Monotonic Mean-Squared Convergence Conditions for Random Pairwise Consensus Synchronization in Wireless Networks

D. Richard Brown, III, *Senior Member, IEEE*, Andrew G. Klein, *Senior Member, IEEE*, and Rui Wang, *Student Member, IEEE*

***Abstract*—Time synchronization is important for a variety of applications in wireless networks including scheduling communication resources, interference avoidance, and data fusion. This paper analyzes the problem of synchronizing nodes in a time-division-duplexed wireless network via consensus methods using only acknowledged message exchanges over existing network traffic. The nodes are assumed to communicate synchronization information randomly and asymmetrically, reflecting the random nature of the timing and the source of synchronization information gleaned from or embedded in existing network traffic. The synchronization method accounts for non-negligible propagation delays which can be disruptive to consensus techniques. To characterize both transient and asymptotic consensus behavior, general results are presented providing necessary and sufficient conditions for monotonic mean squared convergence of a distance from consensus metric at an exponential rate. While the general results apply to a broad class of random consensus models, two models are analyzed in detail: i) random asymmetric gossip and ii) fully connected random broadcast. Bounds are derived for the steady-state distance from consensus in the presence of estimation error. Numerical results are also presented verifying the analysis under different network topologies.**

***Index Terms*—Consensus clock, distributed synchronization, random asymmetric gossip, random broadcast, wireless networks.**

I. INTRODUCTION

SYNCHRONIZATION is the process of establishing a common notion of time among two or more entities. In the context of wired and wireless communication networks, synchronization enables coordination among the nodes in the network and can facilitate scheduling of communication resources, interference avoidance, event detection/ordering, data fusion, and coordinated wake/sleep cycles. Over the last 30 years, a variety of synchronization protocols have been developed for both wired and wireless networks including Network Time Protocol (NTP) [1], Precision Time Protocol (PTP) [2], the Global Positioning System (GPS) [3], [4], and several lightweight protocols for sensor networks, e.g. [5]–[13]. Several modern wireless communication standards, e.g. 802.11, 802.15, and 802.16, also use network synchronization for various functions including time-slotting and power management. In almost all of these examples, the synchronization process requires specific network structures and a dedicated synchronization protocol, both of which result in undesirable overhead.

More recently, researchers have begun to consider the idea of achieving network synchronization through *consensus* techniques, see e.g., [14]–[26]. Rather than synchronizing the nodes in the network to some external "reference" time as in, e.g., NTP and PTP, the idea is to allow the nodes in the network to communicate among themselves and arrive at a common clock rate and offset. Consensus synchronization is appealing in that it does not require specific network structures and has the potential to be embedded in existing peer-to-peer network traffic to reduce or avoid overhead.

The focus of this paper is on consensus synchronization techniques that can be embedded in existing network traffic. Since we assume no dedicated synchronization protocol, the information flow in the network from the perspective of the synchronization function is *random* and potentially *asymmetric*. Random consensus methods for synchronization were considered in [20], [22], [25] and in a more general consensus context in [27]–[33]. Of this prior work, [22], [25], [30], [31], [33] specifically consider random consensus with asymmetric information flows.

In the case of asymmetric information flows, since the average of the states is not preserved, the focus is often on analyzing a "distance from consensus" metric, which is a measure of the state displacement from the current average. Given a state $\boldsymbol{x}[k] = [x_1[k], \dots, x_N[k]]^\top \in \mathbb{R}^N$ at time k, the distance from consensus at time k is defined as

$$d[k] := \frac{1}{N} \left\| \boldsymbol{x}[k] - \mathbf{1}_N \bar{x}[k] \right\|_2^2 \qquad (1)$$

where $\bar{x}[k] := \frac{1}{N}\sum_{i=1}^{N} x_i[k]$ and $\mathbf{1}_N \in \mathbb{R}^N$ is a vector of ones. In the context of synchronization, the state $\boldsymbol{x}[k]$ corresponds to the drifts and/or offsets of the clocks in the network. It has been shown that, under certain conditions, the distance from consensus converges almost surely or in a mean squared sense at an exponential rate, e.g., $\lim_{k\to\infty}(d[k])^{1/k} = \lambda^2$ almost surely for a given constant λ [30], [31]. Similar convergence results for random asymmetric consensus systems are also provided in [22], [25], [33].

