

Resilient Networked Control of Distributed Energy Resources

Alejandro D. Dominguez-Garcia, *Member, IEEE*, Christoforos N.

Hadjicostis, *Senior Member, IEEE*, and

Nitin H. Vaidya, *Fellow, IEEE*

Abstract

This paper considers networked systems and develops a class of distributed algorithms that are resilient against potential packet drops in the communication links between system components. We apply this class of algorithms to the problem of coordinating distributed energy resources (DERs) for service provision to the grid they are connected to, e.g., reactive power support for voltage control. In this problem, each system component can contribute a certain amount of active and/or reactive power, bounded from above and (possibly) below by capacity constraints, and the objective is to coordinate the components so as to collectively provide a predetermined total amount of active and/or reactive power. In the class of algorithms proposed to address this problem, each DER maintains a set of variables and updates them through information exchange with neighboring DERs. We show that, as long as the underlying graph that describes the information exchange between components is strongly connected, and the predetermined total amount of active and/or reactive power does not violate (upper or lower) total capacity constraints, DERs can use this approach to calculate, in a distributed fashion, their fair contribution (subject to their capacity constraints). We show that algorithms reach almost surely convergence to the fair solution, even in the presence of communication link failures.

A. D. Dominguez-Garcia 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. 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.

The work of A. D. Domínguez-García was supported in part by NSF under Career Award ECCS-CAR-0954420. The work of C. N. Hadjicostis was supported in part by the European Commission (EC) Seventh Framework Programme (FP7/2007-2013) under grant agreements INFOS-ICT-223844 and PIRG02-GA-2007-224877. The work of N. H. Vaidya was supported in part by Army Research Office grant W-911-NF-0710287 and NSF award 1059540. Any opinions, findings, and conclusions or recommendations expressed here are those of the authors and do not necessarily reflect the views of the funding agencies, the U.S. government, or the European Commission.

I. INTRODUCTION

Driven by the US-DoE SmartGrid initiative and its European counterpart, electrical energy systems are undergoing radical transformations in structure and functionality in a quest to increase efficiency and reliability. Focusing on the distribution level, proper coordination and control of generation and storage resources [from now on we will refer to them as distributed energy resources (DERs)] provides more flexibility in the provision of ancillary services, which can result in enhanced efficiency and reliability. For example, while the primary function of power electronics interfaces in solar rooftop installations is to control active power, when properly controlled, they can also provide reactive power support for voltage control [1], [2].

In the above scenario, there are several ways in which DERs can be coordinated. A solution can be achieved through a completely centralized control strategy where some decision maker—the *coordinator*—issues a (separate) command to each DER with a request for a certain amount of active or reactive power, so that the DERs collectively provide the total amount of active or reactive power needed. In this centralized solution, however, it is necessary to overlay a communication network connecting the coordinator with each DER, and to maintain knowledge of the DERs that are available and their constraints at any given time. An alternative approach consists of a completely distributed control strategy, in which the coordinator initially relays a request to a limited number of DERs (with which the coordinator can directly communicate) without necessarily knowing the total number of DERs, their individual capacity limitations, or the connections between them). Through a distributed computation over the network, the DERs can determine their own contribution so that collectively they have the same effect as the centralized strategy. In this paper, we pursue this distributed approach.

We consider a network of DERs (also referred to as nodes), each of which can provide a certain amount of active and/or reactive power (bounded from above and from below by upper and lower limits). In our development, we adopt a very general model for the communication modality between nodes, which allows asymmetric information structures, in the sense that if node i can transmit information to another node j , it is not necessarily true that node j can transmit information to node i ; this situation can arise nontrivially in many realistic settings (e.g., when nodes transmit at different power levels or are subject to different degrees of interference in a wireless environment). We only require that each node, apart from seeing

incoming transmissions sent to it by neighboring nodes, also knows the number of nodes that it can transmit information to, which in graph-theoretic terms is referred to as the out-degree of that node. Through a distributed computation over the network, DERs determine (using only information made available by their neighbors) the amount of active and/or reactive power they will provide. The resulting solution should be feasible in the sense that: i) each DER contribution is within its capacity limits; and ii) the DERs collectively provide the amount of active or reactive power requested by the coordinator.

Our starting point to achieve the above objectives is a class of iterative algorithms that rely on two linear iterations [3], where the collective dynamics of each iteration can be described by a discrete-time linear system with no inputs, with a transition matrix that is column stochastic and primitive. In this class of algorithms, each node j in the network maintains two values y_j and z_j , which we refer to as internal states, and updates them (independently of each other) to be, respectively, a weighted linear combination of node j 's own previous internal states, and the previous internal states of its neighboring nodes. It is worth noting that, each node broadcasts the same quantity to all receiving nodes, which simplifies the communication scheme. It is easy to see that, except for the initialization of both iterations, the double-iteration algorithm in [3] (which assumes perfectly reliable communication links) is a particular case of a gossip-based algorithm proposed in [4] (which also assumes perfectly reliable communication links) where the transition matrices describing each linear iteration are allowed to vary as time evolves. However, the purpose of the algorithm in [4] is for the nodes to compute the average of their initial values, not to solve the coordination problem with which we are concerned.

In our previous works, the two-iteration algorithm described above has been utilized in several power system applications that studied the interaction of the algorithm with the electrical network. For example, in [5], we utilize the algorithm for addressing the voltage control problem in distribution networks with deep penetration of renewable-based electricity generation resources and plug-in hybrid vehicles. Similarly, in [6], we utilize the algorithm to coordinate the action of DERs in the provision of reactive power to a node of a sub-transmission network. In both of these applications, it is necessary to consider the interplay between the distributed algorithm for DER coordination with the voltage dynamics of the distribution network. In particular, it is necessary to enforce time-scale separation between the operation of the distributed algorithm and the dynamics of the physical network it controls.

The main focus of this paper is a method to robustify the double-iteration algorithm so that it can tolerate failures in communication links and allow each node to converge to the correct value. Our communication link reliability model assumes that, at each time step, a packet containing information from node i to node j is dropped with some probability. Apart from performing the two iterations as in the non-robustified double-iteration algorithm, we show that the double-iteration algorithm can be robustified by having node i keep track of certain additional variables, and include information about these variables in the messages it broadcasts; similarly, node j will keep track of additional variables and will use them in the update of its internal state. We show that the information contained in these additional variables ensures that the nodes can obtain the information they need despite packet-dropping communication links.

The class of algorithms discussed in this paper is similar in spirit to distributed linear-iterative algorithms for consensus problems (see, e.g., [7], [8] and references therein); however, the end goal here is very different. In consensus problems, the objective is to have the nodes agree on some value. In our setup, the objective is for each node to converge to a value (not necessarily the same) that lies within an interval defined by upper and lower capacity limits, while the sum of the values is equal to the total amount of resource requested by the coordinator. Another difference is the communication modality, as we allow for asymmetric information exchange, whereas in most consensus works, except for a few instances (see, e.g., [9], [10]), symmetric information exchange is assumed.

In terms of communication/computational requirements, each node needs to know the number of its out-neighbors, be able to maintain and broadcast a small number of values, and be able to linearly combine (or simply add up) the values it receives from its in-neighbors. Nodes also need to be able to perform the operations of division, summation, and subtraction, and be able to broadcast two quantities per iteration. The complexity of the computation at each node depends on its in-degree (i.e., the number of values that it receives): at each iteration, node j performs $O(\mathcal{D}_j^-)$ computations.

Some recent works have addressed the consensus and average-consensus problems in the presence of unreliable communication links [11], [12], [13]. Unlike our setup, the work in [11] 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 compensate for it by rescaling their other

weights (so that the resulting transition matrix remains column stochastic). The work in [12] does not require the graph describing the communication network to be undirected and proposes two compensation methods to account for communication link failures. The key in both methods is the fact that, at each time step, the resulting weight matrix is row stochastic; the authors show that the nodes converge almost surely to the same value (which, however is not necessarily the average of the initial values). The work in [13] proposes a strategy that corrects the errors in the quantity (state) iteratively calculated by each node, so that nodes obtain the correct average of the initial values. For this strategy to work, nodes must send acknowledgment messages and retransmit information an appropriate number of times.

The remainder of this paper is organized as follows. Section II introduces the communication model, and briefly describes the non-robust version of the double-iteration algorithm. Section III describes an application example to motivate the use of the proposed algorithms. Section IV describes the proposed strategy to robustify the double-iteration algorithm and demonstrates the use/performance of the algorithms via several examples. Section V characterizes the first and second moment of each iteration. The convergence analysis of the robustified double-iteration algorithm is provided in Section VI. Concluding remarks are presented in Section VII.

II. PRELIMINARIES

In this section, we provide background on graph-theoretic notions, needed to describe the communication network model and the DER networked control setup, and also introduce the basic communication link availability model. In addition, we review a previously proposed double-iteration algorithm that assumes reliable communication [3].

A. Communication Network Model

Let k index discrete time instants; the information exchange between nodes (components) at each time instant k can be described by a directed graph $\mathcal{G}[k] = \{\mathcal{V}, \mathcal{E}[k]\}$, where $\mathcal{V} = \{1, 2, \dots, n\}$ is the vertex set (each vertex—or node—corresponds to a system component), and $\mathcal{E}[k] \subseteq \mathcal{V} \times \mathcal{V}$ is the set of edges, where $(j, i) \in \mathcal{E}[k]$ if node j can receive information from node i at instant k . It is assumed that $\mathcal{E}[k] \subseteq \mathcal{E}$, $\forall k \geq 0$, where \mathcal{E} is the set of edges that describe all possibly available communication links between nodes. 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. All nodes that

can possibly transmit information to node j are called its in-neighbors, and are represented by the set $\mathcal{N}_j^- = \{i \in \mathcal{V} : (j, i) \in \mathcal{E}\}$. For notational convenience, we allow self-loops for all nodes in \mathcal{G} (i.e., $(j, j) \in \mathcal{E}$ for all $j \in \mathcal{V}$). The number of 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 $\mathcal{D}_j^+ = |\mathcal{N}_j^+|$.

The existence of a communication link from node i to node j can be described in probabilistic terms as follows. At instant k , let $x_{ji}[k] = 0$ if $(j, i) \notin \mathcal{E}$, and, for all $(j, i) \in \mathcal{E}$, let $x_{ji}[k]$ be an indicator for the link availability, i.e., $x_{ji}[k] = 1$ with probability q_{ji} if the link is available, and $x_{ji}[k] = 0$ with probability $1 - q_{ji}$ if the link is not available:

$$\Pr\{x_{ji}[k] = m\} = \begin{cases} q_{ji}, & \text{if } m = 1, \\ 1 - q_{ji}, & \text{if } m = 0. \end{cases} \quad (1)$$

We assume that link availability is independent between links and between time steps.

B. DER Networked Control with Reliable Communications

In a networked control setup for DER coordination, each DER j can contribute a certain amount $\pi_j \geq 0$ of active or reactive power, which is upper and lower bounded by capacity limits π_j^{max} and π_j^{min} respectively, which are known to node j . The objective is to coordinate the individual DER contributions π_j , $\forall j$, through a distributed computation over the network so that $\pi_j^{min} \leq \pi_j \leq \pi_j^{max}$, $\forall j$, and collectively the nodes provide a pre-determined total amount $\rho_d = \sum_{j=1}^n \pi_j$ of active or reactive power, as specified by a coordinator.

