Resilient Average Consensus in the Presence of Heterogeneous Packet Dropping Links

Christoforos N. Hadjicostis, Alejandro D. Domínguez-García, and Nitin H. Vaidya

***Abstract*— We address the average-consensus problem for a distributed system whose components (nodes) can exchange information via unreliable interconnections (edges) that form an arbitrary, possibly directed topology (digraph). We consider a general setting where heterogeneous communication links may drop packets with generally unequal probabilities, independently between different links. We develop a distributed linear-iterative algorithm in which nodes maintain and update certain values based on the corresponding values they successfully receive from their in-neighbors. We demonstrate that, even when communication links drop packets with unequal probabilities, the proposed algorithm allows nodes to asymptotically reach average-consensus almost surely, as long as the underlying (possibly directed) communication topology forms a strongly connected digraph. Additionally, we provide a bound on the algorithm convergence rate.**

I. Introduction and Background

Consider a set of interconnected nodes (which could be sensors in a sensor network, routers in a communication network, or unmanned vehicles in a multi-agent system). In the so-called *consensus problem*, each node possesses an initial value and the nodes need to follow a distributed strategy to agree on the same (*a priori* unknown) value by calculating some function of these initial values. If the consensus value is the average of the initial values, then the nodes are said to reach *average consensus*. Consensus (and average consensus) problems have received extensive attention from the control community due to their applicability to topics such as cooperative control, multi-agent systems, and modeling of flocking behavior in biological and physical systems (see, e.g., [1], [2] and references therein).

This paper develops linear-iterative algorithms for average-consensus when the interconnection topology is described by a *directed* graph (digraph) that is not necessarily fully connected and whose edges are *unreliable*, i.e., each edge may drop packets with some probability. Unlike our work in [3] (where we assume that the probability of a packet drop is the same for every link), in this paper, although we assume that, at each time step, a packet containing information from node i to node j (sent through an existing edge in the digraph) is dropped with some probability, and although we assume independence between packet drops at different time steps or different links, we allow the probabilities of packet drops to be heterogeneous at each link.

C. N. Hadjicostis is with the ECE Department of the University of Cyprus, Nicosia, Cyprus, and also with the ECE Department of the University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA. E-mail: chadjic@UCY.AC.CY.

A. D. Domínguez-García and N. H. Vaidya are with the ECE Department of the University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA. E-mail: {aledan,nhv}@ILLINOIS.EDU.

The starting point for this work is an algorithm that relies on two linear iterations [4], [5]; we refer to this algorithm as *ratio-consensus*. It is easy to see that, except for the initialization of both iterations, the ratio-consensus algorithm in [4], [5] (which assumes perfectly reliable communication links) is a particular case of a gossip-based algorithm proposed in [6] (which also assumes perfectly reliable communications), where the transition matrices describing each linear iteration are allowed to vary with time.

Apart from our work in [3], average-consensus in the presence of unreliable communications has also received attention in [7], [8], [9]. The work in [7] assumes that the graph describing the communication network is undirected and, when a communication link fails, it affects communication in both directions. Additionally, nodes have some mechanism to detect link unavailability and can compensate for it by rescaling their other weights The work in [8] can handle directed graphs and proposes two compensation methods to account for communication link failures. However, both compensation methods cannot guarantee consensus to the exact average of the initial values. The authors in [9] propose a strategy that corrects the errors in the state iteratively calculated by each node by acknowledging messages and retransmiting information an appropriate number of times.

II. Preliminaries

Here we provide necessary background on graph-theoretic notions and present the original ratio consensus algorithm in [4], [5], which requires reliable communication links; we also introduce the communication link availability model.

A. Ratio Consensus in the Presence of Reliable Links

The information exchange between nodes can be described by a directed graph $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$, where $\mathcal{V} = \{1, 2, \ldots, n\}$ is the vertex set (each vertex corresponds to a system node), and $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ is the set of edges, where $(j, i) \in \mathcal{E}$ if node j can receive information from node i. Note that $\mathcal{E}$ could be a proper subset of $\mathcal{V} \times \mathcal{V}$, but we require the graph $(\mathcal{V}, \mathcal{E})$ to be strongly connected. For notational convenience (and without loss of generality), we allow self-loops for all nodes in $\mathcal{G}$ (i.e., $(j, j) \in \mathcal{E}$ for all $j \in \mathcal{V}$). All nodes that can possibly transmit information to node j are called its in-neighbors, and are represented by the set $\mathcal{N}_j^- = \{i \in \mathcal{V} : (j, i) \in \mathcal{E}\}$. The number of in-neighbors of j (including itself) is called the in-degree of j and is denoted by $\mathcal{D}_j^- = |\mathcal{N}_j^-|$. The nodes that have j as neighbor (including itself) are called its out-neighbors and are denoted by $\mathcal{N}_j^+ = \{l \in \mathcal{V} : (l, j) \in \mathcal{E}\}$; the out-degree of node j is $\mathcal{D}_j^+ = |\mathcal{N}_j^+|$.

Let $y_0(j)$ be the initial value of node j; assuming that the communication network graph $(\mathcal{G},\mathcal{E})$ is strongly connected, we showed in [4], [5] that average consensus can be asymptotically reached by executing

$$y_{k+1}(j) = \sum_{i\in\mathcal{N}_j^-} \frac{1}{\mathcal{D}_i^+} y_k(i), \tag{1}$$

$$z_{k+1}(j) = \sum_{i\in\mathcal{N}_j^-} \frac{1}{\mathcal{D}_i^+} z_k(i), \tag{2}$$

where $\mathcal{D}_j^+$ is the out-degree of node j, and $z_0(j)=1,\ \forall j$. Then, the nodes can asymptotically calculate the average of their initial values $\overline{y} = \frac{\sum_{l=1}^{n} y_0(l)}{n}$ as

$$\lim_{k\to\infty} \frac{y_k(j)}{z_k(j)} = \overline{y}\ , \forall j \in \mathcal{V}. \tag{3}$$

B. Unreliable Heterogeneous Communication Model

Under the unreliable heterogeneous communication model, the information exchange between nodes (components) at each (discrete) time instant k can be described by a directed graph $\mathcal{G}[k] = \{\mathcal{V},\mathcal{E}[k]\}$, where $\mathcal{G} = (\mathcal{V},\mathcal{E})$ is the digraph that describes the information exchange and $\mathcal{E}[k] \subseteq \mathcal{E},\ \forall k \geq 0$. [At instant k, let $x_k(j,i) = 0$ if $(j,i) \notin \mathcal{E}[k]$ and $x_k(j,i) = 1$ if $(j,i) \in \mathcal{E}[k]$.] The simplest probabilistic model for the communication link from node i to node j can be described as follows. Under our assumption that link availability is independent between links and between time steps, we can model, $x_k(j,i)$ as a Bernoulli random variable:

$$\Pr\{x_k(j,i) = m\} = \begin{cases} q_{ji}, & \text{if } m = 1, \\ 1 - q_{ji}, & \text{if } m = 0. \end{cases} \tag{4}$$

We also define the matrix $Q = [q_{ji}]$ where the entry at the j^{th} row, i^{th} column is q_{ji}, which is taken to be zero for $(j,i) \notin \mathcal{E}$ and can be set to one if the link is always available.