D. R. Brown, III, and R. Wang are with the Department of Electrical and Computer Engineering, Worcester Polytechnic Institute, Worcester MA 01609 USA (e-mail: drb@wpi.edu; rwang@wpi.edu).

A. G. Klein is with the Department of Engineering and Design, Western Washington University, Bellingham, WA 98225 USA (e-mail: andy.klein@wwu.edu).

To the best of our knowledge, all of the results for random consensus systems consider only the *asymptotic* behavior of the system as $k \to \infty$. The transient behavior of consensus systems is generally not studied. This behavior may be important in some applications, however. For example, in the context of synchronization, nodes may first establish a coarse level of synchronization to facilitate time-slotted communication and avoid channel contention. If the synchronization is then refined through consensus techniques and the consensus system has poor transient behavior, this can lead to increased clock dispersion, timeslot collisions, and network disruption.

Contributions:

1) To characterize both the transient and asymptotic consensus behavior, we derive necessary and sufficient conditions for *monotonic* mean squared convergence of the distance from consensus metric for a broad class of random asymmetric (or symmetric) consensus systems. Under these conditions, for any state $\boldsymbol{x}[k]$ with $d[k] > 0$ we have

$$\mathrm{E}\left\{d[k+1] \mid \boldsymbol{x}[k]\right\} < (1-\epsilon)d[k] \tag{2}$$

for some $\epsilon > 0$ where the conditional expectation is performed over the distribution corresponding to the random information flows in the network. Under certain assumptions on the random information flows, these conditions reduce to simple explicit stepsize bounds depending only on the number of nodes in the network.

2) We develop explicit stepsize bounds for monotonic mean squared convergence in two specific scenarios: (i) random asymmetric gossip [22], [30], [31] with a single message exchange in each timeslot and (ii) fully-connected random broadcast [25] with several message exchanges in each timeslot. These results are compared to the stepsize bounds previously developed for general asymptotic convergence in these scenarios.

3) We describe an explicit two-step synchronization approach which corrects both drifts and offsets and accounts for propagation delays. The effect of non-negligible propagation delays on consensus synchronization systems was considered in [18] and [25] where the approach is to allow the clock offsets to achieve consensus with a constant derivative, i.e., a common residual drift, proportional to the sum of the propagation delays. In this paper, we assume time-division-duplexed (TDD) communication with bidirectional message exchanges which allows the initiating node to disambiguate the clock offset from the propagation delay prior to performing local clock compensation. While our synchronization system also converges to a common residual drift due to the asymmetric information flows, the residual drift is not a function of the propagation delays.

4) We analyze the steady-state synchronization performance in the presence of drift and offset estimation errors. The effect of noise has been studied previously for symmetric consensus systems in [19], [27], [32]. In this paper, we leverage our monotonic mean squared convergence results to develop upper bounds on the steady-state distance from consensus metric for symmetric and asymmetric consensus synchronization systems with drift and offset estimation errors.

Related Prior Work: While the body of literature on clock synchronization and consensus is quite large, we summarize here the most relevant prior work with regards to consensus synchronization with random asymmetric information flows.

In [22], a proportional-integral (PI) controller is developed for a random asymmetric gossip consensus synchronization system in which only one pair of nodes communicate in each timeslot. The analysis assumes a fully-connected network with unidirectional messaging, negligible propagation delays, and equiprobable transmit/receive node pairs. Under these conditions, necessary and sufficient conditions on the PI controller adaptation parameter α for asymptotic convergence are developed. In this paper, we analyze a broader class of random messaging models and provide specific results for monotonic mean squared convergence in the random asymmetric gossip scenario considered in [22]. Our synchronization mechanism also accounts for non-negligible propagation delays.