Under the assumption that the communication network graph $(\mathcal{G}, \mathcal{E})$ is strongly connected, and every communication link between nodes is perfectly reliable, i.e., $\Pr\{x_{ji}[k] = 1\} = 1$, $\forall (j, i) \in \mathcal{E}$, $\forall k \geq 0$, it was shown in [3] that the DER coordination problem can be solved by running two separate, appropriately initialized, linear iterations of the form

$$y_j[k+1] = \sum_{i \in \mathcal{N}_j^-} \frac{1}{\mathcal{D}_i^+} y_i[k], \quad (2)$$

$$z_j[k+1] = \sum_{i \in \mathcal{N}_j^-} \frac{1}{\mathcal{D}_i^+} z_i[k], \quad (3)$$

where \mathcal{D}_j^+ (\mathcal{D}_i^+) is the out-degree of node j (i). In particular, i) if the initial conditions in (2) are set to $y_j[0] = \rho_d/m - \pi_j^{\min}$ if j is one of m , $m \geq 1$, nodes contacted initially by the coordinator and $y_j[0] = -\pi_j^{\min}$ otherwise, and ii) if the initial conditions in (3) are set to $z_j[0] = \pi_j^{\max} - \pi_j^{\min}$, then the nodes can asymptotically calculate their contribution π_j as

$$\begin{aligned} \pi_j &:= \lim_{k \rightarrow \infty} \pi_j[k] = \lim_{k \rightarrow \infty} \left(\pi_j^{\min} + \frac{y_j[k]}{z_j[k]} (\pi_j^{\max} - \pi_j^{\min}) \right), \\ &= \pi_j^{\min} + \frac{\rho_d - \sum_{l=1}^n \pi_l^{\min}}{\sum_{l=1}^n \ell_l} (\pi_j^{\max} - \pi_j^{\min}), \end{aligned} \quad (4)$$

which satisfies $\pi_j^{\min} \leq \pi_j \leq \pi_j^{\max}$, $\forall j$, and $\sum_{j=1}^n \pi_j = \rho_d$.

III. MOTIVATING APPLICATION EXAMPLE: TWO-STAGE CONTROL ARCHITECTURE FOR VOLTAGE CONTROL

In this section, we illustrate the application of the algorithms proposed for coordinating DERs to provide reactive power support for voltage control. In an AC power system, reactive power flows impact the voltage of the system buses [14]; thus, by controlling reactive power it is possible to control voltage, which is extremely important to ensure voltage stability in the system. We consider the system displayed on Fig. 1, which we have used in our previous work to illustrate the application of the (non-robust) double-iteration algorithm in (2)–(3) for DER coordination [6]. We use it again here to study the performance of the double-iteration in (2)–(3) in the presence of packet-dropping communication links; the results will provide motivation for the remainder of the paper—the development of a method to handle packet drops in the communication links.

The left side of Fig. 1 describes the topology of a small power system comprised of three electrical generators, denoted by G_1 , G_2 , and G_3 , connected to buses 1, 2, and 3 respectively,

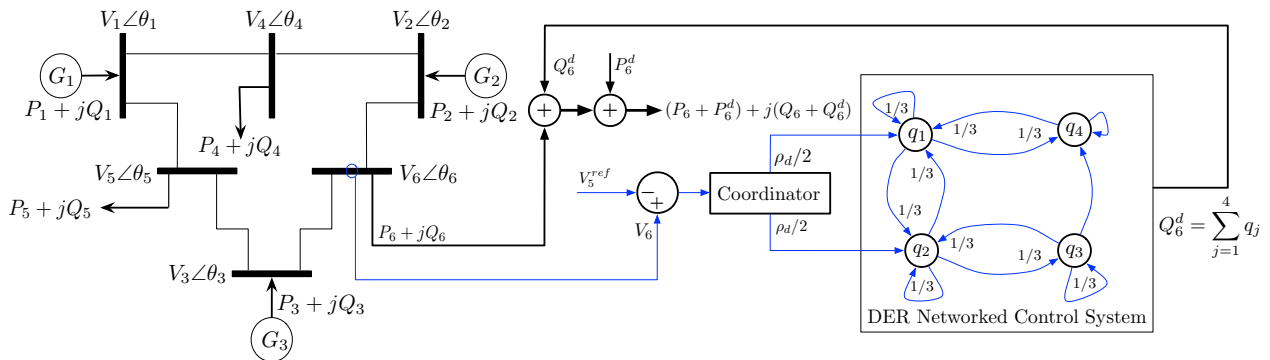


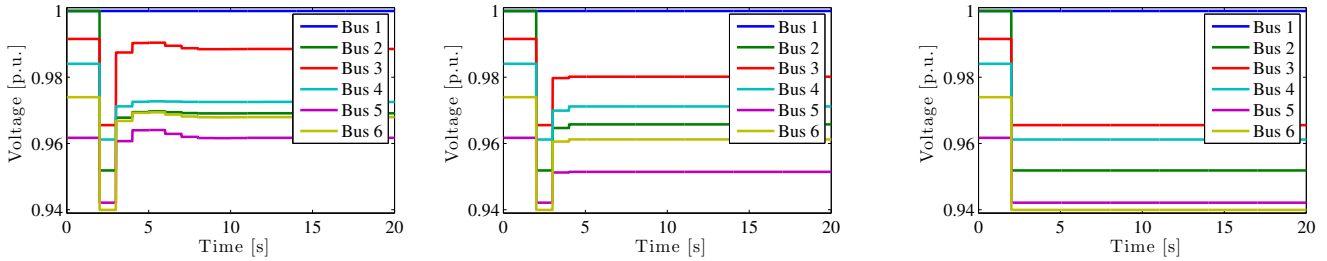
Fig. 1. Two-stage voltage control architecture.

and three loads connected to buses 4, 5, and 6. The complex power injected (or withdrawn) on each of bus i is denoted by $\bar{S}_i = P_i + jQ_i$, and the voltage on each bus i is described by the complex number $\bar{V}_i = V_i \angle \theta_i$. As shown in the picture, each pair of buses is linked by a transmission line. On the right of Fig. 1, we display a two-stage control system that utilizes DERs to control the reactive power injection Q_6^d provided by the four DERs connected to this bus. By controlling Q_6^d , we can regulate V_6 (i.e., bus 6 voltage magnitude) with the objective of maintaining it close to a reference voltage V_6^{ref} .

In order to achieve the voltage regulation objective described above, a local coordination controller—the *coordinator*—takes the difference between V_6 , the actual voltage at node 6, and the reference voltage V_6^{ref} , and issues a command demanding a total amount of reactive power denoted by ρ_d . The coordinator then relays evenly split demands for reactive power to a subset of the DERs that it can directly communicate with (nodes 1 and 2 in our example), so that the total amount ρ_d remains constant. Thus, in Fig. 1, node 1 gets $\rho_d/2$ and node 2 gets $\rho_d/2$. Then, all four nodes run the double-iteration algorithm described in (2) and (3); in particular, and following the notation in (2), we have that

$$\begin{aligned} y_1[k+1] &= \frac{1}{3}y_1[k] + \frac{1}{3}y_2[k] + \frac{1}{3}y_3[k], \\ y_2[k+1] &= \frac{1}{3}y_1[k] + \frac{1}{3}y_2[k] + \frac{1}{2}y_4[k], \\ y_3[k+1] &= \frac{1}{3}y_1[k] + \frac{1}{3}y_3[k], \\ y_4[k+1] &= \frac{1}{3}y_2[k] + \frac{1}{3}y_3[k] + \frac{1}{2}y_4[k], \end{aligned} \tag{5}$$

with $y_1[0] = y_2[0] = 1/2\rho_d$, and $y_3[0] = y_4[0] = 0$. [We also have an identical iteration to the one in (5) for the update of the $z_j[k]$'s. with initial conditions $z_1[0] = z_3[0] = 0.4$, $z_2[0] = 0.2$, and $z_4[0] = 0.1$.] After each DER j has calculated its π_j as described in (4), it sets its reactive power command q_j accordingly, i.e., $q_j = \pi_j$, such that $\rho_d = Q_s^d = \sum_{j=1}^4 q_j$. It is assumed that there is a time-scale separation between coordinator requests and the time it takes the DERs to calculate π_j . This time-scale separation is actually related to the speed with which the algorithm converges, and, in general, it very much depends on the topology of both the communication and the electrical network (see [6] for more details on the overall control system design).



(a) All controls working properly. (b) Failures in DER communications in bus 2. (c) Failures in DER communications in all buses.

Fig. 2. Bus voltages for different malfunction scenarios.

1) *Perfectly reliable links*: Consider the same scenario as in [6], where a contingency shuts down generator G_2 at $t = 2$ s. Before the contingency, generator G_2 was providing voltage control, but was operating as its maximum capacity, therefore, after the contingency G_2 can no longer provide voltage control. The contingency causes a voltage drop on every bus except bus 1 (the slack bus). In this scenario, G_1 is required to produce additional active and reactive power to compensate for the loss of G_2 , but as discussed in [6], it cannot provide voltage control in all load buses, and (as seen in Fig 2(a)) the voltages in buses 5 and 6 drop below 0.95 p.u. Once the coordinators on these buses sense their voltage is low, they initiate a request to the DERs connected to each of them. As seen in Fig. 2(a), within 8 seconds, the voltages in these buses are restored to a value above 0.95 p.u. In the process, the voltage in bus 2 corresponding to the outage generator also recovers, as well as the voltage in bus 3, corresponding to the generator that could not provide voltage control. By setting the DER capacity limits to $\pi_1^{max} = \pi_3^{max} = 0.4$, $\pi_2^{max} = 0.2$, $\pi_4^{max} = 0.1$, and $\pi_1^{min} = \pi_2^{max} = \pi_3^{min} = \pi_4^{max} = 0$, Fig. 2(a) shows the evolution of the $\pi_j[k]$'s after the coordinator makes the first request; it can be seen that the algorithm converges in less than 10 iterations.

2) *Effect of packet-dropping links*: As seen in Fig. 2(b), when the communication links between the DERs connected to bus 2 are not working properly, the voltages in all buses do not recover to the same levels as in Fig. 2(a). To better understand the effect of communication link failures, in Fig. 3, we display back-to-back the evolution of the double-iteration algorithm for the DERs in load bus 2 for the case when their communication network is perfectly reliable and for the case when the communication links drop packages with probability $q_{ji} = 0.5$, $\forall (j, i) \in \mathcal{E}$ at each time step k . In the perfectly reliable case, as displayed in Fig. 3(a), the double-iteration algorithm converges to some nonzero values. On the other hand, when links drop packages, as displayed in Fig. 3(b), the double-iteration algorithm that each DER runs converges to zero

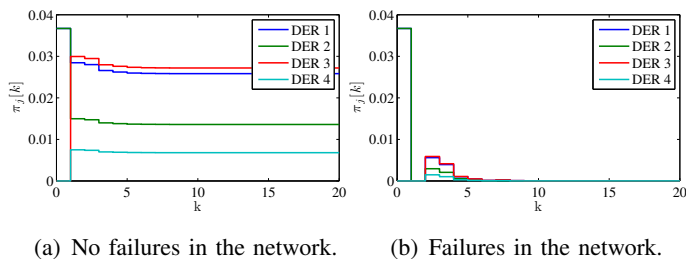


Fig. 3. Responses of the DERs in bus 2.

and therefore, the DERs will not provide any reactive power support even if the coordinator requested it. Finally, Fig. 2(c) shows the case in which link failures affect all the connections between DERs in all three networked control systems.

