Average-Consensus with Switched Markovian Network Links

Kevin Topley, Vikram Krishnamurthy
Department of Electrical Engineering
University of British Columbia
Vancouver, BC V6T 1Z4, Canada
kevint@ece.ubc.ca, vikramk@ece.ubc.ca

George Yin
Department of Mathematics
Wayne State University
Detroit, MI 48202, U.S.A.
gyin@math.wayne.edu

Abstract – Decentralized network estimation of the average initial node value is considered here in a variety of stochastic network settings. Each setting assumes the elements of the network communication graph edge set are modeled as a collection of ergodic Markov chains with slowly switching regime and unknown stationary distribution. In this framework an asymptotic average-consensus is obtained by using a “damped” distributive averaging algorithm in conjunction with an adaptive weighting scheme. The weighting scheme is designed to off-set the unknown probabilities of node communication by associating with each transmission a weight that is inversely proportional to the current estimate of the nodes communication probability. It is shown that for suitably connected graphs with balanced edge sets, and in particular any connected undirected edge set, the weighting scheme and averaging algorithm together yield a network consensus that is bounded to within an arbitrary distance of the average initial node value. The asymptotic node value scaled error measured relative to the node steady-state is also characterized by a vanishing diffusion equation with parameters that approach zero only as the nodes approach consensus. Simulations employ the proposed algorithm to demonstrate its proficiency and illustrate our results.

Keywords: Average-consensus, switching Markovian network links, distributed averaging, two-time-scale model

1 Introduction

Due to its relatively wide application base, a considerable amount of work in recent years has been devoted to obtaining conditions in which an average of node initial values can be accurately estimated by all nodes in a sparsely connected network by sole use of a distributive linear averaging consensus protocol (for a small sample of the literature see [2, 3, 4, 15, 14, 5, 6, 8, 19, 10] for past treatments of this problem). It is evident that the distributive linear averaging consensus protocol is exceptionally tractable analytically, completely decentralized, and in certain settings can also be considered memoryless [18, 12]. The basic protocol assumes the network communication is parameterized by a graph $G = \{V, E, W\}$ with directed edge set $E \subseteq V \times V$, node set $V = \{1, \ldots, n\}$, and weighted adjacency matrix $W$ with elements $W_{ij} \neq 0$ only if $(j, i) \in E$. Under various conditions on $G$, the consensus algorithm we consider results in a “network estimate” that is shared by all nodes, the nodes thus obtain consensus by a coupled action that is similar to the action of autonomous agents in a multi-agent system [14]. The basic consensus problem analyzed by many of the cited works has been utilized in a variety of applications, for example as part of a distributed Kalman filter [17], UAV co-ordination [11], the analysis of swarming behaviors [10], network load balancing [6], and oscillator synchronization [7].

In this paper we consider within a novel framework the same average-consensus problem as studied for instance in [5, 15]. Specifically, we consider how each node can obtain the average of all initial node values by means only of averaging communicated values under a stochastically switching network communication graph $G$. In addition to the existing literature we consider a probabilistic communication model that eliminates the possibility of average-consensus by any linear averaging algorithm with fixed averaging weights. In order to obtain the desired average-consensus we propose a stochastic approximation (SA) implemented independently at each node. The SA acts as an on-line “learning” algorithm and can be used to make local weight adjustments that will asymptotically eliminate the distance between each node state-value and the collective initial average when employing the distributive averaging consensus protocol. We thus effectively solve the average-consensus problem in this setting. The dynamics of the entire network consensus model are presented by simulation, and by then comparing with “naive” averaging schemes we demonstrate clearly the purpose and adequacy of the proposed algorithm.

The outline of the paper is as follows: Section §2 describes the problem formulation and rationale for the proposed consensus algorithm. The algorithm is detailed in Section §3, and a diffusion limit is discussed in §3.1. Possible extensions and an alternative solution is presented in §4. For convenience, we have illustrated the numerical simulations of each algorithm at the time each algorithm is introduced, thus we do not have a separate section devoted only to simulations and numerical evidence. Finally, our conclusions are given in Section §5.
2 Problem Formulation

At discrete-time iterations \( k \in \mathbb{N} \) we consider each node \( i \in \mathcal{V} \) may possibly send a signal to a “neighboring” node \( j \) that belongs to some fixed set \( N_i \subseteq \mathcal{V} \). The signal sent by node \( i \) contains both its current state-value \( s^i \in \mathbb{R} \), as well as its node identification (i.e. the integer \( i \) itself). The receiving node \( j \) thus knows when it has received a signal from node \( i \); however it is not assumed that node \( i \) will know which node, if any, received its signal. In other words, for each \( j \in \mathcal{V} \) the node \( j \) knows each \( i \in \mathcal{V} \) for which \( j \in N_i \), however the node \( j \) does not know \( N_i \). It is clarified now that we allow for the possibility \( i \in N_i \), thus each node \( i \) may “send” itself its current state-value; however this event is treated equivalently to the event of no communication from node \( i \).

