2011 50th IEEE Conference on Decision and Control and
European Control Conference (CDC-ECC)
Orlando, FL, USA, December 12-15, 2011

Corrective Consensus with Asymmetric Wireless Links

Yin Chen† Roberto Tron* Andreas Terzis† Rene Vidal*
yinchen@cs.jhu.edu tron@cis.jhu.edu terzis@cs.jhu.edu rvidal@jhu.edu
Computer Science Department† Center for Imaging Science*
Johns Hopkins University

***Abstract*— Consensus algorithms can be used to compute an average value across a multi-hop network in a distributed way. However, their convergence to the right value is not guaranteed in the presence of random packet losses that are common in real life low-power wireless networks. *Corrective consensus* solves this problem by using a set of auxiliary variables to compensate for the asymmetric state updates caused by packet losses. Nevertheless, one key assumption is that the *probability* of delivering a packet from node i to a neighboring node j is the same as in the reverse direction, from j to i. This assumption might be violated in real life conditions. Our main contribution is showing that corrective consensus converges to the correct average even when this assumption is removed. In addition, we provide a heuristic for modifying the weights used by corrective consensus that specifically considers the unequal probabilities, and we empirically show that this choice can lead to faster convergence.**

I. Introduction

Average consensus [17] is a distributed algorithm that computes the mean of a set of variables held locally by the nodes of a multi-hop network. Doing so might be useful, for instance, when aggregating the measurements of some physical quantity collected by a wireless sensor network, with the goal of obtaining a more accurate estimate. As such, consensus is a key component for a wide range of distributed applications such as Distributed Kalman Filtering [19], Distributed Hypothesis Testing [16], Distributed Linear Support Vector Machine [9], and Distributed Maximum Likelihood Estimation [4].

Consensus algorithms typically require symmetric packet exchanges (i.e., undirected network topology) to converge to the average value [17]. Low-power wireless networks, however, are often characterized by random and asymmetric packet losses [21]. As a consequence, the network topology should be viewed as a time-varying and generally non-balanced directed graph, which prevents convergence towards the average value [11]. Chen et al. introduced *corrective consensus*, which extends the traditional average consensus algorithms to withstand random packet losses [5]. Nevertheless, that work assumes equivalent packet reception ratios between a pair of nodes, i.e., packets can get lost at random, but the probability of delivering a packet from node i to node j is assumed equal to the probability of delivering a packet from node j to node i. Although this property generally holds [21], it may be broken due to interference between different technologies using the same radio spectrum [14].

In this paper, we remove the equal probability assumption and study the convergence behavior of corrective consensus in this setting. Our results prove that corrective consensus still converges to the correct average. Furthermore, we explore an alternative way of applying the auxiliary variables to correct state values during the corrective iterations, and empirically show that this method converges faster than the original corrective consensus algorithm in the case of unequal loss probabilities.

This paper proceeds as follows. Section II gives a brief review of related work and Section III introduces standard and corrective consensus and the terminology used in this paper. In Section IV we first show the convergence of the corrective consensus algorithm after the removal of the equal probability assumption, and then present a different way of using the auxiliary variables. Section V concludes this paper.

II. Related Work

This paper extends the work of Chen et al., which introduced *corrective consensus*, a consensus algorithm that guarantees convergence to the global average in the presence of random packet losses and consequently unbalanced weight matrices [5]. While other works considered similar settings and methodologies, they only showed convergence towards a random value which is generally not equal to the average of initial state values (see [1], [2] and references therein). More recent work applied corrective consensus to the accelerated consensus frameworks [3], [12] for achieving faster convergence rate under the same conditions [7].

To the best of our knowledge, all previous algorithms assume that the packet reception ratios for both communication directions between a pair of nodes are equal. While this assumption generally holds for wireless sensor nodes [21], there are various factors that can break this balance. For example, collocated wireless devices can generate significant interference at one sensor node and, at the same time, introduce only light interference at other nodes that are physically farther [14]. In this work we remove this assumption and allow for unequal packet reception ratios.

We assume that the network graph is still symmetric, although with asymmetric packet loss probabilities. In other words, if node i can send packets to node j, then we assume that also node j can send packets to node i with probability greater than zero. This assumption excludes directed network graphs, and ensures that, for every link, there will be a chance to recover packet drops for both directions. In addition, we only consider the method of [5] and not the accelerated

978-1-61284-799-3/11/$26.00 ©2011 IEEE

version of [7]. We will investigate extensions of our methods in these directions in the future.

III. Background

We consider a multi-hop wireless network formed by a group of N nodes, and model it as a directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$. The vertex set $\mathcal{V} = \{1, 2, \ldots, N\}$ represents the N nodes, and the edge set $\mathcal{E} = \{(i, j) : i, j \in \mathcal{V}, p_{ij} > 0\}$ consists of ordered vertex pairs. A pair $(i, j) \in \mathcal{E}$ represents the directed edge $i \leftarrow j$, which indicates that node i can receive packets from node j. We assume that packets on a link $i \leftarrow j$ can be lost at random and that this process can be modeled as Bernoulli trials with probability of success p_{ij}, which is usually referred to as *packet reception ratio* (PRR) in wireless sensor networks. We assume statistical independence across both time and different links.