III. Resilient Ratio Consensus

In this section we briefly review the resilient ratio-consensus algorithm introduced in [3] to handle unreliable links that drop packets. The analysis there was limited to equal probabilities of packet drops; in this paper we (non-trivially) extend our previous results to the case when packet-drop probabilities are heterogeneous.

A. Resilient Ratio-Consensus Algorithm Formulation

We next describe how to robustify the ratio-consensus algorithm against packet dropping communication links. For the iteration defining the numerator of (3), let $y_k(j)$ be node j's internal state at time instant k, $\mu_k(l,j)$ denote the mass broadcast from node j to each of its out-neighbors l (this is the same for all out-neighbors of node j, i.e., for all $l \in \mathcal{N}_j^+$), and $\nu_k(j,i)$ denote the mass received at node j from node $i \in \mathcal{N}_j^-$. Similarly, for the iteration defining the denominator of (3), let $z_k(j)$ be node j's internal state, $\sigma_k(l,j)$ denote node j's broadcast mass to each out-neighbor $l,\ l \in \mathcal{N}_j^+$ (same for all out-neighbors), and let $\tau_k(j,i)$ denote the total mass received from $i \in \mathcal{N}_j^-$. Then, for all $k \geq 0$, each node j computes

$$y_{k+1}(j) = \frac{1}{\mathcal{D}_j^+} y_k(j) + \sum_{i\in\mathcal{N}_j^-} \big(\nu_k(j,i) - \nu_{k-1}(j,i)\big),$$

$$\mu_k(l,j) = \mu_{k-1}(l,j) + \frac{1}{\mathcal{D}_j^+} y_k(j) = \sum_{t=0}^{k} \frac{1}{\mathcal{D}_j^+} y_t(j), \tag{5}$$

where

$$\nu_k(j,i) = \begin{cases} \mu_k(j,i), & \text{if } (j,i) \in \mathcal{E}[k], \quad k \geq 0, \\ \nu_{k-1}(j,i), & \text{if } (j,i) \notin \mathcal{E}[k], \quad k \geq 0. \end{cases}$$

Similarly, for all $k \geq 0$,

$$z_{k+1}(j) = \frac{1}{\mathcal{D}_j^+} z_k(j) + \sum_{i\in\mathcal{N}_j^-} \big(\tau_k(j,i) - \tau_{k-1}(j,i)\big),$$

$$\sigma_k(l,j) = \sigma_{k-1}(l,j) + \frac{1}{\mathcal{D}_j^+} z_k(j) = \sum_{t=0}^{k} \frac{1}{\mathcal{D}_j^+} z_t(j), \tag{6}$$

where

$$\tau_k(j,i) = \begin{cases} \sigma_k(j,i), & \text{if } (j,i) \in \mathcal{E}[k], \quad k \geq 0, \\ \tau_{k-1}(j,i), & \text{if } (j,i) \notin \mathcal{E}[k], \quad k \geq 0. \end{cases}$$

Packet drops in a particular link affect both iterations in the same way. Also, the initial conditions are set to $y_0(j)$ (initial value of node j) and $z_0(j) = 1$ (as before), with $\mu_{-1}(j,i) = \nu_{-1}(j,i) = \sigma_{-1}(j,i) = \tau_{-1}(j,i) = 0$ for all $(j,i) \in \mathcal{E}$.

B. Vectorized Description of the Resilient Ratio-Consensus

In order to ease the calculations, the iterations in (5)–(6) will be rewritten more compactly in vector form. Using the definition for the indicator variable $x_k(j,i)$ given in (4), which describes the successful transmission of information from node i to node j over an existing, unreliable link, iterations (5) and (6) can be rewritten, for all $k \geq 0$, as

$$\mu_k(l,j) = \begin{cases} \mu_{k-1}(l,j) + \frac{1}{\mathcal{D}_j^+} y_k(j), & \text{if } l \in \mathcal{N}_j^+, \\ 0, & \text{if } l \notin \mathcal{N}_j^+, \end{cases} \tag{7}$$

$$\nu_k(j,i) = \begin{cases} \mu_k(j,i)x_k(j,i) + \\ +\nu_{k-1}(j,i)(1 - x_k(j,i)), & \text{if } i \in \mathcal{N}_j^-, \\ 0, & \text{if } i \notin \mathcal{N}_j^-, \end{cases} \tag{8}$$

$$y_{k+1}(j) = \sum_{i=1}^{n} \big(\nu_k(j,i) - \nu_{k-1}(j,i)\big); \tag{9}$$

$$\sigma_k(l,j) = \begin{cases} \sigma_{k-1}(l,j) + \frac{1}{\mathcal{D}_j^+} z_k(j), & \text{if } l \in \mathcal{N}_j^+, \\ 0, & \text{if } l \notin \mathcal{N}_j^+, \end{cases} \tag{10}$$

$$\tau_k(j,i) = \begin{cases} \sigma_k(j,i)x_k(j,i) + \\ +\tau_{k-1}(j,i)(1 - x_k(j,i)), & \text{if } i \in \mathcal{N}_j^-, \\ 0, & \text{if } i \in \mathcal{N}_j^-, \end{cases} \tag{11}$$

$$z_{k+1}(j) = \sum_{i=1}^{n} \big(\tau_k(j,i) - \tau_{k-1}(j,i)\big). \tag{12}$$