We model the sequence of nodes \( \{j_i(k) : k \in \mathbb{N}\} \) from which node \( i \) receives signals as a Markov chain \( \{X_i(k)\} \) with an \( n \)-dimensional state-space \( S^i \) denoted by the set standard unit vectors \( \{e_1, \ldots, e_n\} \) of \( \mathbb{R}^n \), thus

\[
X^i(k) = e_j \Rightarrow (j,i) \in \mathcal{E}(k) . \tag{2.1}
\]

The above equation represents the possibility of communication from node \( j \) to node \( i \), but makes no statement about communication in the reverse direction. We assume that each Markov chain \( X^i \) has a stationary distribution \( \pi^i \in \mathbb{R}^n \) that is unknown to all nodes. Denoting the \( j^{th} \) element of \( \pi^i \) as \( \pi_{ij} \), we may define

\[
N_i = \{ j : \pi_{ij} > 0, j \in \mathcal{V} \} . \tag{2.2}
\]

By the virtue of including a node i.d. in the communicated signal, each receiving node \( j \in N_i \) can estimate the element \( \pi_{ij} \) by maintaining an empirical measure,

\[
\hat{\pi}_{ij}(k) = \frac{1}{k} \sum_{l=1}^{k} \mathbb{I}(X^i(l) = e_j) , \tag{2.3}
\]

where we have let \( e_j \) denote the \( j^{th} \) standard unit vector in \( \mathbb{R}^n \) and take \( \mathbb{I}(\cdot) \) to be a binary indicator function.

Similar to many recent works on consensus formation, we assign to each node \( i \in \mathcal{V} \) an initial state-value \( s^i(0) \in \mathbb{R} \) with no especial interpreted meaning. The average-consensus problem then requires each node to obtain an accurate estimate of \( \bar{s}(0) = \frac{1}{n} \sum_{i=1}^{n} s^i(0) \). To solve this problem we consider the well-known distributive averaging algorithm,

\[
s^i(k+1) = s^i(k) + \sum_{j=1}^{n} W_{ij}(k+1)(s^j(k) - s^i(k)) , \tag{2.4}
\]

for each \( i \in \mathcal{V} \), where we the weights \( W_{ij}(k) \in \mathbb{R} \) to be time-varying and for simplicity denote \( \mathbb{I}_{ij} = \mathbb{I}(X^i(k) = e_j) \).

It is emphasized that by using the indicator term \( \mathbb{I}_{ij} \), the iteration (2.4) may be applied to a stochastic network setting wherein the set of neighbors transmitting/receiving signals to/from a given node may vary at each iteration. The receiving nodes weight each incoming signal based on the transmitting node i.d.; on the other hand a node does not perform any averaging if no signal arrives. For a fixed communication network (2.4) simplifies to,

\[
s^i(k+1) = s^i(k) + \sum_{j=1}^{n} W_{ij}(k+1)(s^j(k) - s^i(k)) , \quad i \in \mathcal{V} . \tag{2.5}
\]

It can be proven that if the graph is strongly connected, that is if there exists a directed path between any two nodes, then for a balanced weighting scheme, that is for weights satisfying

\[
\sum_{j=1}^{n} W_{ij}(k) = \sum_{j=1}^{n} W_{ji}(k) \quad \forall \ i \in \mathcal{V}, \ k \in \mathbb{N} , \tag{2.6}
\]

the state-value \( s^i \) of each node under (2.5) will asymptotically reach the average-consensus \( \bar{s}(0) \) provided that for each \( k \in \mathbb{N} \) all non-zero off-diagonal weights are positive and \( \sum_{j \neq i} W_{ij}(k) < 1 \) for each \( i \in \mathcal{V} \). Letting \( \mathbb{I} \) denote a column vector of ones with appropriate dimension, it is clear (2.6) can be expressed \( \forall \mathbb{I} = \mathbb{I} \), furthermore we note that the above conditions are sufficient, but not necessary, for the average-consensus to be obtained. These conditions are nonetheless sufficient to solve the average-consensus problem in the current setting, for this reason we do not explore alternative schemes that might involve negative averaging weights.

The advantages of distributive averaging by (2.4) has been clarified in terms of its scalability with the network size \( n \), its robustness to communication link failures, and the potential convergence rate to consensus (for instances of these discussions, see [15, 5, 6, 14, 16]). In contribution to the growing literature on the topic, we seek the average-consensus \( \bar{s}(0) \) for stationary Markov chain sequences \( \{j_i(k) : k \in \mathbb{N}\} \). We furthermore assume each Markov chain \( X^i \) has a slowly switching regime. Specifically we model the distribution \( \pi^i \) as conditional on the state of a slow Markov chain \( \{\theta_k\} \) (for notational convenience we denote the time dependence of \( \theta \) as a subscript, e.g. \( \theta_k \)). We differentiate between the time-scale on which \( X^i \) and \( \theta \) evolve by parameterizing the transition matrix \( P^\varepsilon \) of \( \theta_k \) by using small parameter \( 1 \gg \varepsilon > 0 \) and defining,