IV. ROBUSTIFIED DOUBLE-ITERATION ALGORITHM

Consider the setup of the previous section with a strongly connected graph $(\mathcal{G}, \mathcal{E})$ representing a multicomponent system and the communication links between its components. Each node acts as both sender and receiver at each time step. For each of the two iterations, node j calculates several quantities of interest, which we refer to as: i) internal state; ii) total broadcast mass; and iii) total received mass from each in-neighbor l of node j , i.e., for each node $l \in \mathcal{N}_j^-$. Node j updates the value of its internal state to be a linear combination of its own previous internal state value (weighted by $1/\mathcal{D}_j^+$) and the sum (over all its in-neighbors) of the differences between the two most recently received mass values. At instant time k , the total broadcast mass of i is the sum up to (and including) step k of the weighted value of that node's internal state. Additionally, node j updates the value of the received mass from node l to be either the total broadcast mass sent by node l if the link from l to j is available, or the most recently received mass value from node l . An implicit assumption is that broadcast messages are tagged with the sender's identity so that receiving nodes can determine where messages originated from.

For iteration 1, let $y_j[k]$ be node j 's internal state at time instant k , $\mu_{lj}[k]$ denote the mass broadcast from node j to each of its out-neighbors l (this is the same for each out-neighbor l of node j , i.e., for each $l \in \mathcal{N}_j^+$), and $\nu_{ji}[k]$ denote the mass received at node j from node $i \in \mathcal{N}_j^-$. Similarly, let $z_j[k]$ be node j 's internal state, $\sigma_{lj}[k]$ denote node j 's broadcast mass for each out-neighbor l , $l \in \mathcal{N}_j^+$, and $\tau_{ji}[k]$ denote the total mass received from $i \in \mathcal{N}_j^-$. Then, for

all $k \geq 0$,

$$\begin{aligned} y_j[k+1] &= \frac{1}{\mathcal{D}_j^+} y_j[k] + \sum_{i \in \mathcal{N}_j^-} (\nu_{ji}[k] - \nu_{ji}[k-1]), \\ \mu_{lj}[k] &= \mu_{lj}[k-1] + \frac{1}{\mathcal{D}_j^+} y_j[k] = \sum_{t=0}^k \frac{1}{\mathcal{D}_j^+} y_j[t], \end{aligned} \quad (6)$$

where

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

[Recall that \mathcal{D}_j^+ (\mathcal{D}_i^+) is the number of nodes that node j (i) can transmit information to.] Similarly, for all $k \geq 0$,

$$\begin{aligned} z_j[k+1] &= \frac{1}{\mathcal{D}_j^+} z_j[k] + \sum_{i \in \mathcal{N}_j^-} (\tau_{ji}[k] - \tau_{ji}[k-1]), \\ \sigma_{lj}[k] &= \sigma_{lj}[k-1] + \frac{1}{\mathcal{D}_j^+} z_j[k] = \sum_{t=0}^k \frac{1}{\mathcal{D}_j^+} z_j[t], \end{aligned} \quad (7)$$

where

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

As mentioned earlier, the initial conditions in (6) are set to $y_j[0] = \rho_d/m - \pi_j^{min}$ if j is initially contacted by the leader and $y_j[0] = -\pi_j^{min}$ otherwise, whereas the initial conditions in (7) are set to $z_j[0] = \pi_j^{max} - \pi_j^{min} > 0$. Additionally, $\mu_{ji}[-1] = 0$ and $\nu_{ji}[-1] = 0$ for all $(j, i) \in \mathcal{E}$, and $\sigma_{ji}[-1] = 0$ and $\tau_{ji}[-1] = 0$ for all $(j, i) \in \mathcal{E}$.

Main Result: We shall argue that with the proposed robustification strategy, despite the presence of unreliable communication links (at each time step, each link $(j, i) \in \mathcal{E}$ fails, independently from other links and independently between time steps, with some probability $1 - q_{ji}$), nodes can asymptotically obtain the exact solution to the resource coordination problem as a function of the ratio $y_j[k]/z_j[k]$, i.e.,

$$\pi_j := \lim_{k \rightarrow \infty} \left(\pi_j^{min} + \frac{y_j[k]}{z_j[k]} (\pi_j^{max} - \pi_j^{min}) \right), \quad (8)$$

computed whenever $z_j[k]$ is large enough. In particular, we show that $\lim_{k \rightarrow \infty} y_j[k] = \alpha \lim_{k \rightarrow \infty} z_j[k]$, for $\alpha = \frac{\sum_{j=1}^n y_0(j)}{\sum_{j=1}^n z_0(j)}$ with probability one. We also show that $z_j[k] > 0$ occurs *infinitely often* (i.o.).

Example 1: We complete the application example presented in Section III, assuming $q_{ji} = 0.5$, $\forall (j, i) \in \mathcal{E}$. The evolution of the algorithm right after the first coordinator request (following the contingency) is displayed in Fig. 4. The evolution of the $y_j[k]$'s and $z_j[k]$'s is displayed in Figs. 4(b) and 4(c). Neither of the two iterations converges, but, as shown in Fig. 4(a), the $\pi_j[k]$'s obtained from (8) converge, indicating that the $y_j[k]/z_j[k]$'s converge. \square

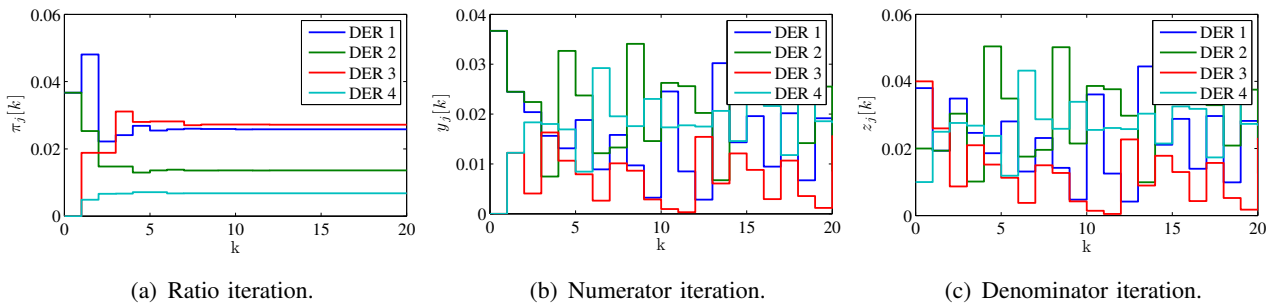


Fig. 4. Robustified algorithm for voltage control in the system of Fig. 1.

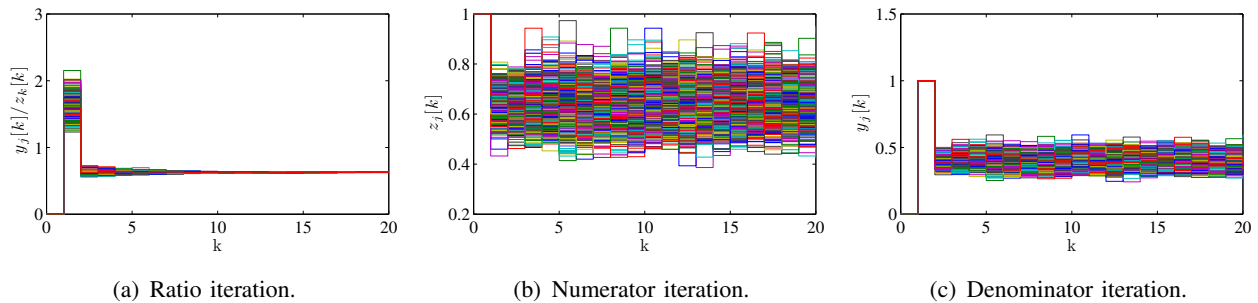
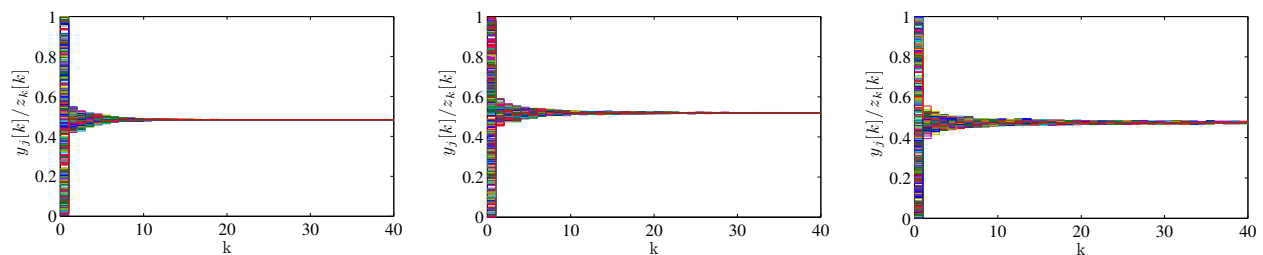


Fig. 5. Robustified algorithm for reaching average consensus in a large graph.



(a) Ratio iteration for $0.45 \leq q_{ji} \leq 0.55$. (b) Ratio iteration for $0.4 \leq q_{ji} \leq 0.6$. (c) Ratio iteration for $0.05 \leq q_{ji} \leq 0.95$.

Fig. 6. Robustified algorithm for reaching average consensus with unequal failure probabilities of varying dispersion.

Example 2: An example of what happens in a larger graph is shown in Fig. 5. Here we consider a graph with 500 nodes, randomly generated by choosing an edge from node i to node j , $1 \leq i, j \leq 500$, $i \neq j$, independently with probability 0.5, and ensuring that the resulting graph is strongly connected. For $q_{ji} = 0.5$, $\forall (j, i) \in \mathcal{E}$, the behavior remains similar to what we observed in Example 1: the ratio $y_j[k]/z_j[k]$ converges to the correct solution even though the individual $y_j[k]$ and $z_j[k]$ do not converge. \square

Example 3: In both Examples 1 and 2, we assumed equal link failure probabilities; in this example, we explore the effect of unequal failure link probabilities on the convergence of the algorithm. As in Example 2, we generate a graph with 500 nodes and for each edge $(j, i) \in \mathcal{E}$, we randomly choose its failure probability q_{ji} to lie between $[0.45, 0.55]$, $[0.4, 0.6]$, or $[0.05, 0.95]$. As shown in Fig. 6, the more dispersion around the mean of q_{ji} (which is $\bar{q}_{ji} = 0.5$ in all three cases), the longer it takes $y_j[k]/z_j[k]$ to converge; but it eventually converges to the correct solution. \square

V. FIRST AND SECOND MOMENT ANALYSIS

In this section, we characterize the first and second moment of the iterations in (6) and (7). The results are then used in Section VI to establish the claims previously stated as *Main Result*. The analysis is restricted to the equal link failure probability case, i.e., $q_{ji} = q$, $\forall (j, i) \in \mathcal{E}$, with $0 < q < 1$. The assumption of equal probabilities simplifies the notation used in the derivation of the expressions that characterize the first and second moment in (6) and (7), however, the techniques used in this paper to develop these expressions (which involve the use of Hadamard and Kronecker products) should also be able to handle unequal link failure probabilities, although the details are left to future work (Example 3 provides some evidence that indeed the algorithm converges even when link failure probabilities are unequal). Additionally, to facilitate notation, we will allow each node j to “drop” the packet carrying its own previous value when updating its value. This way, node j handles its own value in the same way as its neighbors’ values and notation is simplified significantly.

