consensus is formed, illustrated below in Fig.5.

5. CONCLUSIONS

We have shown conditions on the network communication graph edge set and weights that together ensure that under a one-hop distributed linear consensus algorithm a network



Fig. 4. A network that is jointy connected by $\{\mathcal{E}^v, \mathcal{E}^o\}$. The dotted lines indicate observation communication links (\mathcal{E}^o) and solid lines indicate state-value communication links (\mathcal{E}^v) .



Fig. 5. Network average-consensus under fast communication with unit weights for each link in Fig.4. The iterates of the continuous and discrete model are identical in this case, provided the communication weights of observation values are scaled by a factor $O(\mu)$.

average-consensus will be the asymptotic weak limit of sensors that track a set of slowly switching Markov parameters by linear stochastic approximation.

REFERENCES

- C. Chen, Linear System Theory and Design, Oxford Press, pp.234-235, 1997.
- [2] I. Csiszar. The Method of Types, *IEEE Transactions on Information Theory*, vol.44, No.6, 1998, pp.2505-2523.
- [3] I. Schizas, G. Giannakis, S. Roumeliotis, and A. Ribeiro, Consensus in Ad Hoc WSNs with Noisy Links, *IEEE Trans. on Signal Proc.*, vol. 56(4), 2008, pp.1650-1666.
- [4] Y. Hatano and M. Mesbahi, Agreement Over Random Networks, *IEEE Trans. Automat. Control*, vol.50(11), 2005, pp.1867-1872.
- [5] L. Moreau, Stability of Multiagent Systems with Time-dependent Communication Links, *IEEE Trans. Automat. Control*, vol. 50(2), 2005, pp.169-182.
- [6] R. Murray and R. Olfati-Saber, Consensus Problems in Networks of Agents with Switching Topology and Time-Delays, *IEEE Trans. Automat. Control*, vol.49(9), 2004, pp.1520–1532.
 [7] R. Olfati-Saber and J.F. Shamma, "Consensus Filters for Sensor
- [7] R. Olfati-Saber and J.F. Shamma, "Consensus Filters for Sensor Networks and Distributed Sensor Fusion", Proc. of 44th IEEE Conf. on Dec. and Cont., Seville, Spain, 2005.
- [8] R. Olfati-Saber, "Distributed Kalman Filter with Embedded Consensus Filters", Thayer School of Engineering Hanover.
- [9] D. Spanos, R. Olfati-Saber, and R. Murray, "Dynamic Consensus on Mobile Networks", California Inst. of Tech. Pasadena, California.
- [10] M. Porfiri and D. Stilwell, Consensus Seeking Over Weighted Directed Graphs, *IEEE Trans. Automat. Control*, vol. 52(9), 2007, pp.1767-1773.
- [11] K. Liu, and A. Sayeed, Type-Based Decentralized Detection in Wireless Sensor Networks, *IEEE Trans. on Signal Proc.*, vol. 55(5), 2007, pp.1809-1910.
- [12] S. Boyd and L. Xiao, Fast Linear Iterations for Distributed Averaging, Systems and Control Letters, vol.53, 2004, pp.65-78.
- [13] G. Yin, V. Krishnamurthy, and C. Ion, Regime Switching Stochastic Approximation Algorithms with Application to Adaptive Discrete Stochastic Optimization, *SIAM J. Optim.*, vol.14(4), 2004, pp.1187-1215.