On link $i \leftarrow j$, node j can repeatedly transmit the same packet n times. In this case, the *effective PRR*, i.e. the probability of delivering the packet with multiple trials, is equal to $\hat{p}_{ij} := 1 - (1 - p_{ij})^n$. As mentioned in Section I, we assume that for all $i \neq j$, if $p_{ij} > 0$, then $p_{ji} > 0$. We define the $(i, j)^{\text{th}}$ entry of the adjacency matrix $A(t)$ of graph $\mathcal{G}$ as $A_{ij}(t) = 1$ if during the t-th iteration node i receives a packet from node j, and zero otherwise. Also, $A_{ii} = 0$, $\forall i$. The in-degree of node i during iteration t is given by $d_i(t) := \sum_{j=1}^{N} A_{ij}(t)$, the degree matrix of the graph $\mathcal{G}$ is defined as $D(t) := \text{diag}(d_1(t) \cdots d_N(t))$. Then the graph Laplacian is defined as $L(t) := D(t) - A(t)$, and its eigenvalues are located within a disk centered at $\max_{i,t}(d_i(t)) + 0j$ in the complex plane with the radius equal to $\max_{i,t}(d_i(t))$, due to Gershgorin's theorem [10].

A. Standard Consensus

Average consensus assumes that each node $i \in \mathcal{V}$ holds a value $z_i \in \mathbf{R}$ and that the goal is to compute the average $\overline{z} = \sum_{i=1}^{N} z_i / N$ using distributed linear iterations [17]. To do so, each node i defines $x_i(t)$ as the local state variable, initializes it as $x_i(0) = z_i$, and iteratively updates its value with a weighted average of its neighbors' state variables. It can be shown that $x_i(t)$ asymptotically converges to $\overline{z}$ under certain conditions [13], [15], [17], [18].

Specifically, each node i sends its state variable $x_i(t)$ to its one-hop neighbors during each iteration. After receiving their neighbors' states, nodes update their state variables as

$$x_i(t+1) = x_i(t) + \sum_{j=1, j\neq i}^{N} W_{ij}(t)\big(x_j(t) - x_i(t)\big), \quad (1)$$

where $W(t) \in \mathbf{R}^{N\times N}$ is the weight matrix defined as

$$W(t) := I - \epsilon L(t). \quad (2)$$

Since links are subject to random packet losses, $W(t)$ is a random matrix and we denote its expectation as $\mathbf{E}(W)$. Assuming graph $\mathcal{G}$ is connected, we can select the constant ϵ to satisfy the constraint $0 < \epsilon < 1/\max_{i,t}(d_i(t))$, such that the second largest eigenvalue of $\mathbf{E}(W)$ is smaller than 1 in magnitude. As a result, the consensus algorithm shown in (1) will converge to some common value, i.e., $\lim_{t\to\infty} x(t) = \alpha\mathbf{1}$ [20]. When the weight matrix is balanced, i.e., when $\mathbf{1}^T W(t) = \mathbf{1}^T$, then the nodes converge to the correct average $\alpha = \overline{z}$. In general, however, $\alpha \neq \overline{z}$ because $W(t)$ is not always balanced due to the random packet losses [8].

B. Corrective Consensus

The goal of *corrective consensus* is to converge to the consensus value $\overline{z}$ even if $W(t)$ is not always balanced [5]. To do so, corrective consensus introduces a set of auxiliary variables $\phi_{ij}(t)$ on each node i, and updates them as follows:

$$\phi_{ij}(t{+}1) = \phi_{ij}(t) {+} W_{ij}(t)\big(x_j(t) {-} x_i(t)\big), \;\; \phi_{ij}(0) = 0. \quad (3)$$

Comparing with (1), one can see that $\phi_{ij}(t)$ represents the amount of change that node i has made to its state variable $x_i(t)$, due to the past packet exchanges with neighbor node j.

From (3), we have $\phi_{ij} + \phi_{ji} = 0$ if $W_{ij}(t) = W_{ji}(t)$ holds for all t. This happens when the links between node i and j have had only symmetric packet losses. In general, however, it can happen that node j received $x_i(t)$ but $x_j(t)$ was lost, i.e., $W_{ji}(t) \neq W_{ij}(t)$, and consequently $\phi_{ij} + \phi_{ji} \neq 0$, which indicates a possible drift of the global average value of all the state variables away from $\overline{z}$.

In corrective consensus, the nodes first start with the standard consensus iterations, shown in (1), and every k such iterations perform one corrective iteration, as follows

$$x_i(k+1) = x_i(k) - \textstyle\sum_{j=1}^{N} \Delta_{ij}(k)/2, \quad (4)$$

$$\phi_{ij}(k+1) = \phi_{ij}(k) - \Delta_{ij}(k)/2, \quad (5)$$

where we define $\Delta_{ij}(t) = \phi_{ij}(t) + \phi_{ji}(t)$, which is equal to the amount of bias accumulated between node i and j. At a high level, each node i periodically receives $\phi_{ji}(t)$ from its neighbors and uses them to calculate $\Delta_{ij}(t)$, and adjusts its state variable $x_i(t)$ accordingly.

Note that during some iterations, node i may not have access to ϕ_{ji} due to packet losses, and consequently cannot compute Δ_{ij}. When this happens, the corresponding terms in (4) and (5) are omitted.

Finally, it is assumed that each node will try to repeatedly send n times each packet containing $x_i(t)$ in a standard iteration and m times each packet containing $\phi_{ij}(t)$ in a corrective iteration. When the equal probability assumption $p_{ij} = p_{ji}$ holds, corrective consensus is shown to converge to $\overline{z}$, given appropriate values of m and n [5].