A. Vectorized Description of Double-Iteration Algorithm

In order to ease the calculations, the iterations in (6)–(7) will be rewritten more compactly in vector form. Using the definition for the indicator variable $x_{ji}[k]$ given in (1), which describes

the successful transmission of information from node i to node j over an existing, unreliable link, iterations (6) and (7) can be rewritten, for all $k \geq 0$, as

$$\mu_{lj}[k] = \begin{cases} \mu_{lj}[k-1] + \frac{1}{\mathcal{D}_j^+} y_j[k], & \text{if } l \in \mathcal{N}_j^+, \\ 0, & \text{if } l \notin \mathcal{N}_j^+, \end{cases} \quad (9)$$

$$\nu_{ji}[k] = \begin{cases} \mu_{ji}[k]x_{ji}[k] + \nu_{ji}[k-1](1 - x_{ji}[k]), & \text{if } i \in \mathcal{N}_j^-, \\ 0, & \text{if } i \notin \mathcal{N}_j^-, \end{cases} \quad (10)$$

$$y_j[k+1] = \sum_{i=1}^n (\nu_{ji}[k] - \nu_{ji}[k-1]); \quad (11)$$

$$\sigma_{lj}[k] = \begin{cases} \sigma_{lj}[k-1] + \frac{1}{\mathcal{D}_j^+} z_j[k], & \text{if } l \in \mathcal{N}_j^+, \\ 0, & \text{if } l \notin \mathcal{N}_j^+, \end{cases} \quad (12)$$

$$\tau_{ji}[k] = \begin{cases} \sigma_{ji}[k]x_{ji}[k] + \tau_{ji}[k-1](1 - x_{ji}[k]), & \text{if } i \in \mathcal{N}_j^-, \\ 0, & \text{if } i \notin \mathcal{N}_j^-, \end{cases} \quad (13)$$

$$z_j[k+1] = \sum_{i=1}^n (\tau_{ji}[k] - \tau_{ji}[k-1]), \quad (14)$$

where $\mu_{ji}[-1] = \nu_{ji}[-1] = \sigma_{ji}[-1] = \tau_{ji}[-1] = 0$, $\forall j, i$.

Let $A \circ B$ denote the Hadamard (entry-wise) product of a pair of matrices A and B of identical size. Then, for all $k \geq 0$, iteration (9)–(11) can be rewritten in matrix form as

$$M_k = M_{k-1} + P \text{diag}(y_k), \quad (15)$$

$$N_k = M_k \circ X_k + N_{k-1} \circ (U - X_k), \quad (16)$$

$$y_{k+1} = (N_k - N_{k-1})e = [(M_k - N_{k-1}) \circ X_k]e, \quad (17)$$

where $P = [p_{ji}] \in \mathbb{R}^{n \times n}$, with $p_{ji} = \frac{1}{\mathcal{D}_j^+}$, $\forall j \in \mathcal{N}_i^+$ and $p_{ji} = 0$ otherwise; $M_{-1} = N_{-1} = 0$; $y_k = y[k]$; $U \in \mathbb{R}^{n \times n}$, with $[U_{ji}] = 1$, $\forall i, j$; $\text{diag}(y_k)$ is the diagonal matrix that results by having the entries of y_k on the main diagonal; and $e = [1, 1, \dots, 1]^T$ (note that $U = ee^T$). Note that X_k is a matrix whose (j, i) entry is $x_{ji}[k]$. Similar expressions can be obtained for (12)–(14), with z_k replacing y_k , S_k replacing M_k , and T_k replacing N_k .

By defining $A_k := M_k - N_{k-1}$ and $B_k := S_k - T_{k-1}$, iteration (15)–(17) can be rewritten more

compactly as

$$A_k = A_{k-1} \circ (U - X_{k-1}) + P \text{diag}(y_k), \quad k \geq 1, \quad (18)$$

$$y_{k+1} = (A_k \circ X_k)e, \quad k \geq 0. \quad (19)$$

Similarly, we can write iteration (12)–(14) with B_k and z_k replacing A_k and y_k respectively.

For analysis purposes, each matrix in (18)–(19) will be rewritten in vector form by stacking up the corresponding columns.¹ Let $F = [I_n \ I_n \ \dots \ I_n] \in \mathbb{R}^{n \times n^2}$, where I_n is the $n \times n$ identity matrix, and $\tilde{P} = [E_1 P^T \ E_2 P^T \ \dots \ E_n P^T]^T \in \mathbb{R}^{n^2 \times n}$, where $E_i \in \mathbb{R}^{n \times n}$ has $E_i(i, i) = 1$ and all other entries equal zero. [The entries of $E_i P^T \in \mathbb{R}^{n \times n}$ ($P E_i^T = P E_i$) 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, (18)–(19) can be rewritten as

$$a_k = a_{k-1} \circ (u - x_{k-1}) + \tilde{P} y_k, \quad k \geq 1, \quad (20)$$

$$y_{k+1} = F(a_k \circ x_k), \quad k \geq 0, \quad (21)$$

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

$$b_k = b_{k-1} \circ (u - x_{k-1}) + \tilde{P} z_k, \quad k \geq 1, \quad (22)$$

$$z_{k+1} = F(b_k \circ x_k), \quad k \geq 0, \quad (23)$$

where $b_k \in \mathbb{R}^{n^2}$ results from stacking the columns of B_k .

Remark 1: Note that matrices A_k and B_k , and their corresponding vectors a_k and b_k , have some entries that remain at zero for all $k \geq 0$. Specifically, the (j, i) entry of matrices A_k and B_k (and their corresponding entries in a_k and b_k) remain zero if there is no communication link from node i to node j , i.e., $(j, i) \notin \mathcal{E}$. We keep these entries because they simplify the notation in subsequent developments. Perhaps more importantly, we allow all entries of X_k (and thus the vector x_k) to take value 0 or 1 with probabilities $1 - q$ and q respectively, independently between links and time steps. This makes no difference in the evolution of (20)–(21). \square

¹Vectors defined by stacking the columns of a matrix will be denoted with the same small letter as the capital letter of the corresponding matrix, i.e., let $A = [A_{ji}] \in \mathbb{R}^{n \times n}$, then $a = [A_{11} \ \dots \ A_{n1}, A_{12}, \dots, A_{n2}, \dots, A_{1n}, \dots, A_{nn}]^T$.

B. First Moment Analysis

Next, we show that the dynamics of $\bar{a}_k := \mathbf{E}[a_k]$, $\bar{y}_k := \mathbf{E}[y_k]$, $\bar{b}_k := \mathbf{E}[b_k]$, and $\bar{z}_k := \mathbf{E}[z_k]$ can be described via discrete-time linear systems with no inputs, where the corresponding transition matrices are column stochastic and primitive. Furthermore, \bar{y}_k and \bar{z}_k (and also \bar{a}_k and \bar{b}_k) are shown to converge to an identical solution up to scalar multiplication.

Lemma 1: Let a_k , b_k , y_k , and z_k be described by the recurrence relations in (20)–(21), and (22)–(23) respectively. Let the first moments of a_k , y_k , b_k , and z_k be denoted by \bar{a}_k , \bar{y}_k , \bar{b}_k , and \bar{z}_k respectively. Then the evolution of \bar{a}_k , \bar{y}_k , \bar{b}_k , and \bar{z}_k , $\forall k \geq 1$, is governed by

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

$$\bar{y}_{k+1} = [qP + (1 - q)I_n]\bar{y}_k, \quad (25)$$

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

$$\bar{z}_{k+1} = [qP + (1 - q)I_n]\bar{z}_k, \quad (27)$$

where I_m is the $m \times m$ identity matrix, with $\bar{a}_0 = \tilde{P}y_0$, $\bar{y}_1 = qPy_0$, $\bar{b}_0 = \tilde{P}z_0$, and $\bar{z}_1 = qPz_0$.

Proof: Since the development for obtaining \bar{a}_k and \bar{y}_k is parallel to that for obtaining \bar{b}_k and \bar{z}_k , our analysis focuses on the first case. For $k = 0$ in (20)–(21), 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 $\bar{a}_0 = \tilde{P}y_0$, and $\bar{y}_1 = qF\bar{a}_0$; therefore, $\bar{y}_1 = qF\tilde{P}y_0 = qP\bar{y}_0$.

For $k \geq 1$ in (20)–(21), 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

$$\bar{a}_k = \bar{a}_{k-1} \circ (u - \bar{x}_{k-1}) + \tilde{P}\bar{y}_k = (1 - q)\bar{a}_{k-1} + \tilde{P}\bar{y}_k, \quad (28)$$

$$\bar{y}_{k+1} = F(\bar{a}_k \circ \bar{x}_k) = qF\bar{a}_k. \quad (29)$$

Substituting (29) into (28), we obtain $\bar{a}_k = (1 - q)\bar{a}_{k-1} + q\tilde{P}F\bar{a}_{k-1} = [q\tilde{P}F + (1 - q)I_{n^2}]\bar{a}_{k-1}$. Similarly, substituting (28) into (29), we have $\bar{y}_{k+1} = (1 - q)qF\bar{a}_{k-1} + qF\tilde{P}\bar{y}_k = (1 - q)\bar{y}_k + qF\tilde{P}\bar{y}_k = [qP + (1 - q)I_n]\bar{y}_k$. ■

Lemma 2: 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: i) $\lim_{k \rightarrow \infty} \bar{y}_k = \alpha \lim_{k \rightarrow \infty} \bar{z}_k$, and ii) $\lim_{k \rightarrow \infty} \bar{a}_k = \alpha \lim_{k \rightarrow \infty} \bar{b}_k$, where $\alpha = \frac{\sum_{j=1}^n y_0(j)}{\sum_{j=1}^n z_0(j)}$.

Proof: In Lemma 1, it was shown that $\bar{y}_{k+1} = [qP+(1-q)I]\bar{y}_k$ and $\bar{z}_{k+1} = [qP+(1-q)I]\bar{z}_k$ with $\bar{y}_1 = qy_0$, and $\bar{z}_1 = qz_0$. Since P is column stochastic and primitive, it follows that $[qP+(1-q)I]$ is also column stochastic and primitive. Thus, the solutions of (25) and (27) are unique up to scalar multiplication, i.e., $\lim_{k \rightarrow \infty} \bar{y}_k = \alpha \lim_{k \rightarrow \infty} \bar{z}_k$, for some α . Then, from the column stochasticity property: $\sum_{j=1}^n \bar{z}_k(j) = q(\sum_{j=1}^n z_0(j))$ and $\sum_{j=1}^n \bar{y}_k(j) = q(\sum_{j=1}^n y_0(j))$, $\forall k \geq 1$; this implies that $\alpha = \frac{\sum_{j=1}^n y_0(j)}{\sum_{j=1}^n z_0(j)}$. From (20) and (22), it is easy to see that $\tilde{P} \lim_{k \rightarrow \infty} \bar{y}_k = q \lim_{k \rightarrow \infty} \bar{a}_k$ and $\tilde{P} \lim_{k \rightarrow \infty} \bar{z}_k = q \lim_{k \rightarrow \infty} \bar{b}_k$. Since, $\lim_{k \rightarrow \infty} \bar{y}_k = \alpha \lim_{k \rightarrow \infty} \bar{z}_k$, it immediately follows that $\lim_{k \rightarrow \infty} \bar{a}_k = \alpha \lim_{k \rightarrow \infty} \bar{b}_k$. ■

C. 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. In order to characterize Γ_k , Φ_k , Ψ_k , Λ_k , Ξ_k , and Υ_k , we need the results in the following lemma.