Let $A \circ B$ denote the Hadamard (entry-wise) product of a pair of matrices A and B of identical size. Then, for all $k \geq 0$, defining $M_k = [\mu_k(j,i)]$ and $N_k = [\nu_k(j,i)]$, iteration (7)–(9) can be rewritten in matrix form as

$$M_k = M_{k-1} + P\text{diag}(y_k), \tag{13}$$

$$N_k = M_k \circ X_k + N_{k-1} \circ (U - X_k), \tag{14}$$

$$y_{k+1} = (N_k - N_{k-1})e = \big[(M_k - N_{k-1}) \circ X_k\big]e, \tag{15}$$

where $P = [p_{ji}] \in \mathbb{R}^{n \times n}$, with $p_{ji} = \frac{1}{\mathcal{D}_i^+}$, $\forall j \in \mathcal{N}_i^+$ and $p_{ji} = 0$ otherwise; $M_{-1} = N_{-1} = 0$; $U \in \mathbb{R}^{n \times n}$, with $[U_{ji}] = 1,\ \forall i, j$; diag$(y_k)$ is the diagonal matrix that results by having the entries of vector y_k on the main diagonal; and $e = [1, 1, \ldots, 1]^T$ (note that $U = ee^T$). Note that X_k is a matrix whose (j,i) entry is $x_k(j,i)$. Similar expressions can be obtained for (10)–(12), with z_k replacing y_k, S_k replacing M_k, and T_k replacing N_k. By letting $A_k := M_k - N_{k-1}$, iteration (13)–(15) can be rewritten as

$$A_k = A_{k-1} \circ (U - X_{k-1}) + P\text{diag}(y_k), \quad k \geq 1, \tag{16}$$

$$y_{k+1} = (A_k \circ X_k)e, \quad k \geq 0. \tag{17}$$

Similarly, we can write iteration (10)–(12) with $B_k := S_k - T_{k-1}$ and z_k replacing A_k and y_k respectively.

For analysis purposes, each matrix in (16)–(17) will be rewritten in vector form by stacking up the corresponding columns.[1] Let $F = [I_n \ I_n \ \ldots \ I_n] \in \mathbb{R}^{n \times n^2}$, where I_n is the $n \times n$ identity matrix, and $\tilde{P} = [E_1 P^T \ E_2 P^T \ \ldots \ E_n P^T]^T \in \mathbb{R}^{n^2 \times n}$, where $E_i \in \mathbb{R}^{n \times n}$ has $E_i(i,i) = 1$ and all other entries equal zero. [The entries of $E_i P^T \in \mathbb{R}^{n \times n}$ ($PE_i^T = PE_i$) are all zero except for the i^{th} row (column) entries, which are those of the i^{th} row (column) of matrix P^T (P).] Then, (16)–(17) can be rewritten as

$$a_k = a_{k-1} \circ (u - x_{k-1}) + \tilde{P} y_k, \quad k \geq 1, \tag{18}$$

$$y_{k+1} = F(a_k \circ x_k), \quad k \geq 0, \tag{19}$$

where $a_k \in \mathbb{R}^{n^2}$, $x_k \in \mathbb{R}^{n^2}$, and $x_{k-1} \in \mathbb{R}^{n^2}$ result from stacking the columns of matrices A_k, X_k, and X_{k-1}, respectively. Similarly, for the second iteration, we can write

$$b_k = b_{k-1} \circ (u - x_{k-1}) + \tilde{P} z_k, \quad k \geq 1, \tag{20}$$

$$z_{k+1} = F(b_k \circ x_k), \quad k \geq 0, \tag{21}$$

where $b_k \in \mathbb{R}^{n^2}$ results from stacking the columns of B_k. Note that the (j,i) entry of matrix A_k (and B_k) and their corresponding entry in a_k (and b_k) remain zero if there is no communication link from node i to node j, i.e., $(j,i) \notin \mathcal{E}$.

IV. Main Results and Implications

We shall argue that with the resilient ratio-consensus algorithm described in Section III, and despite the presence of unreliable communication links that fail at each time step with (unequal) probabilities (independently from other links and between time steps), nodes can asymptotically reach average consensus. More specifically, we will argue that

$$\lim_{k \to \infty} \frac{y_k(j)}{z_k(j)} = \overline{y},\ \forall j \in \mathcal{V}. \tag{22}$$

with probability one. [Note that $z_k(j) > 0$ for all k since all $z_k(0) = 1$ and self-loops are reliable (i.e., $q_{jj} = 1$ for all j).]

To this end, let $C = \big[\tilde{P}F\text{diag}(q) + (I - \text{diag}(q))\big]$, and $D = \big[I - \tilde{P}F\big]$, where I is the $n^2 \times n^2$ identity matrix, and

$$\Pi = C \otimes C + \big\{[D\text{diag}(q)] \otimes \big[D\big(I - \text{diag}(q)\big)\big]\big\}\, G, \tag{23}$$

where $\otimes$ denotes the Kronecker product, and G is an $n^2 \times n^2$ diagonal matrix with entries $G\big((l-1)n^2 + l, (l-1)n^2 + l\big) = 1,\ \forall l = 1, 2, \ldots, n^2$, and zero otherwise. The following theorem and lemma establish the main convergence results; the implications of these results are illustrated next.

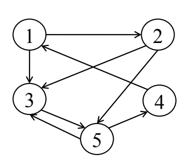
Theorem 1: Let y_k and z_k be the random vectors that result from iterations (18)–(19) and (20)–(21), and define $v_k = y_k - \overline{y} z_k$, where $\overline{y} = \frac{\sum_{l=1}^n y_0(l)}{\sum_{l=1}^n z_0(l)} = \frac{\sum_{l=1}^n y_0(l)}{n}$. If the underlying graph $\mathcal{G}$ is strongly connected and $q_{ji} > 0$ for all $(j,i) \in \mathcal{E}$, then $\|v_k\|_\infty \to 0$ almost surely.

Lemma 1: Let λ_2 denote the eigenvalue of Π with the second largest magnitude, then $\Pr\{\|v_{k+1}\|_\infty > \epsilon\} \leq C'(\|\mathbf{E}[w_0 w_0^T]\|_\infty)^{1/2} k^{\frac{m_2 - 1}{2}} |\lambda_2|^{k/2}$, where m_2 is the algebraic multiplicity of eigenvalue λ_2 and C' is some constant that depends on matrix Π, n, and ϵ.

Example 1: Consider the directed graph in Fig. 1 (where self-loops are not drawn) and set the initial values of the five nodes to be $y_0 = [4, 5, 6, 3, 2]^T$ (with average $\overline{y} = 4$). Assuming no packet drops (perfectly reliable links), we run the iterations in (1)–(2) and plot in Fig. 2(a) the ratio $\frac{y_k(j)}{z_k(j)}$ as a function of the number of iterations k for each node j. We observe that all ratios indeed converge to $\overline{y} = 4$.