IV. Analysis for Unequal Packet Reception Ratios

In this section, we first analyze the convergence properties of corrective consensus when the equal probability ($p_{ij} = p_{ji}$) assumption is removed. We then present an alternative way of dividing the Δ_{ij} and show that it leads to faster convergence in a two-node network. Last, we show simulation results for different dividing rules in a 10-node network.

A. Convergence of Corrective Consensus

Chen et al. proved that corrective consensus converges to the correct average when $p_{ij} = p_{ji}$, $\forall i, j \in \mathcal{V}$ [6]. When this assumption is removed, one can still employ the iterations given by (1), (3), (4) and (5). However, some important properties that were critical for the theoretical analysis are not valid under the new assumption (unequal probability). For instance, we now have $\mathbf{E}(W_{ij}(t)) \neq \mathbf{E}(W_{ji}(t))$.

Fortunately, the results of [6] can be extended to this new case. Notice that if the iterations of corrective consensus converge to a single point, then this point must satisfy $x = \bar{z}\mathbf{1}$, $\Delta_{ij} = 0$, $\forall (i,j) \in \mathcal{E}$. Therefore, if we define the sequence $\tilde{x}(t) := x(t) - \frac{1}{N}\mathbf{1}\mathbf{1}^T x(t)$, it will be sufficient to show that $\lim_{t\to\infty} \tilde{x}(t) = 0$, as it can be deduced from the following.

Theorem 1. *Let $u \in \{1, 2, \ldots\}$ indicate the number of rounds of k standard plus one corrective consensus iterations (i.e., each round has $k+1$ iterations). With appropriate number of repeated packet transmissions n and m, there exist constants $0 < c \leq b < 1$ such that*

$$\mathbf{E}\left(\|\tilde{x}((k+1)u)\|\right) \leq cb^{u-1}\mathbf{E}\left(\|\tilde{x}(0)\|\right). \tag{6}$$

The constants c and b depend only on the packet reception ratios (i.e., p_{ij}) of the links, and the number of repeated transmissions n and m for standard and corrective iteraitons respectively. With sufficiently big n and m, the constants c and b can satisfy the inequality $0 < c \leq b < 1$. It follows that $\|\tilde{x}((k+1)u)\| \to 0$ as $u \to \infty$ almost surely.

The result follows by modifying some of the proofs in [6]. For the exact definition of c and b, and for the details on the proof we refer the reader to the Appendix and [6].

B. Dividing Δ_{ij} in Proportion to Packet Reception Ratio

Corrective consensus, as shown in (4) and (5), divides the Δ_{ij} into halves to amend the state values. This can be seen as letting each node take 50% responsibility of the accumulated error, which is intuitively the appropriate approach when $p_{ij} = p_{ji}$, because nodes i and j are equally likely to cause asymmetric state updates. Following the same argument, one could expect that when $p_{ij} \neq p_{ji}$, it would be more appropriate to divide Δ_{ij} in proportion to the packet reception ratios. For example, instead of dividing in half, we propose the following heuristic

$$x_i(k+1) = x_i(k) - \sum_{j=1}^{N} \frac{p_{ji}}{p_{ij}+p_{ji}} \Delta_{ij}(k), \tag{7}$$

$$\phi_{ij}(k+1) = \phi_{ij}(k) - \frac{p_{ji}}{p_{ij}+p_{ji}} \Delta_{ij}(k). \tag{8}$$

In the above, between a pair of nodes i and j, the node that has a higher outgoing probability is assumed to take a greater responsibility for the accumulated errors. The rationale is as follows. Suppose that node i has a higher outgoing probability, i.e., $p_{ji} > p_{ij}$. Then it is more likely that x_i is delivered to node j and x_j gets updated than vice-versa. In other words, when asymmetric updates occur such that $\Delta_{ij} \neq 0$, it is more likely that x_j was updated but x_i was not, and therefore x_i should be "pulled" towards the average value using Δ_{ij}. This rationale is illustrated in more detail via a two-node example in next section.

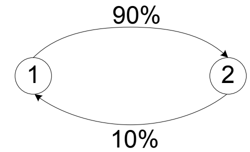

Fig. 1. A two-node network with highly asymmetric packet reception ratios.

C. A Two-node Example

In this simple example, the network has only two nodes and the packet reception ratios differ significantly, as shown in Figure 1. Specifically, the probability of successful delivery of a packet through the link $2 \leftarrow 1$ is $p_{21} = 90\%$, while the probability in the reverse direction is merely $p_{12} = 10\%$. We further assume that the initial state values at node 1 and node 2 are $x_1(0) = 10$ and $x_2(0) = -10$ respectively. Therefore the average value is $\bar{z} = 0$ and the desired consensus configuration is $(x_1, x_2) = (0, 0)$. To simplify the analysis, we assume the parameter k is set to 1, i.e., a corrective iteration takes place after each standard iteration, and consider only the first two iterations. In addition, we set the step size to $\epsilon = \frac{1}{2}$, and assume that $n = m = 1$, i.e., nodes do not repeatedly send the same packets.

Let us start by calculating for the case that $A_{12}(0) = 0$, $A_{21}(0) = 1$, $A_{12}(1) = 0$, $A_{21}(1) = 1$. Here $A_{ij}(t)$ is the (i,j)-th entry of the adjacency matrix, and $A_{ij}(t) = 1$ means that in the t-th iteration, the packet from node j is successfully delivered at node i; and zero otherwise.