Lemma 3: Let x , c and d be random vectors of dimension n . Furthermore, assume that the entries of x are Bernoulli *i.i.d.* random variables such that $\Pr\{x_i = 1\} = q$ and $\Pr\{x_i = 0\} = 1 - q$, $\forall i = 1, 2, \dots, n$, and are independent from c and d . Then $S := \mathbf{E}[(c \circ x)(x \circ d)^T] = q^2 \mathbf{E}[cd^T] + q(1-q) \mathbf{E}[\text{diag}(cd^T)]$, and $T := \mathbf{E}[(c \circ x)((u-x) \circ d)^T] = q(1-q) \mathbf{E}[cd^T] - q(1-q) \mathbf{E}[\text{diag}(cd^T)]$, where $\text{diag}(cd^T)$ is a diagonal matrix with the same diagonal as matrix cd^T .

Proof: The (i, j) , $i \neq j$, entry of S can be obtained as $S_{ij} = \mathbf{E}[c_i x_i d_j x_j]$. Since x_i and x_j are pairwise independent, and independent from c and d , it follows that $\mathbf{E}[c_i x_i d_j x_j] = q^2 \mathbf{E}[c_i d_j]$. For $i = j$, observing that $\mathbf{E}[x_i x_i] = \mathbf{E}[x_i] = q$, $\forall i = 1, \dots, n$, we obtain the corresponding entry of S as $S_{ii} = \mathbf{E}[c_i x_i d_i x_i] = \mathbf{E}[c_i d_i x_i] = q \mathbf{E}[c_i d_i]$. Combining these two facts, it follows that $S := \mathbf{E}[(c \circ x)(x \circ d)^T] = q^2 \mathbf{E}[cd^T] + q(1-q) \mathbf{E}[\text{diag}(cd^T)]$. Similarly, the (i, j) , $i \neq j$, entry of T is given by $T_{ij} = \mathbf{E}[c_i x_i d_j (1-x_j)]$. Since x_i and $(1-x_j)$ are independent, it follows that $\mathbf{E}[c_i x_i d_j (1-x_j)] = \mathbf{E}[c_i d_j] \mathbf{E}[x_i (1-x_j)] = q(1-q) \mathbf{E}[c_i d_j]$. For $i = j$, and observing that $\mathbf{E}[x_i (1-x_i)] = 0$, $\forall i = 1, \dots, n$, the corresponding entry of T can be obtained as $T_{ii} = \mathbf{E}[c_i x_i d_i (1-x_i)] = \mathbf{E}[c_i d_i] \mathbf{E}[x_i (1-x_i)] = 0$. ■

Lemma 4: Consider the second moments of a_k , y_k , b_k , and z_k , and let $\mathbf{E}[a_k a_k^T]$, $\mathbf{E}[y_k y_k^T]$, $\mathbf{E}[b_k b_k^T]$, $\mathbf{E}[z_k z_k^T]$, $\mathbf{E}[a_k b_k^T]$, and $\mathbf{E}[y_k z_k^T]$ be denoted by Γ_k , Φ_k , Ψ_k , Λ_k , Ξ_k , and Υ_k respectively.

Then, the evolutions of $\Gamma_k, \Phi_k, \Psi_k, \Lambda_k, \Xi_k, \Upsilon_k, \forall k \geq 1$, are described by the following iterations (where all I denote $n^2 \times n^2$ identity matrices):

$$\begin{aligned} \Gamma_k &= [q\tilde{P}F + (1-q)I]\Gamma_{k-1}[q\tilde{P}F + (1-q)I]^T \\ &\quad + q(1-q)[I - \tilde{P}F]\text{diag}(\Gamma_{k-1})[I - \tilde{P}F]^T, \end{aligned} \quad (30)$$

$$\Phi_{k+1} = F[q^2\Gamma_k + q(1-q)\text{diag}(\Gamma_k)]F^T, \quad (31)$$

$$\begin{aligned} \Psi_k &= [q\tilde{P}F + (1-q)I]\Psi_{k-1}[q\tilde{P}F + (1-q)I]^T \\ &\quad + q(1-q)[I - \tilde{P}F]\text{diag}(\Psi_{k-1})[I - \tilde{P}F]^T, \end{aligned} \quad (32)$$

$$\Lambda_{k+1} = F[q^2\Psi_k + q(1-q)\text{diag}(\Psi_k)]F^T, \quad (33)$$

$$\begin{aligned} \Xi_k &= [q\tilde{P}F + (1-q)I]\Xi_{k-1}[q\tilde{P}F + (1-q)I]^T \\ &\quad + q(1-q)[I - \tilde{P}F]\text{diag}(\Xi_{k-1})[I - \tilde{P}F]^T, \end{aligned} \quad (34)$$

$$\Upsilon_{k+1} = F[q^2\Xi_k + q(1-q)\text{diag}(\Xi_k)]F^T, \quad (35)$$

with initial conditions $\Gamma_0 = \tilde{P}y_0y_0^T\tilde{P}^T$, $\Phi_1 = \bar{y}_1\bar{y}_1^T + q(1-q)F\text{diag}(\tilde{P}y_0y_0^T\tilde{P}^T)F^T$, $\Psi_0 = \tilde{P}z_0z_0^T\tilde{P}^T$, $\Lambda_1 = \bar{z}_1\bar{z}_1^T + q(1-q)F\text{diag}(\tilde{P}z_0z_0^T\tilde{P}^T)F^T$, $\Xi_0 = \tilde{P}y_0z_0^T\tilde{P}^T$, $\Upsilon_1 = \bar{y}_1\bar{z}_1^T + q(1-q)F\text{diag}(\tilde{P}y_0z_0^T\tilde{P}^T)F^T$.

Proof: For $k = 0$, it follows from Lemma 1 and (20) that $a_0 = \tilde{P}y_0$. Then, $\Gamma_0 = \mathbf{E}[a_0a_0^T] = \tilde{P}\mathbf{E}[y_0y_0^T]\tilde{P}^T = \tilde{P}y_0y_0^T\tilde{P}^T$, $\Phi_1 = \mathbf{E}[y_1y_1^T] = \mathbf{E}[F(a_0 \circ x_0)(x_0 \circ a_0)^T F^T] = F\mathbf{E}[(a_0 \circ x_0)(x_0 \circ a_0)^T]F^T$, and applying the results in Lemmas 1 and 3, it follows that

$$\begin{aligned} \Phi_1 &= q^2F\mathbf{E}[a_0a_0^T]F^T + q(1-q)F\mathbf{E}[\text{diag}(a_0a_0^T)]F^T \\ &= (qF\tilde{P}y_0)(qF\tilde{P}y_0)^T + q(1-q)F\mathbf{E}[\text{diag}(\tilde{P}y_0y_0^T\tilde{P}^T)]F^T \\ &= (qPy_0)(qPy_0)^T + q(1-q)F\text{diag}(\tilde{P}y_0y_0^T\tilde{P}^T)F^T \\ &= \bar{y}_1\bar{y}_1^T + q(1-q)F\text{diag}(\tilde{P}y_0y_0^T\tilde{P}^T)F^T, \end{aligned} \quad (36)$$

where we used the fact that $F\tilde{P} = P$.

For $k \geq 1$, and taking into account that $y_k = F(a_{k-1} \circ x_{k-1})$, it follows that

$$\begin{aligned}
\Gamma_k &= \mathbf{E} \left[(a_{k-1} \circ (u - x_{k-1}) + \tilde{P}y_k)(a_{k-1} \circ (u - x_{k-1}) + \tilde{P}y_k)^T \right] \\
&= \mathbf{E} \left[(a_{k-1} \circ (u - x_{k-1}))(a_{k-1} \circ (u - x_{k-1}))^T \right] \\
&\quad + \mathbf{E} \left[(a_{k-1} \circ (u - x_{k-1}))(a_{k-1} \circ x_{k-1})^T \right] F^T \tilde{P}^T \\
&\quad + \tilde{P}F \mathbf{E} \left[(a_{k-1} \circ x_{k-1})(a_{k-1} \circ (u - x_{k-1}))^T \right] \\
&\quad + \tilde{P}F \mathbf{E} \left[(a_{k-1} \circ x_{k-1})(a_{k-1} \circ x_{k-1})^T \right] F^T \tilde{P}^T.
\end{aligned} \tag{37}$$

Then, from Lemma 3, (37) can be rewritten as

$$\begin{aligned}
\Gamma_k &= (1 - q)^2 \mathbf{E} [a_{k-1}a_{k-1}^T] + q(1 - q) \mathbf{E} [\text{diag}(a_{k-1}a_{k-1}^T)] \\
&\quad + q(1 - q) \mathbf{E} [a_{k-1}a_{k-1}^T] F^T \tilde{P}^T \\
&\quad - q(1 - q) \mathbf{E} [\text{diag}(a_{k-1}a_{k-1}^T)] F^T \tilde{P}^T \\
&\quad + q(1 - q) \tilde{P}F \mathbf{E} [a_{k-1}a_{k-1}^T] \\
&\quad - q(1 - q) \tilde{P}F \mathbf{E} [\text{diag}(a_{k-1}a_{k-1}^T)] \\
&\quad + q^2 \tilde{P}F \mathbf{E} [a_{k-1}a_{k-1}^T] F^T \tilde{P}^T \\
&\quad + q(1 - q) \tilde{P}F \mathbf{E} [\text{diag}(a_{k-1}a_{k-1}^T)] F^T \tilde{P}^T.
\end{aligned} \tag{38}$$

By re-arranging terms in (38) and observing that $\Gamma_{k-1} = \mathbf{E} [a_{k-1}a_{k-1}^T]$ and $\text{diag}(\Gamma_{k-1}) = \mathbf{E} [\text{diag}(a_{k-1}a_{k-1}^T)]$, the result in (30) follows. Additionally, from Lemma 3, it follows that

$$\begin{aligned}
\Phi_{k+1} &= \mathbf{E} [y_{k+1}y_{k+1}^T] = F \mathbf{E} [(a_k \circ x_k)(x_k \circ a_k)^T] F^T \\
&= F \left[q^2 \mathbf{E} [a_k a_k^T] + q(1 - q) \mathbf{E} [\text{diag}(a_k a_k^T)] \right] F^T \\
&= F [q^2 \Gamma_k + q(1 - q) \text{diag}(\Gamma_k)] F^T.
\end{aligned} \tag{39}$$

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 2: Although omitted in the statement of Lemma 4, the dynamics of $\Delta_k = \mathbf{E}[b_k a_k^T]$ and $\Theta_k = \mathbf{E}[z_k y_k^T]$ can also 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 $C = AXB$ (where X is an unknown matrix) can be rewritten as a set of linear equations of the form $(B^T \otimes A)x = c$, where x and c are the vectors that result from stacking the columns of matrices X and C respectively, and \otimes denotes the Kronecker product [15].

Let γ_k be the vector that results from stacking the columns of Γ_k , and G be a diagonal matrix with entries $G((l-1)n^2 + l, (l-1)n^2 + l) = 1$, $\forall l = 1, 2, \dots, n^2$, and zero otherwise. Then, using the ideas above, we can rewrite (30) as

$$\gamma_k = \Pi \gamma_{k-1}, \quad k \geq 1, \quad (40)$$

where $\Pi = [[q\tilde{P}F + (1-q)I] \otimes [q\tilde{P}F + (1-q)I] + q(1-q)([I - \tilde{P}F] \otimes [I - \tilde{P}F])G]$.