\[
P^\varepsilon = I + \varepsilon Q , \tag{2.7}
\]

where we take \( Q \) to be the generator of a continuous-time Markov chain with finite state-space \( \mathcal{M} \). We note that although \( P^\varepsilon \) scales to identity as \( \varepsilon \) vanishes, each of the distributions \( \{\pi^i : i \in \mathcal{V}\} \) remain unchanged, thus for small \( \varepsilon \) the state of \( \theta \) varies slowly relative to the states of each Markov chain \( X^i \). For other works on consensus involving Markovian dynamics see [13]. For works related to consensus in other time-varying or probabilistic settings see [14, 8, 19, 15].

**Connectivity Conditions.** A network consensus by any means will require, by logical necessity, that the communication graph is in some sense strongly connected. In other words, loosely speaking each node must possess over some sufficiently dense collection of time intervals a communication path to every other node in the network. As we consider...
each stationary distribution $\pi_i$ to switch with each $\theta \in \mathcal{M}$, then minimal conditions for strong connectivity in the current setting may be stated,

$$
\bigcup_{\theta \in \mathcal{M}} \cap_{i \in \mathcal{V}} \bigcup_{j \in \mathcal{V}} \{ j : \pi^i_j(\theta) > 0 \} = \mathcal{V} .
$$

(2.8)

For reasons related to the requirements of average-consensus under the distributed weighting scheme, we will assume the edge set $\mathcal{E}$ is balanced, that is

$$
\mathcal{C}(\{ j : (i, j) \in \mathcal{E} \}) = \mathcal{C}(\{ i : (j, i) \in \mathcal{E} \}) \forall i, j \in \mathcal{V} ,
$$

(2.9)

where $\mathcal{C}(\cdot)$ denotes cardinality. The condition (2.9) can be stated $\mathcal{C}(\mathcal{N}_i) = \mathcal{C}(\mathcal{N}_i)$ $\forall i \in \mathcal{V}$. As the neighborhoods $\mathcal{N}_i$ are defined by (2.2), we note that (2.9) then implies,

$$
\sum_{j=1}^{n} I(\pi^{ij}(\theta) > 0) = \sum_{i=1}^{n} I(\pi^{ij}(\theta) > 0) , \text{ for each } \theta \in \mathcal{M} .
$$

(2.10)

It is emphasized the stationary probabilities $\pi^{ij}(\theta)$ themselves need not be balanced, thus we do not require

$$
\sum_{j=1}^{n} \pi^{ij}(\theta) = \sum_{j=1}^{n} \pi^{ij}(\theta) \forall i \in \mathcal{V} \text{ for each } \theta \in \mathcal{M} .
$$

2.1 Motivation for Algorithm.

We here explain why (2.4), as it is expressed above, will not result in the average-consensus $s(0)$ when node communication is modeled by the Markov parameters $\{X(k), \theta_k\}$. We then motivate the adaptive weighting scheme by modifying (2.4) in a way such to improve its ability to reach the average-consensus.

Case 1: Unscaled Averaging. We consider the simplified case when $\mathcal{C}(\mathcal{M}) = 1$ and so $\theta$ is necessarily fixed, generalizations to the case when $\theta$ is time-varying will be made subsequently. Consider the iteration (2.4) with near unit-valued weights, that is for some $1 \ll \epsilon > 0$ take $W_{ij} = (1 - \epsilon)$ if $(j, i) \in \mathcal{E}$ and $W_{ij} = \epsilon$ otherwise. Similar to the proof contained in [19], it can be shown that in this case, even for $\epsilon = 0$, the node state-values under the consensus algorithm (2.4) will almost surely (a.s.) result in an unpredictable consensus point within the convex hull of $s(0)$, see Fig.1 for some general simulations. We note that due to the Markov assumption regarding the reception of signals at each node $i$, there is at most one signal that is received per iteration at any given node and thus for an arbitrarily small $\epsilon$ the near unit averaging scheme satisfies the property

$$
\sum_{j \neq i} W_{ij}(k) < 1 , \forall i \in \mathcal{V} , \forall k \in \mathbb{N} .
$$

(2.11)

We note that for $\epsilon = 0$ this scheme is essentially a “replacement” algorithm, whereby the receiving nodes do not share any of there currently held information. In contrast, for any connected graph with $\epsilon > 0$ the initial information of all nodes will be contained in the consensus value, see Fig.1 - 2 for comparisons. At each time $k$ that (2.11) holds, even strictly, then in the current Markov setting the weighted average of all received signals will still be contained in the convex hull of all node state-values $\{s^i(k) : i \in \mathcal{V}\}$ at the time $k$. In figure 3 we have show an example case when the condition (2.11) at some iterations does not hold. Visibly we can see these occurrences by the spikes which take a node state-value beyond the convex hull of all node state-values at the previous iteration.