Now, consider the case in which all links in the graph in Fig. 1 suffer packet drops with equal probability, independently between links and between time steps, i.e., we take $q_{ji} = 0.8 \equiv q$ for all $(j,i) \in \mathcal{E}$, $i \neq j$, and $q_{jj} = 1$ for all j. In Fig. 2(b) we plot the ratio of $\frac{y_k(j)}{z_k(j)}$ as a function of k; we observe that the ratios indeed converge to $\overline{y} = 4$, but this convergence takes longer than in the case of reliable links.

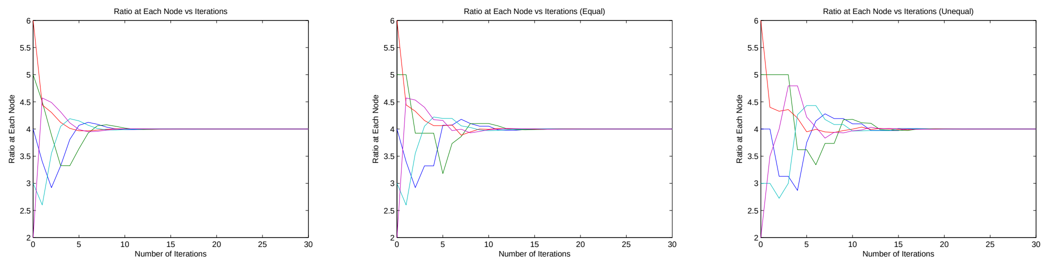
Finally, we consider the case when the probabilities of a packet drop in each link are given by the matrix Q defined in Fig. 1. In Fig. 2(c) we plot the ratio of $\frac{y_k(j)}{z_k(j)}$ as a function of the number of iterations k for the iterations in (7)–(12). The ratios indeed converge to $\overline{y} = 4$, but it takes longer than in the case of reliable links and also appears to take longer than the equal-probability case. Note that for the Q in Fig. 1, the average probability of successful transmission is 0.8. $\Box$



$$Q = \begin{bmatrix} 1 & 0 & 0 & 0.9 & 0 \\ 0.9 & 1 & 0 & 0 & 0 \\ 0.7 & 0.85 & 1 & 0 & 0.6 \\ 0 & 0 & 0 & 1 & 0.9 \\ 0 & 0.85 & 0.7 & 0 & 1 \end{bmatrix}$$

Fig. 1. Small directed graph used for illustrating the resilient algorithm, and the matrix Q that is used in the case when the probabilities of successful transmission for each link are generally unequal.

[1] Vectors defined by stacking the columns of a matrix will be denoted with the same small letter as the capital letter of the corresponding matrix.



(a) $q_{ji} = 1,\ \forall j, i.$ (b) $q_{ji} = 0.8,\ \forall j, i,\ j \neq i.$ (c) Unequally likely packet drops.

Fig. 2. Five-node example: ratio $\frac{y_k(j)}{z_k(j)}$ for each node j, as a function of the number of iterations k and different link failure probability scenarios.

V. Convergence Analysis

First, we characterize the first and second moment of the iterations in (18)–(19), and (20)–(21). The results are then used to sketch the proofs of Theorem 1 and Corollary 1.

A. First Moment Analysis

Let $\overline{a}_k := \mathbf{E}[a_k]$, $\overline{y}_k := \mathbf{E}[y_k]$, $\overline{b}_k := \mathbf{E}[b_k]$, and $\overline{z}_k := \mathbf{E}[z_k]$. We show that $\overline{a}_k$ and $\overline{b}_k$ (and also $\overline{y}_k$ and $\overline{z}_k$) converge to identical vectors up to scalar multiplication. Note that q is a vector that results from stacking up the columns of Q.

Lemma 2: The evolution of $\overline{a}_k$, $\overline{y}_k$, $\overline{b}_k$, and $\overline{z}_k$, $\forall k \geq 1$, is governed by

$$\overline{a}_k = \left[\tilde{P}F\text{diag}(q) + (I_{n^2} - \text{diag}(q))\right]\overline{a}_{k-1}, \quad (24)$$

$$\overline{y}_{k+1} = F\text{diag}(q)\overline{a}_k, \quad (25)$$

$$\overline{b}_k = \left[\tilde{P}F\text{diag}(q) + (I_{n^2} - \text{diag}(q))\right]\overline{b}_{k-1}, \quad (26)$$

$$\overline{z}_{k+1} = F\text{diag}(q)\overline{b}_k, \quad (27)$$

where I_m is the $m \times m$ identity matrix and diag(q) is a diagonal matrix with the entries of vector q defining its diagonal. The initial conditions of the iterations are $\overline{a}_0 = \tilde{P}y_0$, $\overline{y}_1 = F\text{diag}(q)\tilde{P}y_0$, $\overline{b}_0 = \tilde{P}z_0$, and $\overline{z}_1 = F\text{diag}(q)\tilde{P}z_0$.

Proof: Since the development for obtaining $\overline{a}_k$ and $\overline{y}_k$ is parallel to that for obtaining $\overline{b}_k$ and $\overline{z}_k$, our analysis focuses on the first case. For $k = 0$ in (18)–(19), by taking expectations of both sides and noting that packet drops at time step $k = 0$ are independent of the initial values for a_0, it follows that $\overline{a}_0 = \tilde{P}y_0$, and $\overline{y}_1 = F\text{diag}(q)\overline{a}_0$; therefore, $\overline{y}_1 = F\text{diag}(q)\overline{a}_0 = F\text{diag}(q)\tilde{P}\overline{y}_0$. For $k \geq 1$ in (18)–(19), by taking expectations on both sides and noting that packet drops at time step k are independent of previous packet drops and the initial values of a_0, we obtain

$$\begin{aligned} \overline{a}_k &= \overline{a}_{k-1} \circ (u - \overline{x}_{k-1}) + \tilde{P}\overline{y}_k, \\ &= (I_{n^2} - \text{diag}(q))\overline{a}_{k-1} + \tilde{P}\overline{y}_k, \quad (28) \\ \overline{y}_{k+1} &= F(\overline{a}_k \circ \overline{x}_k), \\ &= F\text{diag}(q)\overline{a}_k. \quad (29) \end{aligned}$$

Substituting (29) into (28), we complete the proof. ∎

Lemma 3: The first moments of a_k and b_k (also y_k and z_k) asymptotically converge to the same solution up to scalar multiplication. Specifically, $\lim_{k\to\infty} \overline{a}_k = \overline{y} \lim_{k\to\infty} \overline{b}_k$, and $\lim_{k\to\infty} \overline{y}_k = \overline{y} \lim_{k\to\infty} \overline{z}_k$, where $\overline{y} = \frac{\sum_{l=1}^{n} y_0(l)}{n}$.