Remark 3: Let ψ_k , and ξ_k and δ_k be the vectors that result from stacking the columns of Ψ_k , Ξ_k and Δ_k respectively. Then, it is easy to see that the same recurrence relation as in (40) governs the evolution of ψ_k and ξ_k . \square

The structure and fundamental properties of the matrix Π are established in the next theorem (the proof is provided in the Appendix), 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 1: 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 $\mathcal{V} = \{1, 2, \dots, n\}$ and $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$. Let $F = [I_n \ I_n \ \dots \ I_n] \in \mathbb{R}^{n \times n^2}$, where I_n is the $n \times n$ identity matrix, and $\tilde{P} = [E_1 P^T \ E_2 P^T \ \dots \ E_n P^T]^T \in \mathbb{R}^{n^2 \times n}$, where each $E_i \in \mathbb{R}^{n \times n}$, $i \in \{1, 2, \dots, n\}$, satisfies $E_i(i, i) = 1$ and has all other entries equal to zero. Then, for any q , $0 < q \leq 1$, the matrix Π that defines (40) 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 VI to show that as $k \rightarrow \infty$, the random vector $v_k = y_k - \alpha z_k$, for $\alpha = \frac{\sum_{j=1}^n y_0(j)}{\sum_{j=1}^n z_0(j)}$, converges to $v = 0$ almost surely. Thus, as $k \rightarrow \infty$, and whenever $z_k(i) > 0$, each node i can obtain an estimate of α by calculating $y_k(i)/z_k(i)$.

Lemma 5: Define $w_k = a_k - \alpha b_k$, with $\alpha = \frac{\sum_{j=1}^n y_0(j)}{\sum_{j=1}^n z_0(j)}$, and denote by χ_k the vector that

results from stacking the columns of $\mathcal{X}_k := \mathbf{E}[w_k w_k^T]$. Then, it follows that $\chi_k = \Pi \chi_{k-1}$ with $\chi_0 = \gamma_0 + \alpha^2 \psi_0 - \alpha(\xi_0 + \delta_0)$ and $\sum_{l=1}^{n^4} \chi_0(l) = 0$.

Proof: Since $\mathcal{X}_k := \mathbf{E}[w_k w_k^T] = \mathbf{E}[a_k a_k^T] + \alpha^2 \mathbf{E}[b_k b_k^T] - \alpha(\mathbf{E}[a_k b_k^T] + \mathbf{E}[b_k a_k^T]) = \Gamma_k + \alpha^2 \Psi_k - \alpha(\Xi_k + \Delta_k)$, it follows that $\chi_k = \gamma_k + \alpha^2 \psi_k - \alpha(\xi_k + \delta_k)$. From (40) 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 $\chi_k = \Pi \gamma_{k-1} + \alpha^2 \Pi \psi_{k-1} - \alpha(\Pi \xi_{k-1} + \Pi \delta_{k-1}) = \Pi(\gamma_{k-1} + \alpha^2 \psi_{k-1} - \alpha(\xi_{k-1} + \delta_{k-1})) = \Pi \chi_{k-1}$. In Lemma 4, 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

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

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 $\alpha = \frac{\sum_{j=1}^n y_0(j)}{\sum_{j=1}^n z_0(j)}$, it follows that $\sum_{l=1}^{n^4} \chi_0(l) = \sum_{l=1}^{n^4} (\gamma_0(l) + \alpha^2 \psi_0(l) - \alpha(\xi_0(l) + \delta_0(l))) = 0$. ■

VI. CONVERGENCE ANALYSIS

In this section, we establish that for each j , there is an infinite sequence of time instants $\tau = \{t_1, t_2, \dots\}$ such that $z_j[k] > C$, $\forall k \in \tau$, and almost surely $\lim_{n \rightarrow \infty} \left| \frac{y_j[t_n]}{z_j[t_n]} - \alpha \right| = 0$ where $\alpha = (\sum_{j=1}^n y_0(j)) / (\sum_{j=1}^n z_0(j))$. This result is formally stated in Theorems 3 and 6, but in order to prove this theorems, we first need to establish several ancillary results.

Theorem 2: Let y_k and z_k be the random vectors that result from iterations (20)–(21) and (22)–(23). Define $v_k = y_k - \alpha z_k$, where $\alpha = \frac{\sum_{j=1}^n y_0(j)}{\sum_{j=1}^n z_0(j)}$. Then, $\|v_k\|_\infty \rightarrow 0$ almost surely. Furthermore, for every j , $v_k(j) \rightarrow 0$ as $k \rightarrow \infty$ almost surely.

Proof: The result follows from the first Borel-Cantelli Lemma [16, Theorem 7.3.10]. For all $k \geq 0$ and all $\epsilon > 0$, define the event $E_k(\epsilon) = \{\|v_k\|_\infty > \epsilon\}$. We will first establish an upper bound on $\sum_{k=0}^{\infty} \Pr\{E_k(\epsilon)\}$ by noting that $\Pr\{E_k(\epsilon)\} = \Pr\{\|v_k\|_\infty > \epsilon\} \leq \frac{\mathbf{E}[\|v_k\|_\infty]}{\epsilon}$, thus $\sum_{k=0}^{\infty} \Pr\{E_k(\epsilon)\} \leq \frac{1}{\epsilon} \sum_{k=0}^{\infty} \mathbf{E}[\|v_k\|_\infty] \leq \frac{1}{\epsilon} \sum_{k=0}^{\infty} \mathbf{E}[\|v_k\|_2]$. Note that $\mathbf{E}[\|v_k\|_2] = (\mathbf{E}[v_k^T v_k])^{1/2} = (\text{trace}(\mathbf{E}[v_k v_k^T]))^{1/2} = (\text{trace}(\mathbf{E}[y_k y_k^T]) + \alpha^2 \text{trace}(\mathbf{E}[z_k z_k^T]) - 2\alpha \text{trace}(\mathbf{E}[y_k z_k^T]))^{1/2}$.

We will next show that $\mathbf{E}[\|v_k\|_2] \rightarrow 0$ as $k \rightarrow \infty$ geometrically fast. Using Lemma 4, it can be

established that $\mathbf{E}[v_k v_k^T] = \mathbf{E}[y_k y_k^T] + \alpha^2 \mathbf{E}[z_k z_k^T] - \alpha(\mathbf{E}[y_k z_k^T] + \mathbf{E}[z_k y_k^T]) = F[q^2 \mathcal{X}_{k-1} + q(1 - q)\text{diag}(\mathcal{X}_{k-1})]F^T$, where $\mathcal{X}_{k-1} = \mathbf{E}[w_{k-1} w_{k-1}^T]$ as defined in Lemma 5, 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 1, we showed that Π 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). Since Π is a column stochastic matrix, and Lemma 5 established that $\sum_{l=1}^{n^4} \chi_0(l) = 0$, it follows that $\sum_{l=1}^{n^4} \chi_k(l) = 0, k \geq 0$, and therefore $\lim_{k \rightarrow \infty} \chi_k(l) = 0, \forall l$. Additionally, the convergence of $\chi_k = \Pi \chi_{k-1}$ is geometric with a rate of convergence given by $r = |\lambda_2| + \epsilon$ where λ_2 is the eigenvalue of Π of second largest modulus, which satisfies $|\lambda_2| < \lambda_1 = 1$, and $\epsilon > 0$ is some small constant such that $r < 1$ (see, e.g., [17]). Thus, we have established that $\chi_k(l) \rightarrow 0, \forall l$, geometrically fast, from where it follows that all the entries of $\mathbf{E}[v_k v_k^T]$ go to zero also geometrically fast. Therefore, the trace($\mathbf{E}[v_k v_k^T]$) also goes to zero geometrically fast, so that $\mathbf{E}[\|v_k\|_2]$ also goes to 0 geometrically fast. It immediately follows that $\sum_{k=0}^{\infty} \mathbf{E}[\|v_k\|_2] < \infty$ and therefore $\sum_{k=0}^{\infty} \Pr\{\|v_k\|_{\infty} \geq \epsilon\} < \infty$. Then, from the first Borel-Cantelli Lemma, we have that $\Pr\{\|v_k\|_{\infty} \geq \epsilon \text{ i. o.}\} = 0$. Finally, since $\|v_k\|_{\infty} \geq |v_k(j)|, \forall j$, then, $\Pr\{\|v_k\|_{\infty} \geq \epsilon\} \geq \Pr\{|v_k(j)| \geq \epsilon\}, \forall j$, and thus, $\sum_{k=0}^{\infty} \Pr\{|v_k(j)| \geq \epsilon\} < \infty, \forall j$. Therefore, by Theorem 7.2.4.c of [16], $v_k(j) \rightarrow 0, \forall j$, almost surely. \blacksquare

Corollary 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 .

Proof: It is well-known (see, e.g., [18, 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 $L = \lim_{k \rightarrow \infty} \Pi^k$ is a rank-one column stochastic matrix. It then follows that $\|(\Pi^k - L)\chi_0\|_{\infty} \leq \|\Pi^k - L\|_{\infty} \|\chi_0\|_{\infty} \leq C \|\chi_0\|_{\infty} k^{m_2-1} |\lambda_2|^k$. Since Π is column stochastic and $\sum_{l=1}^{n^4} \chi_0(l) = 0$, it follows that $L\chi_0 = 0$, thus $\|\mathbf{E}[w_k w_k^T]\|_{\infty} = \|\chi_k\|_{\infty} = \|\Pi^k \chi_0\|_{\infty} \leq C \|\chi_0\|_{\infty} k^{m_2-1} |\lambda_2|^k$. Then, since $\mathbf{E}[v_{k+1} v_{k+1}^T] = F[q^2 \mathbf{E}[w_k w_k^T] + q(1 - q)\text{diag}(\mathbf{E}[w_k w_k^T])]F^T$, and by recalling that all the entries of F are either 0 or 1, it is easy to check that $\|\mathbf{E}[v_{k+1} v_{k+1}^T]\|_{\infty} \leq \|\mathbf{E}[w_k w_k^T]\|_{\infty}$, thus $\|\mathbf{E}[v_{k+1} v_{k+1}^T]\|_{\infty} \leq C \|\mathbf{E}[w_0 w_0^T]\|_{\infty} k^{m_2-1} |\lambda_2|^k$. Now, it is easy to establish that $\|\mathbf{E}[v_{k+1} v_{k+1}^T]\|_{\infty} \geq \frac{1}{n^2} \text{trace}(\mathbf{E}[v_{k+1} v_{k+1}^T]) = \frac{1}{n^2} (\mathbf{E}[\|v_{k+1}\|_2])^2 \geq \frac{1}{n^2} (\mathbf{E}[\|v_{k+1}\|_{\infty}])^2$, from where it follows that $\mathbf{E}[\|v_{k+1}\|_{\infty}] \leq n C^{1/2} (\|\mathbf{E}[w_0 w_0^T]\|_2)^{1/2} k^{\frac{m_2-1}{2}} |\lambda_2|^{k/2}$. Define $C' = \frac{n}{\epsilon} C^{1/2}$, for some $\epsilon > 0$; then, the result follows from the fact $\Pr\{\|v_{k+1}\|_{\infty} > \epsilon\} \leq$

$$\frac{\mathbf{E}[\|v_{k+1}\|_\infty]}{\epsilon}.$$

Theorem 2 established 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 $y_k - \alpha z_k = 0$ with probability one. Thus, in this limiting case, each node j can calculate the value of α by taking the ratio $\frac{y_k(j)}{z_k(j)}$, as long as $z_k(j) \neq 0$. Also, Corollary 1 establishes the number of iterations k after which y_k and z_k will satisfy $|y_k - \alpha z_k| \leq \epsilon$, for a given accuracy level ϵ , with some desired probability. This desired probability goes to 1 with a geometric rate governed by the $|\lambda_2|^{1/2}$, where λ_2 is the eigenvalue of Π of second largest modulus. Note that, as also evidenced by the simulations provided in Fig. 4 and Fig. 5, the vectors y_k and z_k do not converge in any way; however, the values y_k and z_k become perfectly aligned almost surely, allowing each node j to calculate $\alpha = \frac{y_k(j)}{z_k(j)}$. The only problem here arises when $y_k(j)$ and $z_k(j)$ have both value zero, which does not constitute a violation of $y_k - \alpha z_k = 0$, but clearly does not allow node j to calculate the desired value α . The next two theorems essentially establish that $z_k(j)$, $j = 1, 2, \dots, n$, will be greater than zero (in fact, greater than a constant C that will be specified) infinitely often. Note that, in subsequent developments, $z_k(j)$ is denoted by $z_j[k]$ in order to remain close to the notation in (12)–(14).