We note, for reference to the possible extensions of our model discussed in §4, that by constraining all weights to $(0, 1/n)$ we may assume up to $n$ signals received at each node per iteration, or in other words the superposition of $n$ Markov chain signal processes observed by each node. Assuming (2.8), the weight constraint $(0, 1/n)$ then ensures an a.s. consensus that will lie strictly within the convex hull $s(0)$,
where we define the vector \( s(k) = [s_1(k), \ldots, s_n(k)]' \in \mathbb{R}^n \), the Markov chains \( X_k = [X^1_k, \ldots, X^n_k]' \in \mathbb{R}^{n \times n} \), and the Laplacian matrix

\[
\mathcal{L}(k) = \text{diag}(\mathcal{I} \ast \mathcal{W}(k) \ast X(k)) - \mathcal{I} \ast \mathcal{W}(k) \ast X(k) .
\] (2.15)

We let \( \ast \) represents element-wise multiplication between two matrices with identical dimensions, the operation \( \text{diag}(\cdot) \) return a diagonal matrix with elements corresponding to those of its vector argument, and \( \mathcal{I} = (I_{ij}) = (I(X^i(k) = e_j)) \in \mathbb{R}^{n \times n} \).

We now assume that in the limit \( \mu \) vanishes the sequence of weights \( \{\mathcal{W}_{ij}(k) : k \in \mathbb{N}\} \) converge to uniformly bounded piece-wise differentiable trajectories \( T(\mathcal{W}_{ij}, \cdot) \). In other words, for a finite \( M > 0 \) we have \( M > T(\mathcal{W}_{ij}, k) \geq 0 \), \( \forall k \in \mathbb{N} \), and also for finite valued \( a_k \in \mathbb{R} \),

\[
\lim_{\mu \to 0} \frac{T(\mathcal{W}_{ij}(k) - T(\mathcal{W}_{ij}, k - 1))}{\mu} = a_k \, , \quad \forall i, j \in \mathcal{V}
\] (2.16)

except at a countable number of times \( k^* \) at which \( T(\mathcal{W}_{ij}, \cdot) \) is not differentiable for each \( i, j \in \mathcal{V} \). In figures 4 - 5 we provide some illustrative trajectories of the limiting algorithm in the present stochastic model for finitely bounded arbitrary fixed and time-varying positive weights satisfying (2.16).

**Figure 4:** An example of the Network Node State-Value asymptotic convergence to an a.s. predictable consensus value when using arbitrary and finitely bounded piece-wise fixed positive valued weights. By (2.19) the consensus point is given by \( \mathbb{I} w^T s(0) \), as indicated by the horizontal line. The only non-differentiable point in the weight trajectories is indicated by the vertical line. We assume \( \mathcal{L}(\mathcal{M}) = 1 \) and thus \( \theta \) is fixed.

In both cases the continuous-time interpolated sequence of node state-values \( s^\mu(t) = s(k) \) if \( t \in [k\mu, (k + 1)\mu) \), will weakly-converge in the limit as \( \mu \) vanishes to solutions \( s(\cdot) \) of the switching ODE,

\[
\frac{d}{dt}s(t) = -\mathcal{L}(t)s(t) \, , \quad t \geq 0 ,
\] (2.17)

where we define the continuous time variable \( t > 0 \) as \( t = t_k^\mu = k\mu \), if \( t \in [k\mu, (k + 1)\mu) \), the vector \( s(t) = [s^1(t), \ldots, s^n(t)]' \in \mathbb{R}^n \), and the Laplacian matrix

**Case 2: Scaled Averaging.** We have argued that for arbitrarily fixed weights in \( (0, 1) \) in the original stochastic network model each node state-value will converge a.s to an unpredictable consensus, illustrated in Fig. 1–2. We do not attempt to characterize the distribution of the consensus resulting from (2.4) in the given stochastic network model and weighting scheme, rather we will demonstrate, in both mathematical arguments and simulated numerical verifications, that for suitably well-behaved positive weights \( \mathcal{W}_{ij}(k + 1) \geq 0 \), \( \forall (i, j) \in \mathcal{E}, \forall k \in \mathbb{N} \) the node state-values will weakly-converge to an a.s. \( \theta \)-dependent consensus within the convex hull of \( s(0) \) under the following scaled consensus algorithm,

\[
s^\mu(k + 1) = s^\mu(k) + \mu \sum_{j=1}^n \mathcal{W}_{ij}(k + 1) I_{ij} (s^\mu(k) - s^\mu(i))
\] (2.13)

in the asymptotic limit \( \mu \to 0 \). In the current stochastic model, the iteration (2.13) can be written more succinctly as,

\[
s(k + 1) = (I - \mu \mathcal{L}(k + 1))s(k) ,
\] (2.14)

the reader may consider [14] for related discussions. However if we assume the weight constraint \((0, 1/n)\) then we may also consider the signal transmission at each node \( i \) being modeled as the Markov chain \( X^i \), thus instead of (2.1) we have

\[
X^i(k) = e_j \Rightarrow (i, j) \in \mathcal{E}(k) .
\] (2.12)

In this setting there can be at most \( n \) signals received by any given node \( j \in \mathcal{V} \), thus if all weights are contained in \( (0, 1/n) \) a consensus within the convex hull of \( s(0) \) will a.s. result provided also (2.8) holds. We propose this setting as it is more likely the signal reception is a stochastic process but the signal transmission is initiated deterministically at each iteration by a given node. Conversely, under (2.1) it is assumed the reception of a signal can be initiated at each node upon every iteration, we further detail the relation between these two assumptions in §4.