Proof: Since P is column stochastic, it follows that $[\tilde{P}F\text{diag}(q) + (I_{n^2} - \text{diag}(q))]$ is also column stochastic (because $0 \leq \text{diag}(q) \leq 1$). In fact, since P is also primitive, it corresponds to an underlying graph $\mathcal{G}$ that is strongly connected. One can easily establish that $[\tilde{P}F\text{diag}(q) + (I_{n^2} - \text{diag}(q))]$ will have a single (reachable) recurrent class as long as each of the edges of the underlying strongly connected graph admit a nonzero probability of transmission (i.e., for each $(j, i) \in \mathcal{E}$, $q_{ji} > 0$ which implies that q will have a positive entry at its $((i-1)n + j)^{th}$ position). The proof is omitted due to space limitations.

From the fact that $[\tilde{P}F\text{diag}(q) + (I_{n^2} - \text{diag}(q))]$ is column stochastic and has a single recurrent class, we know that the solutions of (24) and (26) are unique up to scalar multiplication, i.e., $\lim_{k\to\infty} \overline{a}_k = \alpha \lim_{k\to\infty} \overline{b}_k$, for some α. Then, from the column stochasticity property: $\sum_{j=1}^{n^2} \overline{b}_k(j) = \sum_{j=1}^{n^2} \overline{b}_0(j) = \sum_{j=1}^{n} z_0(j) = n$ and $\sum_{j=1}^{n^2} \overline{a}_k(j) = \sum_{j=1}^{n^2} \overline{a}_0(j) = \sum_{j=1}^{n} y_0(j)$, $\forall k \geq 1$; this implies that $\alpha = \overline{y}$. It immediately follows that $\lim_{k\to\infty} \overline{y}_k = \overline{y} \lim_{k\to\infty} \overline{z}_k$. ∎

B. Second Moment Analysis

Next, we establish that the evolution of $\Gamma_k := \mathbf{E}[a_k a_k^T]$, $\Psi_k := \mathbf{E}[b_k b_k^T]$, $\Xi_k := \mathbf{E}[a_k b_k^T]$, $\Phi_k := \mathbf{E}[y_k y_k^T]$, $\Lambda_k := \mathbf{E}[z_k z_k^T]$, and $\Upsilon_k := \mathbf{E}[y_k z_k^T]$ can be expressed as linear iterations with identical dynamics but different initial conditions. We additionally show that the steady-state solutions of Γ_k, Ψ_k, and Ξ_k (and also Φ_k, Λ_k, and Υ_k) are identical up to scalar multiplication.

Lemma 4: Let x, c and d be random vectors of dimension m. Furthermore, assume that the entries of x are independent (not necessarily identical) Bernoulli random variables such that $\Pr\{x_i = 1\} = q_i$ and $\Pr\{x_i = 0\} = 1 - q_i$, $\forall i = 1, 2, \ldots m$, and are independent from c and d. Then

$$\begin{aligned} S = \mathbf{E}\left[(c \circ x)(x \circ d)^T\right] &= \text{diag}(q)\,\mathbf{E}[cd^T]\text{diag}(q) + \\ &+ \text{diag}(q)\text{diag}(\mathbf{E}\left[cd^T\right])(I_m - \text{diag}(q)), \\ T = \mathbf{E}\left[(c \circ x)\big((u - x) \circ d\big)^T\right] & \\ = \text{diag}(q)\,\mathbf{E}[cd^T](I_m - \text{diag}(q)) &- \\ - \text{diag}(q)\text{diag}(\mathbf{E}\left[cd^T\right])(I_m - \text{diag}(q)), & \end{aligned}$$

where $\text{diag}(\mathbf{E}\left[cd^T\right])$ is a diagonal matrix with the same diagonal as $\mathbf{E}\left[cd^T\right]$.

Proof: The $(i, j), i \neq j$, entry of S can be obtained as $S_{ij} = \mathbf{E}\left[c_i x_i d_j x_j\right]$ and, since x_i and x_j are pairwise independent, and independent from c and d, it follows that $\mathbf{E}\left[c_i x_i d_j x_j\right] = q_i\,\mathbf{E}\left[c_i d_j\right] q_j$. For $i = j$, observing that $\mathbf{E}[x_i x_i] = \mathbf{E}[x_i] = q_i$, $\forall i = 1, \ldots, m$, we obtain that

$S_{ii} = \mathbf{E}\left[c_i x_i d_i x_i\right] = \mathbf{E}\left[c_i d_i x_i\right] = q_i\,\mathbf{E}\left[c_i d_i\right]$. Combining these two facts, the expression for S follows. The expression for T can be similarly established. ■

Lemma 5: The evolutions of Γ_k, Φ_k, Ψ_k, Λ_k, Ξ_k, Υ_k, $\forall k \geq 1$, are described by the following iterations:

$$\begin{aligned}\Gamma_k =& C\Gamma_{k-1}C^T + \\ &+ D\text{diag}(q)\text{diag}(\Gamma_{k-1})(I - \text{diag}(q))D^T, \quad (30)\\ \Phi_{k+1} =& F\text{diag}(q)\Gamma_k\text{diag}(q)F^T + \\ &+ F\text{diag}(q)\text{diag}(\Gamma_k)(I - \text{diag}(q))F^T, \quad (31)\\ \Psi_k =& C\Psi_{k-1}C^T + \\ &+ D\text{diag}(q)\text{diag}(\Psi_{k-1})(I - \text{diag}(q))D^T, \quad (32)\end{aligned}$$

$$\begin{aligned}\Lambda_{k+1} =& F\text{diag}(q)\Psi_k\text{diag}(q)F^T + \\ &+ D\text{diag}(q)\text{diag}(\Psi_k)(I - \text{diag}(q))D^T, \quad (33)\\ \Xi_k =& C\Xi_{k-1}C^T + \\ &+ D\text{diag}(q)\text{diag}(\Xi_{k-1})(I - \text{diag}(q))D^T, \quad (34)\\ \Upsilon_{k+1} =& F\text{diag}(q)\Xi_k\text{diag}(q)F^T + \\ &+ D\text{diag}(q)\text{diag}(\Xi_k)(I - \text{diag}(q))D^T, \quad (35)\end{aligned}$$