In the first iteration, state value $x_1(0)$ is delivered to node 2, but $x_2(0)$ is lost. After receiving the state value, node 2 updates its state as $x_2(1) = x_2(0) + \frac{1}{2}(x_1(0) - x_2(0)) = -10 + (10 - (-10))/2 = 0$. On the other hand, node 1 does not receive any state value and thus does not change its state value and therefore $x_1(1) = x_1(0) = 10$. At this point, the auxiliary variables are $\phi_{12}(1) = 0$ and $\phi_{21}(1) = 10$.

In the second iteration, both nodes would try to correct their state values by running a corrective iteration. Here we suppose $A_{12}(1) = 0$, $A_{21}(1) = 1$, or in other words, $\phi_{12}(1)$ is delivered to node 2, but $\phi_{21}(1)$ does not arrive at node 1. Therefore, when equally dividing Δ_{12}, node 2 corrects its state as $x_2(2) = x_2(1) - \frac{1}{2}(\phi_{12}(1) + \phi_{21}(1)) = 0 - (0+10)/2 = -5$, but node 1 does not receive $\phi_{21}(1)$ and therefore cannot correct its state, so $x_1(2) = x_1(1) = 10$.

On the other hand, when dividing Δ_{12} in proportion to packet reception ratios, we have $x_2(2) = x_2(1) - \frac{p_{12}}{p_{12}+p_{21}}(\phi_{12}(1) + \phi_{21}(1)) = 0 - \frac{10\%}{90\%+10\%}(0+10) = -1$, and $x_1(2) = x_1(1) = 10$. This is closer to the final consensus configuration $(x_1 = x_2 = 0)$ than dividing Δ_{12} equally.

To quantify the distance to the consensus configuration, we define $e(t) := x_1^2(t) + x_2^2(t)$. One can see that $e(2) = 125$ and $e(2) = 101$ respectively for the two dividing methods, and that dividing in proportion to p_{ij} yields better performance.

This analysis can be repeated for all possible instances of packed drops. Table I lists the results for all non-trivial cases, e.g., when A_{12} and A_{21} are both nonzero. The last column lists the probability of each corresponding case. For example, the probability for the first case can be computed as $P(A_{12}(0) = 0) \times P(A_{21}(0) = 1) \times P(A_{12}(1) = 0) \times$

Dividing Δ_{12}	$A_{12}(0)$	$A_{21}(0)$	$A_{12}(1)$	$A_{21}(1)$	$x_1(1)$	$x_2(1)$	$\phi_{12}(1)$	$\phi_{21}(1)$	$x_1(2)$	$x_2(2)$	$e(2)$	Probability
Half & half	0	1	0	1	10	0	0	10	10	-5	125	0.6561
Proportional	0	1	0	1	10	0	0	10	10	-1	**101**	0.6561
Half & half	0	1	1	1	10	0	0	10	5	-5	50	0.0729
Proportional	0	1	1	1	10	0	0	10	1	-1	**2**	0.0729
Half & half	0	1	1	0	10	0	0	10	5	0	25	0.0081
Proportional	0	1	1	0	10	0	0	10	1	0	**1**	0.0081
Half & half	1	0	0	1	0	-10	-10	0	0	-5	**25**	0.0081
Proportional	1	0	0	1	0	-10	-10	0	0	-9	81	0.0081
Half & half	1	0	1	1	0	-10	-10	0	5	-5	**50**	0.0009
Proportional	1	0	1	1	0	-10	-10	0	9	-9	162	0.0009
Half & half	1	0	1	0	0	-10	-10	0	5	-10	**125**	0.0001
Proportional	1	0	1	0	0	-10	-10	0	9	-10	181	0.0001

TABLE I

CONSENSUS ITERATIONS IN THE TWO-NODE NETWORK. PROPORTIONAL DIVIDING PERFORMS BETTER IN THE MORE PROBABLE CASES.

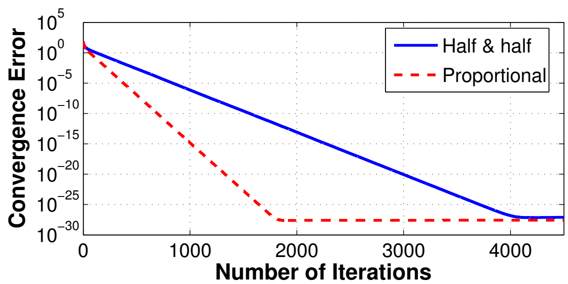

Fig. 2. Convergence speed for different ways of dividing Δ_{ij} in a 10-node line topology.

$P(A_{21}(1) = 1) = 0.9 \times 0.9 \times 0.9 \times 0.9 = 0.6561$.

One can see from Table I that proportionally dividing Δ_{ij} yields faster convergence for the more probable cases and therefore converges faster in expectation. The expected error $\mathbf{E}\big(e(2)\big)$ for the cases in Table I is 86.12 when dividing by half and 67.24 when dividing in proportion to p_{ij}.

We do not claim that the choice of weights for splitting the Δ_{ij} in (7) is optimal. However, in practice, as it will be shown in the next section, this choice is superior to using equal weights. In our future work we will investigate criteria for choosing the weights optimally.