**Figure 3:** An example of the Network Node State-Value convergence to, again, an unpredictable consensus value when using arbitrarily fixed positive valued weights some of which are greater than 1. The spikes that disrupt the convex hull of the node state-values indicate the discrete-times \( k^* \in \mathbb{N} \) for which the constraint (2.11) did not hold.
Figure 5: Two examples of the Network Node State-Value asymptotic convergence to an a.s. predictable consensus value when using arbitrary piece-wise differentiable positive valued sinusoidal functions $T(W_{ij}, t) = (I(t < t_1)(a_1 + b_1 \cdot \sin(t)) + I(t < t_2)(a_2 + b_2 \cdot \cos(t)))/\sum_{i \neq j} W_{ij}(\theta_i) > 0$ of the weights. We let $L = 1$, $t_1 = 100$ and fix $(a_1, b_1, a_2, b_2)$ as arbitrary constants. The a.s consensus point is given by (2.19) and is indicated by the horizontal line. The only non-differentiable point in the weight trajectories is indicated by the vertical line. We assume $C(M) = 1$ and thus $\theta$ is fixed.

In this case each node $i \in \mathcal{V}$ can estimate by the empirical measure (2.3) the relative frequencies with which it receives signals from its respective neighbors. By then weighting these signals inversely proportional to their frequency, given that $C(M) > 1$, the only reason an average-consensus does not occur in this setting is the result of the “naive” distributive averaging that occurs when estimates of $\pi(\theta_k)$ are not sufficiently close to the actual communication probabilities $\pi(\theta_k)$. We consider below separately the two cases $C(M) = 1$ and $C(M) > 1$. We find that in the former case the asymptotic network consensus can arbitrarily well-approximate (a.w.a) the average-consensus on a time-scale that is an order of magnitude smaller than the time-scale on which average-consensus may be a.w.a when $C(M) > 1$.

We clarify the abbreviation "v a.w.a. w" is used whenever there is an asymptotic element-wise a.s. equality between two matrices $v$ and $w$ of the same dimension.

In the limit as $\mu$ vanishes the empirical measure (2.3) may be expressed for sufficiently large $k$ as,

$$
\hat{\pi}_k + \mu(X_k - \hat{\pi}_k), \, \hat{\pi}_0 = X(0), \, i \in \mathcal{V}.
$$

Thus we find by [9] that the interpolated sequence of each node estimate $\hat{\pi}(\theta_k)$ will weakly-converge to a solution $\hat{\pi}(\cdot)$ of the switching ODE,

$$
d\hat{\pi}(t) = -\hat{\pi}(t) + \pi(t), \, \hat{\pi}(0) = X(0), \, t \geq 0.
$$

Under the weighting scheme,

$$
W_{ij}(k) = \frac{\alpha}{\pi^ij(k)}, \, \alpha > 0,
$$

the interpolated sequence of each weight $W_{ij}(k)$ will then weakly-converge to the finitely bounded piece-wise differentiable trajectory

$$
W_{ij}(t) = \alpha/(e^{-t}(\hat{\pi}^ij(0) - \pi^ij(\theta_i)) + \pi^ij(\theta_i)),
$$

and thus the asymptotic Laplacian $L(t)$ can be expressed element-wise as,

$$
L_{ij}(t) = \begin{cases} 
\alpha T^ij/(e^{-t}c_{ij}(t) + 1) & \text{if } i \neq j \\
\alpha \sum_{j \neq i} T^ij/(e^{-t}c_{ij}(t) + 1) & \text{if } i = j
\end{cases}
$$

where we now define $T^ij = \pi^ij(\theta_i) > 0$ and $c_{ij}(t) = \pi^ij(t_{ij})/\pi^ij(\theta_i)$, the sequence $\{t_1, \ldots, t_L\}$ indicates the times at which $\theta_i$ switches state, and $t_{ij} = \max\{t_l : t_l \leq t\}$.