with initial conditions

$$\begin{aligned}\Gamma_0 &= \tilde{P}y_0y_0^T\tilde{P}^T,\\ \Phi_1 &= F\text{diag}(q)\tilde{P}y_0y_0^T\tilde{P}^T\text{diag}(q)F^T + \\ &\quad +F\text{diag}(q)\text{diag}(\tilde{P}y_0y_0^T\tilde{P}^T)(I - \text{diag}(q))F^T,\\ \Psi_0 &= \tilde{P}z_0z_0^T\tilde{P}^T,\\ \Lambda_1 &= F\text{diag}(q)\tilde{P}z_0z_0^T\tilde{P}^T\text{diag}(q)F^T + \\ &\quad +F\text{diag}(q)\text{diag}(\tilde{P}z_0z_0^T\tilde{P}^T)(I - \text{diag}(q))F^T,\\ \Xi_0 &= \tilde{P}y_0z_0^T\tilde{P}^T,\\ \Upsilon_1 &= F\text{diag}(q)\tilde{P}y_0z_0^T\tilde{P}^T\text{diag}(q)F^T + \\ &\quad +F\text{diag}(q)\text{diag}(\tilde{P}z_0z_0^T\tilde{P}^T)(I - \text{diag}(q))F^T.\end{aligned}$$

Proof: For $k = 0$, it follows from Lemma 2 and (18) that $a_0 = \tilde{P}y_0$. Then, $\Gamma_0 = \mathbf{E}[a_0a_0^T] = \tilde{P}\,\mathbf{E}[y_0y_0^T]\tilde{P}^T = \tilde{P}y_0y_0^T\tilde{P}^T$, $\Phi_1 = \mathbf{E}[y_1y_1^T] = \mathbf{E}\left[F(a_0\circ x_0)(x_0\circ a_0)^TF^T\right] = F\,\mathbf{E}\left[(a_0\circ x_0)(x_0\circ a_0)^T\right]F^T$, and applying the results in Lemmas 2 and 4, it follows that

$$\begin{aligned}\Phi_1 &= F\text{diag}(q)\,\mathbf{E}\left[a_0a_0^T\right]\text{diag}(q)F^T + \quad (36)\\ &\quad + F\text{diag}(q)\text{diag}(\mathbf{E}\left[a_0a_0^T\right])(I - \text{diag}(q))F^T\\ &= F\text{diag}(q)\tilde{P}y_0y_0^T\tilde{P}^T\text{diag}(q)F^T + \quad (37)\\ &\quad + F\text{diag}(q)\text{diag}(\tilde{P}y_0y_0^T\tilde{P}^T)(I - \text{diag}(q))F^T\ .\end{aligned}$$

Taking into account that $y_k = F(a_{k-1}\circ x_{k-1})$, it follows that for $k \geq 1$ we have

$$\begin{aligned}\Gamma_k =& \mathbf{E}\left[\left(a_{k-1}\circ(u - x_{k-1}) + \tilde{P}y_k\right)\right.\\ &\left.\left(a_{k-1}\circ(u - x_{k-1}) + \tilde{P}y_k\right)^T\right]\\ =& \mathbf{E}\left[\left(a_{k-1}\circ(u - x_{k-1})\right)\left(a_{k-1}\circ(u - x_{k-1})\right)^T\right]\\ &+ \mathbf{E}\left[\left(a_{k-1}\circ(u - x_{k-1})\right)\left(a_{k-1}\circ x_{k-1}\right)^T\right]F^T\tilde{P}^T\\ &+ \tilde{P}F\,\mathbf{E}\left[\left(a_{k-1}\circ x_{k-1}\right)\left(a_{k-1}\circ(u - x_{k-1})\right)^T\right]\\ &+ \tilde{P}F\,\mathbf{E}\left[\left(a_{k-1}\circ x_{k-1}\right)\left(a_{k-1}\circ x_{k-1}\right)^T\right]F^T\tilde{P}^T;\end{aligned}$$

then, by applying Lemma 4 four times and rearranging terms, we write the expression above as in (30). Using $y_{k+1} = F(a_k\circ x_k)$ and applying Lemma 4 once more, (31) results.

The expressions for Ψ_k, Λ_{k+1}, Ξ_k, and Υ_{k+1} can be derived in a similar fashion and are omitted for brevity. ■

Remark 1: Although omitted in the statement of Lemma 5, $\Delta_k = \mathbf{E}[b_ka_k^T]$ and $\Theta_k = \mathbf{E}[z_ky_k^T]$ can be easily obtained by noting that $\Delta_k = \Psi_k^T$ and $\Theta_k = \Upsilon_k^T$. □

Next, we show that the steady-state solutions of Γ_k, Ψ_k, Ξ_k and Δ_k are identical up to scalar multiplication. To see this, we will rewrite (30), (32), and (34) in vector form using Kronecker products. For matrices C, A, and B of appropriate dimensions, the matrix equation $C = AXB$ can be rewritten as $(B^T\otimes A)x = c$, where x and c are the vectors that result from stacking the columns of X and C respectively, and $\otimes$ denotes the Kronecker product [10].

Let γ_k be the vector that results from stacking the columns of Γ_k, and let Π be the matrix defined in (23). Then, using the ideas above, we can rewrite (30) as

$$\gamma_k = \Pi\gamma_{k-1},\ k \geq 1. \quad (38)$$

If we let ψ_k, ξ_k, δ_k be the vectors that result from stacking the columns of Ψ_k, Ξ_k and Δ_k respectively, then it also easy to see that the same recurrence relation as in (38) governs the evolution of ψ_k, ξ_k and δ_k.

The structure and fundamental properties of the matrix Π are established in the next theorem (the proof is omitted due to space limitations), from where it follows that the steady-state solutions of γ_k, ψ_k, ξ_k and δ_k (and therefore Γ_k, Ψ_k, Ξ_k and Δ_k) are identical up to scalar multiplication.