In [25], a random broadcast consensus synchronization system with asymmetric information flows and propagation delays was studied. The synchronization mechanism in [25] uses unidirectional messaging and it was shown that, if the stepsize is chosen such that $0 < \mu < \frac{4}{\max|C_i|}$ where $\max|C_i|$ is the maximum number of neighbors among all nodes in the network, the system asymptotically converges to a common clock with a common residual drift. The residual drift is a consequence of the initial drifts and the non-negligible propagation delays in the network. The random messaging model in [25] assumes each node equiprobably transmits or receives timing information in each timeslot. In this paper, we analyze a broader class of random messaging models and provide specific results for monotonic mean squared convergence in a fully-connected random broadcast scenario with consensus dynamics identical to [25]. In this specific scenario, our necessary and sufficient condition on the stepsize for monotonic mean squared convergence reduces to $0 < \mu < \frac{4}{N}$ whereas [25] specifies $0 < \mu < \frac{4}{N-1}$ for asymptotic (not necessarily monotonic) convergence. We also specify a bidirectional messaging model whereas [25] uses unidirectional messages. While the use of bidirectional messages requires additional network resources, it allows for disambiguation of clock offsets and propagation delays prior to local clock corrections and allows for a fully-decentralized implementation of the synchronization protocol.

In [30], [31], [33], a broad class of random consensus systems is studied in a general (non-synchronization) context. Mean squared and almost sure asymptotic convergence results are developed as discussed previously with [33] incorporating a random messaging model with temporally correlated link activations. The analysis in these papers assumes that the linear consensus dynamics are represented by a stochastic matrix, i.e., the elements of the consensus update matrix are non-negative and the row sums are all equal to one. The convergence results in [33] further assume these matrices to have strictly positive diagonals. In this paper, we also develop general convergence results that apply to a broad class of random asymmetric and symmetric consensus systems but focus on necessary and sufficient conditions for *monotonic* mean squared convergence. Somewhat surprisingly, although monotonic mean squared convergence is stricter than general asymptotic mean squared convergence, our stepsize condition can be looser than the stepsize conditions developed in [30], [31]. For example, in an asymmetric gossip scenario with equiprobable links, our stepsize condition reduces to $0 < \mu < \frac{N}{N-1}$ whereas [30], [31] specifies a stepsize of $0 < \mu < 1$. As a consequence, we show that monotonic mean squared convergence can be achieved with non-stochastic consensus update matrices containing negative diagonal elements.

The consensus literature also includes several studies of *deterministic* consensus algorithms, e.g., [34]–[40]. Under certain conditions, it has been shown that such algorithms can exhibit uniform exponential convergence such that, for any starting time k_0 and any initial state $\boldsymbol{x}[k_0]$,

$$\|\boldsymbol{x}[k] - \mathbf{1}_N\bar{x}\|_2 \leq c\lambda^{k-k_0} \|\boldsymbol{x}[k_0] - \mathbf{1}_N\bar{x}\|_2$$

for all $k \geq k_0$, where c is a constant, $0 < \lambda < 1$, and $\bar{x}$ is the consensus value. Note that c can be larger than one, hence these results do not necessarily imply that the convergence to the consensus state is monotonic. These results are also based on the existence of a deterministic sequence of information flows and do not directly translate to a scenario with random information flows.

Outline: The rest of the paper is organized as follows. We present the general consensus model and monotonic mean squared convergence theorem in Section II. We then introduce the system model, local clock model, and drift and offset estimators in Section III. Random consensus-based drift and offset compensation is discussed in Section IV. The convergence behavior of this synchronization system with perfect estimates is analyzed in Section V. An analysis of the steady-state distance from consensus with non-zero estimation errors is presented in Section VI. Numerical results are given in Section VII, followed by conclusions in Section VIII. A proof of the main theorem is provided in the Appendix.