D. Simulation Results

We run the corrective consensus in a 10-node linear topology where $p_{ij} = 10\%$, $\forall i = j + 1$ and $p_{ij} = 90\%$, $\forall i = j - 1$. Figure 2 presents the convergence results for the two methods of dividing Δ_{ij}. The y-axis draws $e(t) := \sum_{i=1}^{N}(x_i(t) - \overline{z})^2$, and each curve was averaged over 100 independent experiments with the initial state values following a uniform distribution. During the experiments, we set $n = 1$, $m = 5$ and $k = 1$. One can see that the original corrective consensus, i.e, dividing Δ_{ij} into halves, indeed converges to the average value, as show in Section IV-A. In addition, it is also obvious that dividing Δ_{ij} according to the values of p_{ij} and p_{ji} gives significantly faster ($\approx$ 2.3 times) convergence speed than the original corrective consensus.

V. CONCLUSION

In this paper we have shown the convergence of corrective consensus with unequal packet reception ratios. We have also studied the effect of assigning the correction values according to packet reception ratios, which empirically converges faster than the original corrective consensus algorithm. Our future work will further investigate its convergence properties.

REFERENCES

[1] T. Aysal and K. Barner. Convergence of consensus models with stochastic disturbances. *IEEE Transactions on Information Theory*, 56(8):4101 –4113, 2010.

[2] T. Aysal, A. Sarwate, and A. Dimakis. Reaching consensus in wireless networks with probabilistic broadcast. In *Allerton Conference on Communication, Control, and Computing*, pages 732 –739, 2009.

[3] T. C. Aysal, B. N. Oreshkin, and M. J. Coates. Accelerated distributed average consensus via localized node state prediction. *Trans. Sig. Proc.*, 57(4):1563–1576, 2009.

[4] S. Barbarossa and G. Scutari. Decentralized maximum-likelihood estimation for sensor networks composed of nonlinearly coupled dynamical systems. *IEEE Transactions on Signal Processing*, 55(7):3456–3470, 2007.

[5] Y. Chen, R. Tron, A. Terzis, and R. Vidal. Corrective consensus: Converging to the exact average. In *IEEE Conference on Decision and Control*, pages 1221–1228, 2010.

[6] Y. Chen, R. Tron, A. Terzis, and R. Vidal. On corrective consensus: Converging to the exact average. Technical report, Computer Science Department, Johns Hopkins University, 2010.

[7] Y. Chen, R. Tron, A. Terzis, and R. Vidal. Accelerated corrective consensus: Converge to the exact average at a faster rate. In *IEEE American Control Conference*, 2011.

[8] F. Fagnani and S. Zampieri. Average consensus with packet drop communication. *SIAM Journal of Control Optimization*, 48(1):102–133, 2009.

[9] P. A. Forero, A. Cano, and G. B. Giannakis. Consensus-based distributed linear support vector machines. In *ACM/IEEE International Conference on Information Processing in Sensor Networks*, pages 35–46, 2010.

[10] R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge Univ. Press, 1987.

[11] D. Kingston and R. Beard. Discrete-time average-consensus under switching network topologies. In *IEEE American Control Conference*, 2006.

[12] E. Kokiopoulou and P. Frossard. Polynomial filtering for fast convergence in distributed consensus. *IEEE Transactions on Signal Processing*, 57(1):342–354, 2009.

[13] T. Li and J.-F. Zhang. Consensus conditions of multi-agent systems with time-varying topologies and stochastic communication noises. *IEEE Transactions on Automatic Control*, 55(9), 2010.

[14] C.-J. M. Liang, N. B. Priyantha, J. Liu, and A. Terzis. Surviving Wi-Fi interference in low power ZigBee networks. In *ACM Conference on Embedded Networked Sensor Systems*, pages 309–322. ACM, 2010.

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

[16] R. Olfati-saber, E. Franco, E. Frazzoli, and J. S. Shamma. Belief consensus and distributed hypothesis testing in sensor networks. In *Network Embedded Sensing and Control Workshop*, volume 331, pages 169–182, 2006.

[17] R. Olfati-Saber and R. Murray. Consensus problems in networks of agents with switching topology and time-delays. *IEEE Transactions on Automatic Control*, 49(9):1520–1533, 2004.
[18] W. Ren, R. W. Beard, and E. M. Atkins. Information consensus in multivehicle cooperative control. *IEEE Control Systems Magazine*, 27(2):71–82, 2007.
[19] I. Schizas, G. Giannakis, S. Roumeliotis, and A. Ribeiro. Consensus in Ad-Hoc WSNs with noisy links—Part II: Distributed estimation and smoothing of random signals. *IEEE Transactions on Signal Processing*, 56(4):1650–1666, April 2008.
[20] A. Tahbaz-Salehi and A. Jadbabaie. On consensus over random networks. In *Allerton Conference on Communication, Control and Computing*, pages 1315–1321, 2006.
[21] J. Zhao and R. Govindan. Understanding Packet Delivery Performance In Dense Wireless Sensor Networks. In *International Conference on Embedded Networked Sensor Systems*, pages 1–13, 2003.

APPENDIX

CONVERGENCE OF CORRECTIVE CONSENSUS

We want to analyze the convergence properties of corrective consensus under unequal packet reception ratios. Similar to [6], we prove Theorem 1 from Section IV-A in two steps: first assume that the ϕ_{ij}'s are always successfully delivered, and then remove this assumption.

A. Non-lossy ϕ_{ij}

In the case of non-lossy ϕ_{ij}, we need to show that Theorem 2, 3, 4 and 5 in [6] hold. Specifically, unequal packet reception ratios mean that $\mathbf{E}(W_{ij}(t)) \neq \mathbf{E}(W_{ji}(t))$, which mainly affects the proofs for Theorem 3 and leave the other three theorems intact. Therefore, in what follows we will adopt the very steps in the proof of Theorem 3 in [6] and report only the necessary modifications.