Theorem 2: Let $P \in \mathbb{R}^{n\times n}$ be a column stochastic and primitive weight matrix associated with a directed graph $\mathcal{G} = \{\mathcal{V},\mathcal{E}\}$, with $\mathcal{V} = \{1,2,\ldots,n\}$ and $\mathcal{E} \subseteq \mathcal{V}\times\mathcal{V}$. Let $F = [I_n\ I_n\ \ldots\ I_n] \in \mathbb{R}^{n\times n^2}$, where I_n is the $n\times n$ identity matrix, and $\tilde{P} = [E_1P^T\ E_2P^T\ \ldots\ E_nP^T]^T \in \mathbb{R}^{n^2\times n}$, where each $E_i \in \mathbb{R}^{n\times n},\ i \in \{1,2,\ldots,n\}$, satisfies $E_i(i,i) = 1$ and has all other entries equal to zero. Then, for any[2] vector q, $0 < q \leq 1$, the matrix Π that defines (38) is column stochastic, and it has a single eigenvalue of maximum magnitude at value one.

The next two lemmas establish that the first and second moments of a_k and b_k (also y_k and z_k) converge to the same solution up to a scalar multiplication. These lemmas are used in Section V-C to show that as $k\to\infty$, the random vector $v_k = y_k - \overline{y}z_k$ converges to $v = 0$ almost surely. Thus, as $k\to\infty$ (and since $z_k(j) > 0$), each node j can obtain an estimate of $\overline{y}$ by calculating $y_k(j)/z_k(j)$.

Lemma 6: Define $w_k = a_k - \overline{y}b_k$ and denote by χ_k the vector that results from stacking the columns of $\mathcal{X}_k := \mathbf{E}[w_kw_k^T]$. Then, it follows that $\chi_k = \Pi\chi_{k-1}$ with $\chi_0 = \gamma_0 + \overline{y}^2\psi_0 - \overline{y}(\xi_0 + \delta_0)$ and $\sum_{l=1}^{n^4}\chi_0(l) = 0$.

[2]Taking the vector q to be strictly positive implies that all $q_{ji}, i,j \in \{1,2,...,n\}$, need to be strictly positive, including pairs $(j,i)\notin\mathcal{E}$. It can be shown that this can be done without loss of generality because non-existing links are excluded from the structure of matrix P (and matrix $\tilde{P}$); in other words, the fact that their corresponding q's have nonzero values has no effect on the iterations.

Proof: Since $\mathcal{X}_k := \mathbf{E}[w_k w_k^T] = \mathbf{E}[a_k a_k^T] + \overline{y}^2\,\mathbf{E}[b_k b_k^T] - \overline{y}(\mathbf{E}[a_k b_k^T] + \mathbf{E}[b_k a_k^T]) = \Gamma_k + \overline{y}^2\Psi_k - \overline{y}(\Xi_k + \Delta_k)$, it follows that $\chi_k = \gamma_k + \overline{y}^2\psi_k - \overline{y}(\xi_k + \delta_k)$. From (38) and subsequent discussion, it follows that $\gamma_k = \Pi\gamma_{k-1}$, $\psi_k = \Pi\psi_{k-1}$, $\xi_k = \Pi\xi_{k-1}$, and $\delta_k = \Pi\delta_{k-1}$, thus $\chi_k = \Pi\gamma_{k-1} + \overline{y}^2\Pi\psi_{k-1} - \overline{y}(\Pi\xi_{k-1} + \Pi\delta_{k-1}) = \Pi(\gamma_{k-1} + \overline{y}^2\psi_{k-1} - \overline{y}(\xi_{k-1} + \delta_{k-1})) = \Pi\chi_{k-1}$.

Now, in Lemma 5, it was shown that $\Gamma_0 = \tilde{P} y_0 y_0^T \tilde{P}^T$, $\Psi_0 = \tilde{P} z_0 z_0^T \tilde{P}^T$, and $\Xi_0 = \tilde{P} y_0 z_0^T \tilde{P}^T = \Delta_0$. Since γ_0, ψ_0, ξ_0, and δ_0 result from stacking the columns of Γ_0, Ψ_0, Ξ_0, and Δ_0, it follows that

$$\sum_{l=1}^{n^4} \gamma_0(l) = \sum_{i=1}^{n^2}\sum_{j=1}^{n^2} \Gamma_0(i,j) = \left(\sum_{i=1}^{n} y_0(i)\right)^2,$$
$$\sum_{l=1}^{n^4} \psi_0(l) = \sum_{i=1}^{n^2}\sum_{j=1}^{n^2} \Psi_0(i,j) = \left(\sum_{i=1}^{n} z_0(i)\right)^2,$$
$$\sum_{l=1}^{n^4} \xi_0(l) = \left(\sum_{i=1}^{n} y_0(i)\right)\left(\sum_{i=1}^{n} z_0(i)\right),$$
$$\sum_{l=1}^{n^4} \delta_0(l) = \left(\sum_{i=1}^{n} z_0(i)\right)\left(\sum_{i=1}^{n} y_0(i)\right),$$

where the last equality in each of the above expressions is obtained by taking into account that i) matrix $\tilde{P}$ is column stochastic by construction, and ii) for any $a, b \in \mathbb{R}^n$, we have that $\sum_{i=1}^{n}\sum_{j=1}^{n}(ab^T)(i,j) = (\sum_{l=1}^{n} a_l)(\sum_{l=1}^{n} b_l)$. Since $\overline{y} = \frac{\sum_{j=1}^{n} y_0(j)}{\sum_{j=1}^{n} z_0(j)}$, it follows that $\sum_{l=1}^{n^4}\chi_0(l) = \sum_{l=1}^{n^4}(\gamma_0(l) + \overline{y}^2\psi_0(l) - \overline{y}(\xi_0(l) + \delta_0(l))) = 0$. ■

C. Convergence of the Resilient Ratio-Consensus Algorithm

Here, we only sketch the proofs for Theorem 1 and Corollary 1 (stated in Section IV); these proofs are similar to the corresponding ones in [3], and therefore are omitted.

Theorem 1 establishes that, in the limit as the number of iterations k becomes large, the values of vectors y_k and z_k will be perfectly aligned so that $v_k = y_k - \overline{y} z_k = 0$ with probability one. Thus, in the limit, each node j can calculate the value of $\overline{y}$ by obtaining $\frac{y_k(j)}{z_k(j)}$, *as long as* $z_k(j) \neq 0$.