Theorem 3: Consider a strongly connected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ and the iteration in (12)–(14), with $x_{ji}[k]$, $(j, i) \in \mathcal{E}$, $k = 0, 1, 2, \dots$, as defined in (1). For every $j = 1, 2, \dots, n$, define the event $E_k^j = \{z_j[kn] \geq C\}$, $k \geq 1$, where $C = \frac{n}{(n+m)(\mathcal{D}_{\max}^+)^{n-1}}$, $\mathcal{D}_{\max}^+ = \max_{j \in \mathcal{V}} \{\mathcal{D}_j^+\}$, $n = |\mathcal{V}|$, and $m = |\mathcal{E}|$. Let ζ_k^j denote the indicator of the event E_k^j , $k \geq 1$, i.e., $\zeta_k^j = 1$ whenever E_k^j occurs and $\zeta_k^j = 0$ otherwise. Then, $\Pr\{z_j[(k+1)n] \geq C \mid \zeta_k^j, \zeta_{k-1}^j, \dots, \zeta_1^j\} \geq q^n$, $\forall j$.

Proof: Since for every j , $z_j[0] > 0$, $\forall j$, it follows from (22)–(23) that, for every j , $z_j[k] \geq 0$, $k \geq 0$. Then, the total mass \mathcal{M}_{k+1} in the system, defined as $\mathcal{M}_{k+1} := \sum_{j=1}^n z_j[k+1] + \sum_{(j,i) \in \mathcal{E}} (\sigma_{ji}[k] - \tau_{ji}[k-1])(1 - x_{ji}[k])$, satisfies $\mathcal{M}_{k+1} = n$, for all $k = 0, 1, 2, \dots$. [This follows from the fact that $\mathcal{M}_0 = \sum_{j=1}^n z_j[0] = n$ and the observation that $\mathcal{M}_{k+1} := \sum_{j=1}^n z_j[k+1] + \sum_{(j,i) \in \mathcal{E}} (\sigma_{ji}[k] - \tau_{ji}[k-1])(1 - x_{ji}[k]) = \sum_{(j,i) \in \mathcal{E}} (\sigma_{ji}[k] - \tau_{ji}[k-1])x_{ji}[k] + \sum_{(j,i) \in \mathcal{E}} (\sigma_{ji}[k] - \tau_{ji}[k-1])(1 - x_{ji}[k]) = \sum_{(j,i) \in \mathcal{E}} (\sigma_{ji}[k] - \tau_{ji}[k-1]) = \sum_{j=1}^n z_j[k] + \sum_{(j,i) \in \mathcal{E}} (\sigma_{ji}[k-1] - \tau_{ji}[k-2])(1 - x_{ji}[k-1])$ which is equal to \mathcal{M}_k .] The definition of \mathcal{M}_{k+1} involves the summation of $n + m$ nonnegative quantities, namely, $z_j[k+1]$ for $j = 1, 2, \dots, n$ and $m_{ji}[k+1] := (\sigma_{ji}[k] - \tau_{ji}[k-1])(1 - x_{ji}[k])$ for $(j, i) \in \mathcal{E}$. We can think of the \mathcal{M}_{k+1} 's as follows: $z_j[k+1]$ is the mass at node j , whereas $m_{ji}[k+1]$ is the mass waiting to get transferred to node j from node i . Since all of these quantities are nonnegative, at least one of them is larger or equal to $\frac{n}{n+m}$.

Regardless of whether this quantity is associated with a node (say j^*) or a link (say (j^*, i^*)), this mass has at least one way of reaching any node j of interest in graph \mathcal{G} via a path of length at most $n - 1$: in particular, there is at least one path of length at most $n - 1$ from node j^* to j and all the links in this path have weight at least $\frac{1}{\mathcal{D}_{\max}^+}$. If all these links are activated, which occurs with probability q^{n-1} (q^n in the case of link (j^*, i^*) because the mass needs to first transfer to j^*), then a fraction $(\frac{1}{\mathcal{D}_{\max}^+})^{n-1}$ of the mass transfers to j in at most n steps. It follows that $\Pr\{z_j[(k+1)n] \geq C \mid \zeta_k, \zeta_{k-1}, \dots, \zeta_1\} \geq q^n$. ■

Given a sequence of events $E_1, E_2, \dots, E_n, \dots$ defined on some probability space, the next theorem states the 1912 Borel criterion for establishing whether the event that infinitely many of the E_k occur, denoted by $\{E_k \text{ i.o.}\}$, happens with probability one or zero (see, e.g., [19], [20]). This is used in Theorem 5 to establish that $E_k^j = \{z_j[kn] \geq C\}$, $k \geq 1$, occurs infinitely often for all nodes j . Finally, in Theorem 6, we show that, whenever $z_j[k] \geq C$, each j can obtain an estimate of α by calculating $y_j[k]/z_j[k]$ and this estimate will almost surely converge to α .

Theorem 4 (Borel, 1912 [19]): Let $\{E_k\}$, $k = 1, 2, \dots$, be a sequence of events defined on some probability space. Let ζ_k be the indicator function of event E_k and let $\Pr\{E_k \mid \zeta_{k-1}, \zeta_{k-2}, \dots, \zeta_1\}$ denote the conditional probability of event E_k given the outcome of previous trials. If $0 < p'_k \leq \Pr\{E_k \mid \zeta_{k-1}, \zeta_{k-2}, \dots, \zeta_1\} \leq p''_k$ for every k , whatever $\zeta_1, \zeta_2, \dots, \zeta_{k-1}$, then i) $\Pr\{E_k \text{ i.o.}\} = 0$ if $\sum_{k=1}^{\infty} p''_k < 1 \leq \infty$, and ii) $\Pr\{E_k \text{ i.o.}\} = 1$ if $\sum_{k=1}^{\infty} p'_k = \infty$.

Theorem 5: Consider the iteration in (12)–(14). For every $j = 1, 2, \dots, n$, define the event $E_k^j = \{z_j[kn] \geq C\}$, $k \geq 1$, where $C = \frac{n}{(n+m)(\mathcal{D}_{\max}^+)^{n-1}}$, $\mathcal{D}_{\max}^+ = \max_{j \in \mathcal{V}}\{\mathcal{D}_j^+\}$, $n = |\mathcal{V}|$, and $m = |\mathcal{E}|$. Then, $\Pr\{E_k \text{ i.o.}\} = 1$.

Proof: Theorem 3 established that, for every j , $\Pr\{z_j[(k+1)n] \geq C \mid \zeta_k^j, \zeta_{k-1}^j, \dots, \zeta_1^j\} \geq q^n$. Define $p'_k = q^n$, then $\sum_{k=1}^{\infty} p'_k = \infty$, and by the second assertion of Theorem 4, it follows that, $\Pr\{E_k^j \text{ i.o.}\} = 1$, $\forall j$. ■

Theorem 6: For each j , let $\tau = \{t_1, t_2, \dots\}$ be an increasing sequence such that $z_j[k] > C$, $\forall k \in \tau$. Then, almost surely $\lim_{n \rightarrow \infty} \left| \frac{y_j[t_n]}{z_j[t_n]} - \alpha \right| = 0$.

Proof: Assume that $\alpha > 0$. Then since $z_j[k] \geq C$, $\forall k \in \tau$, whenever $y_j[t_n] - \alpha z_j[t_n] > 0$, it follows that $\frac{y_j[t_n]}{z_j[t_n]} - \alpha \leq \frac{y_j[t_n] - \alpha z_j[t_n]}{C}$. Also, in the proof of Theorem 3, we established that $\mathcal{M}_k = n$, $k \geq 0$, from where it follows that $z_j[t_n] \leq n$, therefore, whenever $y_j[t_n] - \alpha z_j[t_n] \leq 0$, $\frac{y_j[t_n]}{z_j[t_n]} - \alpha \geq \frac{y_j[t_n] - \alpha z_j[t_n]}{n}$. In Theorem 2, we established that $|y_j[k] - \alpha z_j[k]| \rightarrow 0$ almost surely, thus, almost surely, the subsequence $|y_j[t_n] - \alpha z_j[t_n]| \rightarrow 0$. Then, since $C < n$, almost surely,

$\lim_{n \rightarrow \infty} \left| \frac{y_j[t_n]}{z_j[t_n]} - \alpha \right| \leq \lim_{n \rightarrow \infty} \left| \frac{y_j[t_n] - \alpha z_j[t_n]}{C} \right| = 0$. The proof when $\alpha \leq 0$ is similar and therefore omitted. ■

VII. CONCLUDING REMARKS

In this paper, we proposed a method to ensure robustness of a class of linear-iterative distributed algorithms against unreliable communication links that may drop packets. As illustrated, this class of algorithms can be used to coordinate the action of DERs to provide reactive power support for voltage control. Other possible applications include the coordination of plug-in hybrid electric vehicles to provide active power for up and down regulation services, and the control of demand response resources in demand response programs.

APPENDIX

Proof of Theorem 1: We show first column stochasticity of matrix Π . Let $C = q\tilde{P}F + (1-q)I$ and $D = I - \tilde{P}F$, so that $\Pi = C \otimes C + q(1-q)(D \otimes D)G$. We will establish that $C \otimes C$ is column stochastic and also show that the column sums of $D \otimes D$ are all zero. By construction, the entries of the i^{th} column of $\tilde{P} \in \mathbb{R}^{n^2 \times n}$ are all zero, with the possible exception of the ones indexed by $((i-1)n+j, i)$, $i, j = 1, 2, \dots, n$, each of which corresponds to the (j, i) entry of matrix P . Then, it follows that $\sum_{l=1}^{n^2} \tilde{P}_{li} = \sum_{j=1}^n P_{ji} = 1, \forall i = 1, 2, \dots, n^2$. The matrix $\tilde{P}F \in \mathbb{R}^{n^2 \times n^2}$ is also column stochastic by construction, as it results from horizontally concatenating n times the matrix \tilde{P} , i.e., $\tilde{P}F = [\tilde{P} \tilde{P} \dots \tilde{P}]$; therefore, the matrix C is also column stochastic. The Kronecker product of C with itself, results in an $n^4 \times n^4$ block matrix of the form $C \otimes C = [C_1 \ C_2 \dots \ C_{n^2}]$, where $C_j = [c_{1j}C^T \ c_{2j}C^T \ \dots \ c_{n^2j}C^T]^T$. Then, it follows that the sum of the entries of the l^{th} column of C_j is $\sum_{m=1}^{n^4} C_j(m, l) = (\sum_{i=1}^{n^2} c_{ij})(\sum_{r=1}^{n^2} c_{rl})$. Since $\sum_{i=1}^{n^2} c_{ij}$ and $\sum_{r=1}^{n^2} c_{rl}$ are the sum of the entries of the j^{th} and l^{th} columns of $C = q\tilde{P}F + (1-q)I_{n^4}$, it follows that $\sum_{m=1}^{n^4} C_j(m, l) = 1$; thus, $C \otimes C$ is column stochastic.