We note that under (3.5) by defining $L^\alpha(t) = (L_{ij}^\alpha(t))$ as

$$
L_{ij}^\alpha(t) = \begin{cases} 
-\pi^ij/(e^{-t}c_{ij}(t) + 1) & \text{if } i \neq j \\
\sum_{j \neq i} -\pi^ij/(e^{-t}c_{ij}(t) + 1) & \text{if } i = j
\end{cases}
$$

we then have

$$
\int_t^0 L(u)du = \int_0^t L^\alpha(u)du, \, \forall \alpha > 0.
$$

3 Adaptive and Damped Weights

We assume the node nodes do not know $\{\pi^ij(\theta_k) : i \in \mathcal{V}\}$, or more specifically each node $i \in \mathcal{V}$ does not know $\pi^i$. The only times at which the matrix $W$ switches is indicated by $t_1, \ldots, t_L$ and is shown in [15]. In this case, as follows from [15], the asymptotic consensus is the average $\bar{s}(0)$ if and only if $W \ast \pi = (W \ast \pi)^j \ast \mathbb{I}$ and as well (2.8) holds. This basic result leads to the proposed damped weighting scheme described in §3.

3.1 Almost Sure Convergence to Consensus with Time-Varying Weights

Let $\theta_t = \theta_t$ be the Markov chain of states $\theta_t$ of the switching ODE, where we take $t_0 = 0$, $t_{L+1} = t$, and the set of points $\{t_1, \ldots, t_L\}$ indicate the times at which $T(W_{ij}, \cdot)$ is not differentiable for each $i, j \in \mathcal{V}$. We can see by (2.19) that taking the limit $t \to \infty$ the network steady-state can be expressed

$$
limit_{t \to \infty} s(t) = \lim_{t \to \infty} \prod_{i=0}^{L} e^{-\sum_{j=0}^{L} T^ij/(e^{-t}c_{ij}(t) + 1)} s(0),
$$

where the second line assumes we know each time $t_i$ at which $T(W_{ij}, \cdot)$ is not differentiable.

If the weights are fixed then $dT(W_{ij}, t)/dt = 0$ for all $t \geq 0$ and so the right-hand side of (2.19) simplifies to $\lim_{t \to \infty} e^{-L}(s(0)$ and can be computed as $w_i^j s(0)$ where $w_i \in \mathbb{R}^n$ satisfies both $w_i^j L = 0$ and $w_i^j \mathbb{I} = 1$, as shown in [15]. In this case, as follows from [15], the asymptotic consensus is the average $\bar{s}(0)$ if and only if $W \ast \pi = (W \ast \pi)^j \ast \mathbb{I}$ and as well (2.8) holds. This basic result leads to the proposed damped weighting scheme described in §3.

3.2 Continuous-Time Almost Sure Convergence to Consensus with Time-Varying Weights

The only times at which the matrix $W$ switches is indicated by $t_1, \ldots, t_L$ and is shown in [15]. In this case, as follows from [15], the asymptotic consensus is the average $\bar{s}(0)$ if and only if $W \ast \pi = (W \ast \pi)^j \ast \mathbb{I}$ and as well (2.8) holds. This basic result leads to the proposed damped weighting scheme described in §3.

3.3 Almost Sure Convergence to Consensus with Time-Varying Weights

The only times at which the matrix $W$ switches is indicated by $t_1, \ldots, t_L$ and is shown in [15]. In this case, as follows from [15], the asymptotic consensus is the average $\bar{s}(0)$ if and only if $W \ast \pi = (W \ast \pi)^j \ast \mathbb{I}$ and as well (2.8) holds. This basic result leads to the proposed damped weighting scheme described in §3.

3.4 Almost Sure Convergence to Consensus with Time-Varying Weights

The only times at which the matrix $W$ switches is indicated by $t_1, \ldots, t_L$ and is shown in [15]. In this case, as follows from [15], the asymptotic consensus is the average $\bar{s}(0)$ if and only if $W \ast \pi = (W \ast \pi)^j \ast \mathbb{I}$ and as well (2.8) holds. This basic result leads to the proposed damped weighting scheme described in §3.
Case 1: Fixed $\theta$ ($C(\mathcal{M}) = 1$). It is clear by (3.6) that for any fixed $\theta \in \mathcal{M}$ each element of the Laplacian $L^{\alpha}$ monotonically approaches a steady-state such that $L^{\alpha}$ is balanced, specifically $L^{\alpha}$ will element-wise monotonically approach the matrix $L^{bal}$ as defined,

$$L^{bal}_{ij}(t) = \begin{cases} T_{ij} \pi_i & \text{if } i \neq j \\ \sum_{j \neq i} T_{ij} \pi_i & \text{if } i = j \end{cases}.$$ \hspace{1cm} (3.8)

In other words $L^{\alpha}$ a.w.a. $L^{bal}$ for sufficiently large $t$.

It follows then taking $\alpha$ sufficiently small we can for any $t > 0$ ensure $tL^{bal}$ will a.w.a. $\int_0^t L(u)du$. This implies by (2.19) and the work in [15] that for sufficiently small $\alpha$ the average-consensus $\bar{s}(0)$ will a.w.a. the asymptotic network consensus point.

Figure 6: Network Weak-Convergence to Consensus. When $\alpha$ decreases the rate of averaging of slows, but the asymptotic network consensus point becomes closer to the average-consensus.