Notation: Vectors and matrices are denoted by boldface letters. The $N \times N$ identity matrix is denoted $\boldsymbol{I}_N$, and $\mathbf{1}_N$ denotes a length N vector of all ones. We use $\mathrm{E}\{\cdot\}$, $(\cdot)^\top$, and $\|\cdot\|_2$ for expectation, transposition, and Euclidean norm, respectively. Definitions are also denoted by $:=$. The symmetric and idempotent matrix $\boldsymbol{Q} \in \mathbb{R}^{N\times N}$ is defined as

$$\boldsymbol{Q} := \boldsymbol{I}_N - \frac{1}{N}\mathbf{1}_N\mathbf{1}_N^\top. \tag{3}$$

Finally, for a real symmetric matrix $\boldsymbol{X}$ we define

$$\lambda_{\max}(\boldsymbol{X}) := \max \operatorname{eig}(\boldsymbol{X}). \tag{4}$$

II. Monotonic Mean Squared Convergence

This section provides general monotonic mean square convergence results for a class of consensus systems that will be applied in the context of synchronization for the remainder of the paper. We consider a state vector $\boldsymbol{x}[k] \in \mathbb{R}^N$ with random dynamics governed by

$$\boldsymbol{x}[k+1] = (\boldsymbol{I}_N + \mu\boldsymbol{R}[k])\,\boldsymbol{x}[k] \tag{5}$$

where $\mu > 0$ is a stepsize parameter and $\boldsymbol{R}[k] \in \mathbb{R}^{N\times N}$ is a matrix randomly drawn from some finite set $\mathcal{R}$ satisfying the property $\boldsymbol{R}[k]\mathbf{1}_N = \mathbf{0}$ for all $R[k] \in \mathbb{R}$. In the context of synchronization, the state vector $\boldsymbol{x}[k]$ corresponds to clock drifts and/or offsets and the $\boldsymbol{R}[k]$ matrices correspond to clock corrections resulting from random interactions among the nodes in the network. Given the dynamics in (5), we are interested in characterizing the conditions under which the distance from consensus metric as defined in (1) exhibits monotonic mean squared convergence according to (2).

We assume $\boldsymbol{R}[k]$ is independent and identically distributed (i.i.d.) for all k and denote

$$\bar{\boldsymbol{R}} := \mathrm{E}\{\boldsymbol{R}[k]\} \tag{6}$$

$$\bar{\boldsymbol{S}} := \mathrm{E}\left\{\boldsymbol{R}^\top[k]\boldsymbol{Q}^\top\boldsymbol{Q}\boldsymbol{R}[k]\right\}. \tag{7}$$

We further define the matrix $\boldsymbol{U} \in \mathbb{R}^{N\times(N-1)}$ as a matrix composed of $N-1$ orthonormal columns all orthogonal to $\mathbf{1}_N$.

A. General Monotonic Mean Squared Convergence

The following theorem establishes necessary and sufficient conditions on the stepsize μ such that the system (5) exhibits monotonic mean squared convergence to consensus.

Theorem 1 (General Monotonic Mean Squared Conv.): Given the distance from consensus metric specified in (1) and random dynamics specified in (5), there exists $\epsilon > 0$ such that

$$\mathrm{E}\{d[k+1] \mid \boldsymbol{x}[k]\} < (1-\epsilon)d[k] \tag{8}$$

for all $\boldsymbol{x}[k] \in \mathbb{R}^N$ such that $d[k] > 0$ if and only if

$$\mu\lambda_{\max}\left(\boldsymbol{U}^\top(\bar{\boldsymbol{R}}^\top + \bar{\boldsymbol{R}} + \mu\bar{\boldsymbol{S}})\boldsymbol{U}\right) < 0. \tag{9}$$

A proof of this theorem is provided in Appendix A. Note that the condition $d[k] > 0$ implies that $\boldsymbol{x}[k] \neq \mathbf{1}_N\bar{x}[k]$, i.e., the system is not already in a state of consensus at time k. If the system were in consensus at time k, then it is straightforward to see that $\boldsymbol{x}[k+\ell] = \boldsymbol{x}[k]$ and $d[k+\ell] = 0$ for all $\ell = 0, 1, \ldots$.

Note that Theorem 1 provides necessary and sufficient conditions for *monotonic* mean squared convergence. Hence, these conditions can be considered sufficient, but not necessary, for asymptotic (non-monotonic) mean squared convergence.