First, we observe that Equations (24)-(29) in [6] remain unaffected. Therefore, we have

$$\mathbf{E}\|\tilde{x}(k+1)\| = \mathbf{E}\|\tilde{x}(k)\| + \frac{1}{2}\sum_{s=0}^{k-1}\mathbf{E}\sqrt{\sum_{j=1}^{N}\Big(\sum_{i=1}^{N}\delta_{ij}(s)\Big)^2}, \quad (9)$$

where $\delta_{ij}(t) = (W_{ij}(t) - Wji(t))(\tilde{x}_j(t) - \tilde{x}_i(t))$. Each term of the summation in Equation (9) can be bounded by using Jensen's inequality. Therefore, we can rewrite Equation (30) in [6] while taking into account that $\mathbf{E}(\delta_{ij}(t)) \neq 0$ as

$$\begin{aligned}
&\mathbf{E}\left[\Big(\sum_{i=1}^{N}\delta_{ij}(s)\Big)^2\Big|\tilde{x}(s)\right] = \mathbf{E}\left[\sum_{i=1}^{N}\sum_{l=1}^{N}\delta_{ij}(s)\delta_{lj}(s)\Big|\tilde{x}(s)\right] \\
&\le \mathbf{E}\left[\frac{1}{2}\sum_{i=1}^{N}\sum_{l=1}^{N}\left(\delta_{ij}^2(s) + \delta_{lj}^2(s)\right)\Big|\tilde{x}(s)\right] \\
&= \mathbf{E}\left[N\sum_{i=1}^{N}\delta_{ij}^2(s)\Big|\tilde{x}(s)\right] \\
&= N\sum_{i=1}^{N}\mathbf{E}\left[(W_{ij}(s) - W_{ji}(s))^2\right](\tilde{x}_j(s) - \tilde{x}_i(s))^2 \\
&= N\sum_{i=1}^{N}(\hat{p}_{ij} + \hat{p}_{ji} - 2\hat{p}_{ij}\hat{p}_{ji})\epsilon^2(\tilde{x}_j(s) - \tilde{x}_i(s))^2 \\
&\le N(\hat{p}_{\hat{\imath}\hat{\jmath}} + \hat{p}_{\hat{\jmath}\hat{\imath}} - 2\hat{p}_{\hat{\imath}\hat{\jmath}}\hat{p}_{\hat{\jmath}\hat{\imath}})\epsilon^2\sum_{i=1}^{N}(\tilde{x}_j(s) - \tilde{x}_i(s))^2 \\
&= N\tilde{p}\epsilon^2\textstyle\sum_{i=1}^{N}(\tilde{x}_j(s) - \tilde{x}_i(s))^2 \\
&= N\tilde{p}\epsilon^2\Big(\sum_{i=1}^{N}\tilde{x}_j^2(s) + \sum_{i=1}^{N}\tilde{x}_i^2(s) - 2\sum_{i=1}^{N}\tilde{x}_j(s)\tilde{x}_i(s)\Big) \quad (10) \\
&= N\tilde{p}\epsilon^2\Big(N\tilde{x}_j^2(s) + \|\tilde{x}(s)\|^2\Big)
\end{aligned}$$

Recall that the effective PRR of the link $(i, j) \in E$ is $\hat{p}_{ij} := 1 - (1 - p_{ij})^n$, where n is the number of repeated transmissions in each standard iteration. In Equation (10) we define $(\hat{\imath}, \hat{\jmath}) = \arg\max_{1\le i,j\le N} \hat{p}_{ij} + \hat{p}_{ji} - 2\hat{p}_{ij}\hat{p}_{ji}$, and we use a shorthand $\tilde{p} = \hat{p}_{\hat{\imath}\hat{\jmath}} + \hat{p}_{\hat{\jmath}\hat{\imath}} - 2\hat{p}_{\hat{\imath}\hat{\jmath}}\hat{p}_{\hat{\jmath}\hat{\imath}}$ for convenience. Comparing the above result to Equation (30) in [6], one can see that the only difference is the coefficient N and the definition of $\tilde{p}$. The same applies to Equation (31)-(35) in [6]. Hence, we have

$$\mathbf{E}\|\tilde{x}((k+1)u)\| \le c^u\|\tilde{x}(0)\|, \quad (11)$$

where

$$c = \left(\overline{\lambda_2}^k + \frac{1}{2}\epsilon\sqrt{2\tilde{p}N^2}\frac{1 - \overline{\lambda_2}^k}{1 - \overline{\lambda_2}}\right), \quad (12)$$

and $\overline{\lambda_2} = \mathbf{E}(|\lambda_2(W(t))|)$, i.e., the expected second largest eigenvalue of the weight matrix. Note that we can enforce $c < 1$ by tuning the stepsize ϵ and the number of repeated packet transmissions n. Specifically, $\tilde{p}$ can be made arbirarily close to 0 by increasing n, as explained by Theorem 5 in [6].

Note that Equation (11) is different from the one in Theorem 1 in Section IV-A, and this is because in this section we are using the assumption that ϕ_{ij} is non-lossy, but Theorem 1 is for the more general case that ϕ_{ij} can be lossy, which will be addressed in next section.