Since $\tilde{P}F$ is column stochastic, the column-sums of $D = I - \tilde{P}F$ are zero. The Kronecker product of D with itself is of the form $D \otimes D = [D_1 \ D_2 \dots \ D_{n^2}]$, where $D_j = [d_{1j}D^T \ d_{2j}D^T \ \dots \ d_{n^2j}D^T]^T$. Using similar arguments as above, it follows that $\sum_{m=1}^{n^4} D_j(m, l) = (\sum_{i=1}^{n^2} d_{ij})(\sum_{r=1}^{n^2} d_{rl}) = 0$, which implies that the column-sums of $D \otimes D$ are zero. The only thing left to establish that Π is column stochastic is to show that all entries of Π are nonnegative (from where it immediately follows that $\Pi = C \otimes C + q(1-q)(D \otimes D)G$ is column stochastic).

We argue nonnegativity of Π as follows: due to the sparsity structure of G , the only nonzero entries of $(D \otimes D)G$ will be in columns $(k-1)n^2 + k$, $k = 1, 2, \dots, n^2$; thus except for entries in these columns, the entries of Π will be identical to the corresponding entries in $C \otimes C$. From the structure of $\tilde{P}F$, entries of $C \otimes C$ and $q(1-q)(D \otimes D)$ can, respectively, take one of the following three forms: i) $(qp_{ij} + (1-q))(qp_{lm} + (1-q))$ and $q(1-q)(1-p_{ij})(1-p_{lm})$; ii) $qp_{ij}(qp_{lm} + (1-q))$ and $-q(1-q)p_{ij}(1-p_{lm})$; or iii) $q^2p_{ij}p_{lm}$ and $q(1-q)p_{ij}p_{lm}$, where $p_{ij} \geq 0$ and $p_{lm} \geq 0$ are the (i, j) and (l, m) entries of matrix P . For i), the corresponding entry of Π is $(qp_{ij} + (1-q))(qp_{lm} + (1-q)) + q(1-q)(1-p_{ij})(1-p_{lm}) = qp_{ij}p_{lm} + (1-q)$, satisfying $0 \leq qp_{ij}p_{lm} + (1-q) \leq 1$. For ii), the corresponding entry of Π is $qp_{ij}(qp_{lm} + (1-q)) - q(1-q)p_{ij}(1-p_{lm}) = qp_{ij}p_{lm}$ satisfying $0 \leq qp_{ij}p_{lm} \leq 1$. For iii), the entry of Π is $q^2p_{ij}p_{lm} + q(1-q)p_{ij}p_{lm} = qp_{ij}p_{lm}$, satisfying $0 \leq qp_{ij}p_{lm} \leq 1$.

To prove the second assertion, we will show first that matrix $\tilde{P}F$ can be written via a permutation of its indices in the form

$$\begin{bmatrix} U & V \\ 0 & W \end{bmatrix}, \quad (41)$$

where U is an irreducible column stochastic matrix and $\lim_{k \rightarrow \infty} W^k = 0$. Since $\tilde{P}F$ is column stochastic, we can assume that it corresponds to the weight matrix of some graph $\tilde{\mathcal{G}} = \{\tilde{\mathcal{V}}, \tilde{\mathcal{E}}\}$. We will show that this graph has a single recurrent class plus a few transient states, from which the decomposition of $\tilde{P}F$ in (41) follows. Let $\tilde{\mathcal{V}} = \{(1, 1), (2, 1), \dots, (n, 1), (1, 2), (2, 2), \dots, (n, 2), \dots, (n, n-1), (1, n), (2, n), \dots, (n, n)\}$. From the structure of $\tilde{P}F$, for any node $(i, j) \in \tilde{\mathcal{V}}$, one-step transitions out of (i, j) are to nodes of the form (m, i) , with $i \in \mathcal{N}_m^-$, where \mathcal{N}_m^- is the set in-neighbors of node m in the graph \mathcal{G} (with weight matrix P). Also, from the structure of $\tilde{P}F$, there are possibly several rows of $\tilde{P}F$ with all entries equal to zero, which means that a node (i, j) that is associated with such row cannot be reached from any other node; however, as already argued, from nodes of the form (i, j) , it is possible to reach nodes of the form (m, i) , where $i \in \mathcal{N}_m^-$. The nodes corresponding to rows with all entries being zero are transient. Also, the only nonzero diagonal entries of $\tilde{P}F$ correspond to diagonal entries of P , which are smaller than one; thus nodes that are not reachable from any other node cannot be disconnected.

Next we will show that all non-transient nodes form a single recurrent class (as already argued all nonzero diagonal entries are strictly smaller than one which means there cannot be

absorbing nodes). This follows from the fact that the graph \mathcal{G} is strongly connected, which means that for any $j, l \in \mathcal{V}$, there exists a path between j and l . Let i_1, i_2, \dots, i_t denote the nodes traversed along the path between j and l . We will show next that for any two non-transient nodes $(i, j), (r, l) \in \tilde{\mathcal{V}}$ there exists a path. As already argued, from (i, j) one can reach in a single hop any node of the form (m, i) , where m is a neighbor of node i in the graph \mathcal{G} . Since i_1 is the first node traversed in the path between j and l , it follows that $(i_1, i) \in \tilde{\mathcal{V}}$ can be reached in one step from (i, j) . By repeatedly using this argument, it follows that the sequence of nodes $(i_1, i), (i_2, i_1), \dots, (i_t, i_{t-1}), (r, i_t)$ forms a path between (i, j) and (r, l) , which means that any non-transient node can be reached by any other non-transient node; thus, the set of non-transient nodes forms a single recurrent class. Clearly, the vertex set $\tilde{\mathcal{V}}$ can be decomposed into a single recurrent class and possibly several transient nodes. By re-ordering the nodes, $\tilde{P}F$ can be rewritten as in (41) (see, e.g., [17, p. 126]). Furthermore, since Q in (41) is irreducible, it follows that $qQ + (1 - q)I$ (where I is the identity matrix) is primitive. Then, $C = q\tilde{P}F + (1 - q)I$ has a unique largest eigenvalue of value one, i.e., $\lambda_1 = 1$, and $1 > |\lambda_2| \geq \dots \geq |\lambda_{n^2}|$. Let $\sigma(C) = \{\lambda_1, \lambda_2, \dots, \lambda_{n^2}\}$. Then, $\sigma(C \otimes C) = \{\lambda_i \lambda_j, i = 1, \dots, n, j = 1, \dots, n\}$, including algebraic multiplicities in both cases [15, p. 245]. Since $\lambda_1 = 1$ is unique (multiplicity one) and $|\lambda_i| < 1, i = 2, \dots, n^2$, it follows that the eigenvalue of $C \otimes C = [q\tilde{P}F + (1 - q)I] \otimes [q\tilde{P}F + (1 - q)I]$ of largest magnitude also takes value 1 and is unique. Since $C \otimes C$ is column stochastic and $\lambda_1 = 1$ is unique, we know that either $C \otimes C$ is also primitive or it can be decomposed (through index permutation) to a matrix of the form in (41) for some U', V' , and W' such that where U is a primitive matrix and $\lim_{k \rightarrow \infty} W^k = 0$ [17]. Due to the sparsity structure of G , the only nonzero entries of $(D \otimes D)G$ will be in columns $(k - 1)n^2 + k, k = 1, 2, \dots, n^2$, thus except for entries in the aforementioned columns, the nonzero entries of Π are the same as those in $C \otimes C$. For all other columns in Π (that include nonzero entries in $(D \otimes D)G$), it was shown above that the nonzero entries of Π are strictly positive, thus Π has the same sparsity structure as $C \otimes C$, which means that Π can be decomposed in the form of (41) (for some L'', M'', N''), and the resulting upper-right block is primitive. Thus, Π has a unique largest eigenvalue at one.

REFERENCES

- [1] M. Baran and I. El-Markabi, "A multiagent-based dispatching scheme for distributed generators for voltage support on distribution feeders," *IEEE Trans. Power Syst.*, vol. 22, no. 1, pp. 52–59, Feb. 2007.

- [2] K. Rogers *et al.*, “An authenticated control framework for distributed voltage support on the smart grid,” *IEEE Trans. Smart Grid*, vol. 1, no. 1, pp. 40–47, June 2010.
- [3] A. D. Domínguez-García and C. N. Hadjicostis, “Distributed algorithms for control of demand response and distributed energy resources,” in *Proc. IEEE Conference on Decision and Control*, 2011.
- [4] F. Benezit *et al.*, “Weighted gossip: Distributed averaging using non-doubly stochastic matrices,” in *Proc. IEEE International Symposium on Information Theory*, June 2010.
- [5] B. A. Robbins and A. D. Domínguez-García, “Distributed algorithms for voltage control in electrical networks,” Power Systems Engineering Research Center, PSERC Publication 11-07, September 2011.
- [6] B. A. Robbins, A. D. Domínguez-García, and C. N. Hadjicostis, “Control of distributed energy resources for reactive power support,” in *Proc. of North American Power Symposium*, August 2011.
- [7] 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.
- [8] S. Sundaram and C. N. Hadjicostis, “Distributed function calculation and consensus using linear iterative strategies,” *IEEE J. Sel. Areas Commun.*, vol. 26, no. 4, pp. 650–660, May 2008.
- [9] R. Olfati-Saber and R. M. Murray, “Agreement problems in networks with directed graphs and switching topology,” in *Proc. of American Control Conference*, vol. 4, 2003, pp. 4123–4132.
- [10] B. Gharesifard and J. Cortés, “Distributed strategies for generating weight-balanced and doubly stochastic digraphs,” 2010. [Online]. Available: <http://arxiv.org/abs/0911.0232>
- [11] S. Patterson, B. Bamieh, and A. El Abbadi, “Distributed average consensus with stochastic communication failures,” in *Proc. IEEE CDC*, Dec. 2007, pp. 4215–4220.
- [12] F. Fagnani and S. Zampieri, “Average consensus with packet drop communication,” *SIAM Journal on Control and Optimization*, vol. 48, no. 1, pp. 102–133, 2009.
- [13] Y. Chen *et al.*, “Corrective consensus: Converging to the exact average,” in *Proc. IEEE Conference Decision and Control*, Dec. 2010, pp. 1221–1228.
- [14] T. Van Cutsem and C. Vorunas, *Voltage Stability of Electric Power Systems*. Boston, MA: Kluwer Academic Publishers, 1998.
- [15] R. Horn and C. Johnson, *Topics in Matrix Analysis*. New York, NY: Cambridge University Press, 1991.
- [16] G. Grimmett and D. Stirzaker, *Probability and Random Processes*. Oxford University Press, 1992.
- [17] E. Seneta, *Non-negative Matrices and Markov Chains*, revised printing ed. New York, NY: Springer, 2006.
- [18] R. Horn and C. Johnson, *Matrix Analysis*. New York, NY: Cambridge University Press, 1985.
- [19] E. Borel, “Sur un probleme de probabilites relatif aux fractions continues,” *Mathematische Annalen*, vol. 72, pp. 578–587, 1912.
- [20] S. W. Nash, “An extension of the Borel-Cantelli lemma,” *The Annals of Mathematical Statistics*, vol. 25, no. 1, pp. 165–167, 1954.