Example 1: Consider the $N = 3$ dimensional state with

$$\boldsymbol{R}_1 = \begin{bmatrix} 0 & 0 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad \boldsymbol{R}_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & -1 \end{bmatrix}$$

where $\mathrm{Prob}[\boldsymbol{R}[k] = R_i] = 0.5$ for $i = 1, 2$. We can calculate

$$\bar{\boldsymbol{R}} = \begin{bmatrix} 0 & 0 & 0 \\ \frac{1}{2} & -\frac{1}{2} & 0 \\ \frac{1}{2} & 0 & -\frac{1}{2} \end{bmatrix} \text{ and } \bar{\boldsymbol{S}} = \begin{bmatrix} \frac{2}{3} & -\frac{1}{3} & -\frac{1}{3} \\ -\frac{1}{3} & \frac{1}{3} & 0 \\ -\frac{1}{3} & 0 & \frac{1}{3} \end{bmatrix}.$$

The $\boldsymbol{U} = [\boldsymbol{u}_1, \boldsymbol{u}_2]$ matrix can be further specified with columns

$$\boldsymbol{u}_1 = \frac{1}{\sqrt{6}}\begin{bmatrix} 2 \\ -1 \\ -1 \end{bmatrix} \text{ and } \boldsymbol{u}_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$$

which results in

$$\boldsymbol{U}^\top(\bar{\boldsymbol{R}}^\top + \bar{\boldsymbol{R}} + \mu\bar{\boldsymbol{S}})\boldsymbol{U} = \begin{bmatrix} -1+\mu & 0 \\ 0 & -1+\frac{\mu}{3} \end{bmatrix}.$$

For $\mu > 0$, we see that $\lambda_{\max} = -1 + \mu$. Hence $\mu\lambda_{\max} < 0$ for $0 < \mu < 1$ and we have monotonic mean squared convergence over this range of μ. When $\mu = 0$, $\mu\lambda_{\max} = 0$ and we do not have monotonic mean squared convergence. This is obvious, of course, since $\mu = 0$ implies constant dynamics in (5). For $\mu < 0$, $\lambda_{\max} = -1 + \frac{\mu}{3} < 0$, and $\mu\lambda_{\max} > 0$ for all $\mu < 0$. Hence $0 < \mu < 1$ is the range of stepsizes over which this system exhibits monotonic mean squared convergence.

B. Exponential Convergence Rate

Since Theorem 1 applies to *all* $\boldsymbol{x}[k] \in \mathbb{R}^N$ such that $d[k] > 0$, there exists $\epsilon > 0$ when (9) is satisfied such that

$$\mathrm{E}\left\{d[k] \mid \boldsymbol{x}[0]\right\} < (1-\epsilon)^k d[0].$$

Hence, when (9) is satisfied, the average distance from consensus metric exhibits an exponential convergence rate from any finite initial state and $\lim_{k\to\infty} \mathrm{E}\{d[k] \mid \boldsymbol{x}[0]\} = 0$. Since

$$\epsilon < -\mu\lambda_{\max}\left(\boldsymbol{U}^\top(\bar{\boldsymbol{R}}^\top + \bar{\boldsymbol{R}} + \mu\bar{\boldsymbol{S}})\boldsymbol{U}\right) \tag{10}$$

the convergence rate $(1-\epsilon)^k$ is conservative as it assumes that $\boldsymbol{x}[0], \boldsymbol{x}[1], \dots$ are all in the direction of the eigenvector corresponding to the worst-case eigenvalue. Hence, the convergence rate $(1-\epsilon)^k$ is an upper (worst-case) bound on the convergence rate of the average distance from consensus metric.

C. Special Monotonic Mean Squared Convergence

The following corollary provides a simple expression for the range of stepsizes that admit monotonic mean squared convergence for the special case when $\bar{\boldsymbol{R}}^\top + \bar{\boldsymbol{R}} = -\theta\bar{\boldsymbol{S}}$ for some constant $\theta > 0$ and when $\boldsymbol{U}^\top\bar{\boldsymbol{S}}\boldsymbol{U}$ is positive definite. Note that the latter condition was satisfied by the example in the previous section but the former condition was not satisfied. These conditions are satisfied under certain symmetry conditions on the node interactions, e.g., when the network is *balanced* and *connected* as discussed in Section V. We present this result here in a general context and specialize it to specific synchronization scenarios in Section V.