The result in (11) is equivalent to Theorem 3 in [6] with the new critical value c defined in (12). By far, we have shown the convergence properties of corrective consensus under nonequal probabilities in the case of non-lossy ϕ_{ij}.

B. Lossy ϕ_{ij}

Next we will remove the non-lossy ϕ_{ij} assumption. In this case, we need to show that Theorem 6, 7 and 8 in [6] hold. Similar to the previous case, only Theorem 6 is affected and the other two theorems are intact with a trivial change to Theorem 8. In what follows, we will make necessary modifications to the proof of Theorem 6, and then point out the trivial change to Theorem 8.

As in [6], we use the random stationary variables $v_{ij}(u)$ to indicate the reception status of ϕ_{ij}, defined as $v_{ij}(u) = 1$ if node i receives ϕ_{ji} from node j at the u-th corrective iteration, and zero otherwise. We will also denote $q_{ij} = P(v_{ij} = 1) = 1 - (1 - p_{ij})^m$, with m being the number of repeated transmissions in each corrective iteration. In other words, q_{ij} is the effective PRR of the link $i \leftarrow j$ during corrective iterations.

In Theorem 6 [6], it can be verified that Equations (38)-(44) still hold without modification. However, $\mathbf{E}(\delta_{ij}) \neq 0$, therefore it is not straightforward to obtain Equation (46) in [6]. We will proceed by first analyzing the following term

$$\begin{aligned}
&\mathbf{E}\left[\Big(\textstyle\sum_{i=1}^{N}\delta_{ij}(s)v_{ij}(u)\Big)^2\Big|\tilde{x}(s)\right] \\
&= \mathbf{E}\left[\textstyle\sum_{i=1}^{N}\sum_{l=1}^{N}\delta_{ij}(s)\delta_{lj}(s)v_{ij}(u)v_{lj}(u)\Big|\tilde{x}(s)\right]
\end{aligned}$$

$$\leq \mathbf{E}\Big[\frac{1}{2}\sum_{i=1}^{N}\sum_{l=1}^{N}\big(\delta_{ij}^2(s)+\delta_{lj}^2(s)\big)v_{ij}(u)v_{lj}(u)\Big|\tilde{x}(s)\Big] \tag{13}$$
$$\leq \mathbf{E}\Big[\frac{1}{2}\sum_{i=1}^{N}\sum_{l=1}^{N}\big(\delta_{ij}^2(s)+\delta_{lj}^2(s)\big)\Big|\tilde{x}(s)\Big]$$

in which we use the fact that, by definition, $0 \leq v_{ij}(u) \leq 1$.

Combining (10) and (13), we obtain

$$\mathbf{E}\|\tilde{x}(\hat{u}-1)\| + \mathbf{E}\left[\frac{1}{2}\sum_{s=\hat{u}-k-1}^{\hat{u}-1}\Big\{\sum_{j=1}^{N}\Big[\sum_{i=1}^{N}\delta_{ij}(s)v_{ij}(u)\Big]^2\Big\}^{\frac{1}{2}}\right]$$
$$\leq c\mathbf{E}\|\tilde{x}(\hat{u}-k-1)\|,$$

in which the definition of c is given in (12), and we use $\hat{u} = (k+1)u$ as a shorthand. This result is identical to Equation (46) in [6] except for the different definition of c.

Next, Equation (47) and the first three rows of (48) in [6] hold and we get

$$\mathbf{E}\|\tilde{x}(\hat{u})\| \leq c\mathbf{E}\|\tilde{x}(\hat{u}-k-1)\| + \frac{1}{2}\sum_{r=1}^{u-1}\sum_{s=(k+1)(r-1)}^{(k+1)r-2} V_{u,s,r}(\delta_{ij}, v_{ij}). \tag{14}$$

where

$$V_{u,s,r}(\delta_{ij}, v_{ij}) = \mathbf{E}\sqrt{\sum_{j=1}^{N}\Big[\sum_{i=1}^{N}\delta_{ij}(s)v_{ij}(u)\Lambda_{ij}(u-1,r)\Big]^2}$$
$$\leq \mathbf{E}\left[\sqrt{\mathbf{E}\Big[\sum_{j=1}^{N}\Big[\sum_{i=1}^{N}\delta_{ij}(s)v_{ij}(u)\Lambda_{ij}(u-1,r)\Big]^2\Big|\tilde{x}(s)\Big]}\right], \tag{15}$$

in which we define

$$\Lambda_{ij}(u,r) = \prod_{l=r}^{u}\Big(1-\frac{1}{2}\big(v_{ij}(l)+v_{ji}(l)\big)\Big).$$

The innermost term on the second row of (15) can be expanded as

$$\Big[\sum_{i=1}^{N}\delta_{ij}(s)v_{ij}(u)\Lambda_{ij}(u-1,r)\Big]^2$$
$$=\sum_{i=1}^{N}\sum_{l=1}^{N}\delta_{ij}(s)\delta_{lj}(s)v_{ij}(u)v_{lj}(u)\Lambda_{ij}(u-1,r)\Lambda_{lj}(u-1,r)$$
$$\leq\frac{1}{2}\sum_{i=1}^{N}\sum_{l=1}^{N}(\delta_{ij}^2(s)+\delta_{lj}^2(s)) \times v_{ij}(u)v_{lj}(u)\Lambda_{ij}(u-1,r)\Lambda_{lj}(u-1,r)$$
$$=\sum_{i=1}^{N}\delta_{ij}^2(s)v_{ij}(u)\Lambda_{ij}(u-1,r)\sum_{l=1}^{N}v_{lj}(u)\Lambda_{lj}(u-1,r)$$
$$\leq \sum_{i=1}^{N}\delta_{ij}^2(s)v_{ij}(u)\Lambda_{ij}(u-1,r)N$$