Case 2: Slowly time-varying $\theta$ ($C(\mathcal{M}) > 1$). If $\theta$ is slowly time-varying then $L^{\alpha}$ does not monotonically approach $L^{bal}$, rather at every time $\theta$ switches state each element of $L^{\alpha}(t)$ will possess a discontinuous jump change in value that is not necessarily closer to the corresponding elements in $L^{bal}$. Thus, although for any given sample path of $\theta$ the transformation (3.7) will still hold, if there is no time $t^*$ after which $\theta_t$ will remain fixed then there will be no time at which $L^{\alpha}(t)$ will monotonically approach $L^{bal}$. Thus even for vanishing values of $\alpha$ the equality (3.7) does not imply $tL^{bal}$ will a.w.a. $\int_0^t L(u)du$, and hence average-consensus is not necessarily obtained.

However, by the recent results [12] it has been shown that if $\theta$ itself has a $C(\mathcal{M}) = m$-dimensional stationary distribution $\nu \in \mathbb{R}^m$ satisfying $\nu^TQ = 0$, then considering the limit as $\mu$ vanishes and $k\mu \to \infty$, a piece-wise constant interpolation of the sequence $\{\tilde{\pi}_i(k)\}$ will weakly-converge to the weighted average $\tilde{\pi}_i$ of the distributions $\{\pi_i(\theta_k) : \theta \in \mathcal{M}\}$,

$$\tilde{\pi}_i = \sum_{l=1}^m \pi_i(\theta^l)\nu_l,$$ \hspace{1cm} (3.9)

where we denote the state-space of $\theta$ as $\mathcal{M} = \{\theta^1, \ldots, \theta^m\}$ and $\nu_l$ denotes the stationary probability that $\theta = \theta^l$. Like-wise the sequence of communication stationary distributions $\{\pi_i(\theta_k) : i \in \mathcal{V}\}$ will on the time-scale $k\mu \to \infty$ each respectively weakly-converge to the stationary measures $\{\pi_i^*: i \in \mathcal{V}\}$, where we define $\pi_i^* = \sum_{l=1}^m \pi_i(\theta^l)\nu_l$. Thus on this larger time-scale both the estimates $\tilde{\pi}_i$ and distribution of the communication signals $\{X^i_k\}$ weakly-converge to a common distribution (3.9) for each $i \in \mathcal{V}$.

Under the weighting scheme (3.3) it can be shown that for $\alpha = O(\mu)$ the interpolated state-values under the doubly-scaled consensus algorithm,

$$s^i(k+1) = s^i(k) + \mu \alpha \sum_{j=1}^n W_{ij}(k+1) \mathcal{I}_{ij}(k+1)(s^j(k) - s^i(k))$$ \hspace{1cm} (3.10)

will for each $i \in \mathcal{V}$ weakly-converge in the limit $\mu \to 0$ on the larger continuous-time scale $t = t_k = k\mu \alpha = k\mu s$ if $t \in [k\mu, (k+1)\mu]$. \hspace{1cm} (3.11)

where, due to the weak-convergence of both $\tilde{\pi}_i$ and $\{X^i_k\}$ to (3.9) for each $i \in \mathcal{V}$, we have each non-zero off-diagonal element of $\mathcal{L}$ equal to unity. Provided then (2.10) holds, under the strong connectivity condition (2.8) an a.s. asymptotic consensus to $\bar{s}(0)$ will result on the larger time-scale $k\mu \to \infty$.

3.1 Diffusion Limit.

For finitely bounded positive weights both of the iterations (2.13) and (3.10) result asymptotically in a network node state-value dynamic that is a.s. strictly convex,

$$\text{conv}\{s(t)\} \subset \text{conv}\{s(t+\epsilon)\} \text{ for all } \epsilon > 0.$$ \hspace{1cm} (3.12)

Due to these convex dynamics the node state-values weakly-converge to trajectories that, under the connectivity condition (2.8), will a.s. become an identical scalar, say $\bar{s}(0)$.

By defining the scaled tracking error,

$$v^i_k = \frac{s^i(k) - \bar{s}(0)}{\sqrt{\mu}},$$ \hspace{1cm} (3.13)
and the corresponding interpolated sequence \( \{ v^i_k(t) = v^i_k(k), \ t \in [k \mu, (k + 1) \mu) \} \), it can be shown as well as verified numerically that the sequence of scaled errors \( \{ v^i_k(k) \} \) will weakly-converge to solutions \( \{ v^i(\cdot) : i \in \mathcal{V} \} \) of the switching diffusion,

\[
dv_i(t) = -v_i(t) \mathcal{L}^i(t)s^i_{\text{dif}}(t) \, dt + s^i_{\text{dif}}(t) \Sigma_{\pi_i}^{1/2}(\theta) \mathcal{W}^i \, dw
\]

(3.14)

where \( w(\cdot) \) represents a standard Brownian motion and we define \( \mathcal{L}^i(t) \) as the \( i^{th} \) row of \( \mathcal{L}(t) \) and likewise \( \mathcal{W}_i \) the \( i^{th} \) row of the \( \mathcal{W}_i \). By doing so the nodes would avoid averaging that is dictated by a sequence of unbalanced Laplacians \( \mathcal{L}(t) \), \( 0 \geq t < t^* \). In this solution then each node simply waits to begin averaging of incoming signals only until after some measure of variation in its own estimate of \( \pi^i(\theta) \) reaches a sufficiently small upper bound, thus implying the estimate is stable and hence accurate (see [1] for related discussions). By the weighting scheme (3.3), presuming the balanced edge condition (2.10) and strong connectivity (2.8), an a.s. asymptotic consensus to \( \bar{s}(0) \) will then result.