Collary 2 (Special Monotonic Mean Squared Conv.): Given $\bar{\boldsymbol{R}}^\top + \bar{\boldsymbol{R}} = -\theta\bar{\boldsymbol{S}}$ for some constant $\theta > 0$ and $\boldsymbol{U}^\top\bar{\boldsymbol{S}}\boldsymbol{U}$ positive definite,

$$\mu\lambda_{\max}\left(\boldsymbol{U}^\top(\bar{\boldsymbol{R}}^\top + \bar{\boldsymbol{R}} + \mu\bar{\boldsymbol{S}})\boldsymbol{U}\right) < 0$$

if and only if $0 < \mu < \theta$.

Proof: Given $\bar{\boldsymbol{R}}^\top + \bar{\boldsymbol{R}} = -\theta\bar{\boldsymbol{S}}$ for some constant $\theta > 0$, we can write

$$\boldsymbol{U}^\top(\bar{\boldsymbol{R}}^\top + \bar{\boldsymbol{R}} + \mu\bar{\boldsymbol{S}})\boldsymbol{U} = \boldsymbol{U}^\top\bar{\boldsymbol{S}}\boldsymbol{U}(\mu - \theta).$$

Since $\boldsymbol{U}^\top\bar{\boldsymbol{S}}\boldsymbol{U}$ is positive definite and $\theta > 0$, it follows that

$$\mu\lambda_{\max}\left(\boldsymbol{U}^\top(\bar{\boldsymbol{R}}^\top + \bar{\boldsymbol{R}} + \mu\bar{\boldsymbol{S}})\boldsymbol{U}\right) < 0$$

if and only if $0 < \mu < \theta$. ■

In light of Theorem 1, Corollary 2 implies that monotonic mean squared convergence is achieved if and only if $0 < \mu < \theta$ when the conditions on $\bar{\boldsymbol{R}}$ and $\bar{\boldsymbol{S}}$ are satisfied.

Example 2: Consider a $N = 2$ dimensional state with

$$\boldsymbol{R}_1 = \begin{bmatrix} -1 & 1 \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad \boldsymbol{R}_2 = \begin{bmatrix} 0 & 0 \\ 1 & -1 \end{bmatrix}$$

where $\mathrm{Prob}[\boldsymbol{R}[k] = R_i] = 0.5$ for $i = 1, 2$. We can calculate

$$\bar{\boldsymbol{R}} = \begin{bmatrix} -0.5 & 0.5 \\ 0.5 & -0.5 \end{bmatrix} \quad \text{and} \quad \bar{\boldsymbol{S}} = \begin{bmatrix} 0.5 & -0.5 \\ -0.5 & 0.5 \end{bmatrix}.$$

Given any $\boldsymbol{U}$ orthogonal to $\mathbf{1}_2$, it is straightforward to verify $\boldsymbol{U}^\top\bar{\boldsymbol{S}}\boldsymbol{U} > 0$. Also observe that $\theta = 2$. Corollary 2 then implies that this consensus system will exhibit monotonic mean squared convergence if and only if $0 < \mu < 2$. Note that for $1 < \mu < 2$, the consensus update matrix $\boldsymbol{W}[k] = \boldsymbol{I}_2 + \mu\boldsymbol{R}[k]$ is non-stochastic in this example.

While the results in Theorem 1 and Corollary 2 can be applied to a wide range of consensus problems, the rest of this paper focuses specifically on consensus synchronization with random asymmetric node interactions.

III. System Model

We assume a time-slotted wireless network with N nodes. We further assume a general communication model where, in a given timeslot, a randomly selected set of "initiating" nodes $\mathcal{I} \in \{1, \dots, N\}$ transmit messages to other nodes in the network. Each node $i \in \mathcal{I}$ receives replies from a random (or possibly deterministic) set of "responding" nodes $\mathcal{J}_i \in \{1, \dots, N\}$ and then estimates the relative drifts and/or offsets in these replies to adjust its local clock. Note that the synchronization information flows from the nodes in $\mathcal{J}_i$ to node i for each $i \in \mathcal{I}$. The messages and replies are assumed to be TDD.

A. Reference Time and Local Time

The nodes in the network do not possess a common notion of time. We will use the notation t to refer to some notion of reference time, i.e., the "true" time, in the system. All time-based quantities such as propagation delays and/or frequencies are specified in reference time unless otherwise noted.