Theorem 1 proof sketch: The result follows from the first Borel-Cantelli Lemma. For all $k \geq 0$, and all $\epsilon > 0$, the key is to upper bound $\sum_{k=0}^{\infty}\Pr\{\|v_k\|_\infty > \epsilon\}$ by $\frac{1}{\epsilon}\sum_{k=0}^{\infty}\mathbf{E}\left[\|v_k\|_2\right]$ and then establish that $\mathbf{E}\left[\|v_k\|_2\right] \to 0$ as $k \to \infty$ geometrically fast. To this end, we can show by Lemma 5 that $\mathbf{E}[v_k v_k^T]$ can be written as a function of $\mathcal{X}_{k-1} = \mathbf{E}[w_{k-1} w_{k-1}^T]$ as defined in Lemma 6. Thus, the evolution of $\mathbf{E}[v_k v_k^T]$ is governed by the evolution of $\mathcal{X}_{k-1}$ or by χ_{k-1} (the vector that results from stacking the columns of $\mathcal{X}_{k-1}$). In Theorem 2, we showed that Π is column stochastic and has a unique eigenvector (with all entries strictly positive) associated to the largest eigenvalue $\lambda_1 = 1$. Then, the solution of $\chi_k = \Pi\chi_{k-1}$ is unique and equal to this eigenvector (up to scalar multiplication); but Lemma 6 established that $\sum_{l=1}^{n^4}\chi_0(l) = 0$, therefore $\lim_{k\to\infty}\chi_k(l) = 0$, $\forall l$. Additionally, the convergence of $\chi_k = \Pi\chi_{k-1}$ is geometric with a rate of convergence given by $|\lambda_2|$ where λ_2 is the eigenvalue of Π of second largest modulus, which satisfies $|\lambda_2| < \lambda_1 = 1$ [11]. Some additional manipulations of $\mathbf{E}[v_k v_k^T]$ lead to the result. □

Lemma 1 establishes the number of iterations k after which y_k and z_k will satisfy $|y_k - \overline{y} z_k| \leq \epsilon$, for a given accuracy level ϵ, with some desired probability. This probability goes to 1 with a geometric rate governed by $|\lambda_2|^{1/2}$, where λ_2 is the eigenvalue of Π of second largest modulus.

Lemma 1 proof sketch: It is well-known (see, e.g., [10, Thm. 8.5.1]) that $\|\Pi^k - L\|_\infty \leq C k^{m_2-1}|\lambda_2|^k$, for some constant $C = C(\Pi)$, where $L = \lim_{k\to\infty}\Pi^k$ is a rank-one column stochastic matrix. It then follows that $\|(\Pi^k - L)\chi_0\|_\infty \leq \|\Pi^k - L\|_\infty\|\chi_0\|_\infty \leq C\|\chi_0\|_\infty k^{m_2-1}|\lambda_2|^k$, but since $\sum_{l=1}^{n^4}\chi_0(l) = 0$, we have that $L\chi_0 = 0$, and $\|\,\mathbf{E}[v_{k+1}v_{k+1}^T]\|_\infty \leq \|\,\mathbf{E}[w_k w_k^T]\|_\infty \leq C\|\chi_0\|_\infty|\lambda_2|^k$. After realizing that $\|\,\mathbf{E}[v_{k+1}v_{k+1}^T]\|_\infty \geq \frac{1}{n^2}(\mathbf{E}[\|v_{k+1}\|_\infty])^2$, the result follows from some additional manipulations.

Remark 2: In our recent work [12], we have followed an alternative approach to establish convergence of the ratio-consensus algorithm that involves the use of coefficients of ergodicity used in the analysis of non-homogeneous Markov chains (see, e.g., [11]). This approach involves rewriting (13)–(15) slightly differently, by defining $A_k := M_k - N_k$ $\big($instead of $A_k := M_k - N_{k-1}$ as in (16)–(17)$\big)$. □

VI. Concluding Remarks

In this paper, we proposed and analyzed a method to ensure resiliency of a linear-iterative distributed algorithm for average consensus against unreliable heterogeneous communication links that may drop packets with generally unequal probabilities. Asymptotic convergence of all nodes to average consensus with probability one was established and a bound on the rate of convergence was also obtained. Future work will characterize convergence in the presence of unreliable heterogenous communication links that can independently drop packets following individual finite state Markov models.

References

[1] J. Tsitsiklis, “Problems in decentralized decision making and computation,” Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge, MA, 1984.

[2] R. Olfati-Saber, J. Fax, and R. Murray, “Consensus and cooperation in networked multi-agent systems,” *Proc. of the IEEE*, vol. 95, no. 1, pp. 215–233, Jan. 2007.

[3] A. D. Domínguez-García, C. N. Hadjicostis, and N. Vaidya, “Resilient networked control of distributed energy resources,” *IEEE Journal on Selected Areas in Comm.: Smart Grid Communications Series*, vol. 30, no. 6, pp. 1137–1148, Jul. 2012.

[4] A. D. Domínguez-García and C. N. Hadjicostis, “Coordination and control of distributed energy resources for provision of ancillary services,” in *Proc. IEEE SmartGridComm*, Oct. 2010.

[5] A. D. Domínguez-García and C. N. Hadjicostis, “Distributed algorithms for control of demand response and distributed energy resources,” in *Proc. IEEE Conf. on Decision and Control*, 2011.

[6] F. Benezit *et al.*, “Weighted gossip: Distributed averaging using non-doubly stochastic matrices,” in *Proc. IEEE International Symposium on Information Theory*, Jun. 2010.

[7] S. Patterson, B. Bamieh, and A. El Abbadi, “Distributed average consensus with stochastic communication failures,” in *Proc. IEEE Conf. on Decision and Control*, Dec. 2007.

[8] F. Fagnani and S. Zampieri, “Average consensus with packet drop communication,” *Siam Journal on Control and Optimization*, vol. 48, no. 1, pp. 102–133, Feb. 2009.

[9] Y. Chen, R. Tron, A. Terzis, and R. Vidal, “Corrective consensus: Converging to the exact average,” in *Proc. IEEE Conf. on Decision and Control*, Dec. 2010.

[10] R. Horn and C. Johnson, *Matrix Analysis*. New York, NY: Cambridge University Press, 1985.

[11] E. Seneta, *Non-negative Matrices and Markov Chains*, revised printing ed. New York, NY: Springer, 2006.

[12] N. H. Vaidya, C. N. Hadjicostis, and A. D. Domínguez-García, “Robust consensus over unreliable links: Analysis using coefficients of ergodicity,” in *Proc. IEEE Conf. on Decision and Control*, 2012.