Alternative Solution. Assuming \( \theta \) were fixed, each node could avoid averaging during some initial period \( t \in [0, t^* \) for which the estimates \( \{ \pi_i(t) : i \in \mathcal{V} \} \) are in far proximity of the actual communication probabilities \( \{ \pi_i(\theta) : i \in \mathcal{V} \} \). By doing so the nodes would avoid averaging that is dictated by a sequence of unbalanced Laplacians \( \mathcal{L}(t) \), \( 0 \geq t < t^* \). In this solution then each node simply waits to begin averaging of incoming signals only until after some measure of variation in its own estimate of \( \pi^i(\theta) \) reaches a sufficiently small upper bound, thus implying the estimate is stable and hence accurate (see [1] for related discussions). By the weighting scheme (3.3), presuming the balanced edge condition (2.10) and strong connectivity (2.8), an a.s. asymptotic consensus to \( \bar{s}(0) \) will then result.

If \( \theta \) is slowly time-varying then there is no guarantee that on the bounded time-scale \( t = t^a_k \) \( k \mu \), \( k \in [k \mu, (k + 1) \mu) \), a sufficiently small upper bound on the variation of the estimates \( \{ \pi_i(t) : i \in \mathcal{V} \} \) is ever reached. Thus again, similar to above, we require \( \alpha = O(\mu) \) so that the averaging algorithm will possess dynamics on the larger time-scale \( k \mu \to \infty \), and the estimates \( \pi_i \) converge to the steady-state (3.9) that is identical to the signal communication probabilities \( \pi_i \). On this larger time-scale, once some measure of variation in \( \pi_i \) reaches a sufficiently low upper bound, the node averaging may commence and a network consensus that arbitrarily closely approximates the average-consensus \( \bar{s}(0) \) will be obtained.

Extensions of the Stochastic Network Model. Due to the asymptotic limit as \( \mu \) vanishes, all of the above results hold when, rather than a single signal source at each iteration that behaves as a Markov chain \( X^i(k) \), there are multiple signal sources that result from an arbitrary finite number \( d \) of Markov chains \( \{ X^i_1(k), \ldots, X^i_d(k) \} \). The state of each of each Markov chain \( X^i \) determines the source of a signal at each discrete-time \( k \in \mathbb{N} \).

Alternatively we could, as described above Case 2 in §2.1,
assume the signal transmitted from each node $i$ would be received only by one node $j_i(k) \in V$ at each iteration $k$. The sequence of nodes $\{j_i(k)\}$ may then be modeled as the Markov chain $X_i(k)$ with stationary distributions $\pi_i(\theta)$ and same switching properties as above. Again a generalization could be made for multiple transmissions from each node with the receiving nodes dictated by the states of some finite number $d$ Markov chains $\{X_i^l(k), \ldots, X_i^d(k)\}$. Also any mixture of these reception and transmission models are, permissible, since for any family of stationary distributions \{\pi_i^l : i \in V, l = 1, \ldots, l_i\} the matrix $\pi_R$, defined with every $i^{th}$ row equal to $\sum_{l=1}^{d} \pi_i^l$, which thus implies the reception of signals at node $i$ is modeled as the $l_i$ Markov chains, can be readily equated on the off-diagonals to a matrix $\pi_T$ that satisfies $\pi_T^T \pi_T = \mathbb{I}^{d}$. This implies the transmission of signals at node $i$ is modeled as a set of $l_i$ Markov chains. All that is required is the transformation $\pi_{R,i} = 1 - \sum_{j \neq i} \pi_{R,j}^{(i)}$ for all $i \in V$, which does not affect the consensus iterations (2.13) or (3.10) in any way due to the null effect of a transmission from node $j$ to node $i$, thus the sequence of iterates will asymptotically be identical and the characterizations (2.17) - (3.15) hold.

5 Conclusions

We have proposed a weight adaptation and damped averaging algorithm that together ensure weak-convergence to a consensus bounded within an arbitrarily small distance of the average initial node state-values in a network with slowly switching but piece-wise fixed communication link probabilities, assuming a balanced edge set in each regime. An asymptotic diffusion equation was obtained to characterize the node state-value scaled error relative to the asymptotic consensus point, this diffusion vanishes as the values approach a consensus steady-state limit. By simulations the necessity of the proposed algorithm in the given stochastic network setting was demonstrated.

References