None of the nodes have knowledge of the reference time t. The local time at node i is modeled as

$$t_i = t + \Delta_i(t)$$

where $\Delta_i(t)$ is a non-stationary random process that captures the effect of clock drift, fixed local time offset, local oscillator phase noise, and frequency instability [41].

Over a short interval T, a reasonable model of the local clock offset can be written as [42], [43]

$$\begin{bmatrix} \Delta_i[k+1] \\ \beta_i[k+1] \end{bmatrix} = \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \Delta_i[k] \\ \beta_i[k] \end{bmatrix} \tag{11}$$

where $\beta_i[k] := \frac{d}{dt}\Delta_i(t)|_{t=kT}$ is a dimensionless quantity representing the nominal relative rate of the clock at node i with respect to the reference time at time kT, $\Delta_i[k]$ is the local clock offset at time kT, and where we have assumed the stochastic behavior of the oscillator to be negligible. We can also write $\beta_i = 1 + \alpha_i$ where $|\alpha_i|$ represents the rate deviation from nominal and is bounded by the long-term stability specifications of an oscillator. Even for low cost oscillators, $|\alpha_i| \ll 1$ with typical values being $|\alpha_i| \in [10^{-8}, 10^{-5}]$ as discussed in [44].

B. Pairwise Drift Estimation

We define the pairwise drift between nodes i and j as observed at node i as

$$\beta_{j,i}[k] := \beta_j[k] - \beta_i[k].$$

Since nodes derive their symbol rate and carrier frequency from the same local oscillator that drives the local clock, any message between a pair of nodes in the network allows for the estimation of pairwise clock drift at the physical layer through carrier frequency and/or symbol rate offset estimation.

As an example, consider a wireless link with nominal carrier frequency of ω_c. Suppose node i initiates a message exchange and node j responds to node i by transmitting an unmodulated carrier $\cos(\omega_c(\beta_j[k]t + \Delta_j[k]))$ to node i. The down-mixed signal at node i can be written in the reference timebase as

$$\begin{aligned} r(t) &= 2 \cdot \mathsf{LPF}\left\{\cos\left(\omega_c\left(\beta_j[k]t + \Delta_j[k]\right) + \phi_{j,i}\right)\right. \\ &\qquad \left.\times \cos\left(\omega_c\left(\beta_i[k]t + \Delta_i[k]\right)\right)\right\} \\ &= \cos\left(\omega_c\beta_{j,i}[k]t + \omega_c\left(\Delta_j[k] - \Delta_i[k]\right) + \phi_{j,i}\right) \end{aligned}$$

where $\phi_{j,i}$ is the phase of the propagation channel from node j to node i and LPF denotes standard low-pass filtering. In node i's timebase, we have

$$\begin{aligned} r(t_i) &= \cos\left(\frac{\omega_c\beta_{j,i}[k]t_i}{\beta_i[k]} + \text{constant phase terms}\right) \\ &= \cos\left(\frac{\omega_c\beta_{j,i}[k]t_i}{1 + \alpha_i[k]} + \text{constant phase terms}\right) \\ &\approx \cos\left(\omega_c\beta_{j,i}[k]t_i + \text{constant phase terms}\right) \end{aligned}$$

where the approximation results from assumption that $|\alpha_i[k]| \ll 1$. From the observation $r(t_i)$, node i can use standard frequency estimation techniques, e.g. [45], to form an estimate of the pairwise drift $\beta_{j,i}[k]$.

Pairwise clock drifts can also be estimated by observing multiple timestamped messages from another node in the network [46]. Without restricting ourselves to a particular method for pairwise drift estimation, we denote the pairwise drift estimate at node $i \in \mathcal{I}$ based on the response from node $j \in \mathcal{J}_i$ in timeslot k as

$$\hat{\beta}_{j,i}[k] = \beta_{j,i}[k] + \zeta_{j,i}[k]$$

where $\zeta_{j,i}[k]$ is the pairwise drift estimation error.