where we use the property that $0 \leq \Lambda_{ij}(u-1,r) \leq 1$ and $0 \leq v_{ij}(u) \leq 1$, $\forall i,j,u,r$. Then, we can write the second innermost term on the second row of (15) as

$$\mathbf{E}\Big[\sum_{j=1}^{N}\Big[\sum_{i=1}^{N}\delta_{ij}(s)v_{ij}(u)\Lambda_{ij}(u-1,r)\Big]^2\Big|\tilde{x}(s)\Big]$$
$$\leq \mathbf{E}\Big[N\sum_{j=1}^{N}\sum_{i=1}^{N}\delta_{ij}^2(s)v_{ij}(u)\Lambda_{ij}(u-1,r)\Big|\tilde{x}(s)\Big]$$
$$= N\sum_{j=1}^{N}\sum_{i=1}^{N}\mathbf{E}(\delta_{ij}^2(s)|\tilde{x}(s))\mathbf{E}v_{ij}(u)\mathbf{E}\Lambda_{ij}(u-1,r)$$
$$= N\sum_{j=1}^{N}\sum_{i=1}^{N}\mathbf{E}(\delta_{ij}^2(s)|\tilde{x}(s))q_{ij}(1-\frac{q_{ij}}{2}-\frac{q_{ji}}{2})^{u-r}$$
$$\leq Nq_{\hat{\imath}\hat{\jmath}}(1-\frac{q_{\hat{\imath}\hat{\jmath}}}{2}-\frac{q_{\hat{\jmath}\hat{\imath}}}{2})^{u-r}\sum_{j=1}^{N}\sum_{i=1}^{N}\mathbf{E}(\delta_{ij}^2(s)|\tilde{x}(s))$$
$$\leq N(1-\frac{q_{\hat{\imath}\hat{\jmath}}}{2}-\frac{q_{\hat{\jmath}\hat{\imath}}}{2})^{u-r}\sum_{j=1}^{N}\sum_{i=1}^{N}\mathbf{E}(\delta_{ij}^2(s)|\tilde{x}(s))$$
$$\leq N(1-\frac{q_{i^*j^*}}{2}-\frac{q_{j^*i^*}}{2})^{u-r}\sum_{j=1}^{N}\sum_{i=1}^{N}\mathbf{E}(\delta_{ij}^2(s)|\tilde{x}(s))$$
$$= N\tilde{q}^{u-r}\sum_{j=1}^{N}\sum_{i=1}^{N}\mathbf{E}(\delta_{ij}^2(s)|\tilde{x}(s))$$

in which we define $(\hat{\imath},\hat{\jmath}) = \arg\max_{1\leq i,j\leq N} q_{ij}(1-\frac{q_{ij}}{2}-\frac{q_{ji}}{2})^{u-r}$ and $(i^*,j^*) = \arg\min_{1\leq i,j\leq N; q_{ij}>0} q_{ij}+q_{ji}$, and use the property $q_{\hat{\imath}\hat{\jmath}} \leq 1$. In addition, we define $\tilde{q} = 1-\frac{q_{i^*j^*}}{2}-\frac{q_{j^*i^*}}{2}$ as a shorthand.

Using the above results, we can revise Equation (48) in [6] as follows

$$\mathbf{E}\sqrt{\sum_{j=1}^{N}\Big[\sum_{i=1}^{N}\delta_{ij}(s)v_{ij}(u)\Lambda_{ij}(u-1,r)\Big]^2}$$
$$\leq \mathbf{E}\left[\sqrt{\mathbf{E}[\sum_{j=1}^{N}\Big[\sum_{i=1}^{N}\delta_{ij}(s)v_{ij}(u)\Lambda_{ij}(u-1,r)\Big]^2\Big|\tilde{x}(s)]}\right]$$
$$\leq \mathbf{E}\left[\sqrt{N\tilde{q}^{u-r}\sum_{j=1}^{N}\sum_{i=1}^{N}\mathbf{E}(\delta_{ij}^2(s)|\tilde{x}(s))}\right]$$
$$\leq \tilde{q}^{\frac{u-r}{2}}\sqrt{2\tilde{p}N^2}\epsilon\mathbf{E}\|\tilde{x}(s)\|$$

Finally, Equations (49), (50) and (51) in [6] remain the same except for the additional coefficient N and the different definitions of the critical value c and the shorthands $\tilde{q}$ and $\tilde{p}$. Hence, we have proved Theorem 6 in [6]. This implies the statement of Theorem 1 in Section IV-A with b satisfying $\tilde{q} \leq \frac{b(b-c)}{2b-c}$, and we will have

$$\mathbf{E}\left(\|\tilde{x}((k+1)u)\|\right) \leq cb^{u-1}\mathbf{E}\left(\|\tilde{x}(0)\|\right). \tag{16}$$

One can see that Equation (11) is a special case of Equation (16), in which b degenerates to $b = c$.

Now Theorem 8 in [6] should be modified to state that such value b always exists with an appropriate choice of the number of repeated transmissions m, which is trivial to show because $\tilde{q}$ can be made arbitrarily close to 0 by tuning the value of m.