C. *Pairwise Offset Estimation*

The pairwise offset estimators described in this section use bidirectional message exchanges to disambiguate pairwise clock offsets from propagation delays. Without this disambiguation, propagation delay can lead to bias in the clock offset estimates and can prevent the consensus metric from converging to zero [18], [25]. We denote the propagation delay from node i to node j as $\psi_{i,j}$. Since all of the message exchanges in the system are assumed to be TDD, we assume reciprocal propagation delays $\psi_{i,j} = \psi_{j,i}$ in each link. Basic electromagnetic principles have long established that channel reciprocity holds at the antennas when the channel is accessed at the same frequency in both directions [47]. Channel reciprocity can also be quite accurate at intermediate-frequency (IF) and/or baseband if a reciprocal transceiver architecture is used [48] and can be further improved through transceiver calibration techniques to remove I/Q imbalance effects [49], [50].

The sender/receiver protocol [7], as shown in Fig. 1, is one example of how nodes can disambiguate pairwise clock offsets from propagation delays. Given a packet transmitted by node i in local time $t_i^{(a)}$, it arrives at node j in local time $t_j^{(b)} = t_i^{(a)} + \psi_{i,j} + \Delta_j - \Delta_i$. The response from node j contains the local timestamps $t_j^{(b)}$ and $t_j^{(c)}$ and arrives at node i at local time $t_i^{(d)} = t_j^{(c)} + \psi_{i,j} + \Delta_i - \Delta_j$. After receiving the response, node i can compute the pairwise clock offset to node j as

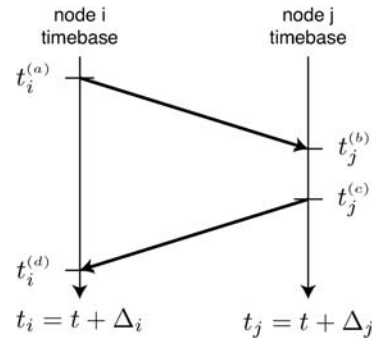


Fig. 1. Sender/receiver acknowledged message exchange.

$$\frac{\left(t_j^{(b)} - t_i^{(a)}\right) - \left(t_i^{(d)} - t_j^{(c)}\right)}{2} = \Delta_j - \Delta_i = \Delta_{j,i}$$

where it is assumed that any pairwise drift between nodes i and j is negligible.

The timestamp-free synchronization protocol [51] is another example of how pairwise clock offsets can be estimated and through acknowledged message exchanges, except through physical layer characteristics of the transmissions and without the use of timestamps. Again, without restricting ourselves to a particular method for pairwise offset estimation, we denote the pairwise offset estimate at node $i \in \mathcal{I}$ based on the response from node $j \in \mathcal{J}_i$ in timeslot k as

$$\hat{\Delta}_{j,i}[k] = \Delta_{j,i}[k] + \eta_{j,i}[k]$$

where $\eta_{j,i}[k]$ is the pairwise offset estimation error.

IV. Consensus Synchronization Dynamics

This section describes how pairwise drift and offset estimates can be used to correct local clocks and arrive at consensus for the clock drifts and offsets. The goal is not to force $\beta_i[k] = 1$ and $\Delta_i[k] = 0$ or to achieve "average consensus" where the global average of the drifts and offsets are preserved over time [28]. Rather, the goal is to drive the clock drifts and offsets to *common* values $\bar{\beta}$ and $\bar{\Delta}$ across the network, both of which may differ from the averages in earlier timeslots (this is sometimes called "alignment" rather than "consensus", e.g., [15], [18]). For conceptual simplicity, we describe the network synchronization protocol as a two-step process similar to [16]: (i) drift compensation and (ii) offset compensation. In practice, both drift and offset compensation can be performed simultaneously since pairwise drift estimates can often be inferred "for free" from physical layer characteristics of normal network traffic. Simultaneous drift and offset compensation is considered in Section VII.

A. *Step 1: Drift Compensation*

During the drift compensation step, each node $i \in \mathcal{I}$ that initiated a message exchange in timeslot k uses the pairwise drift estimates $\hat{\beta}_{j,i}[k]$ for $j \in \mathcal{J}_i$ to adjust its local clock drift. We define the drift vector at time k as

$$\boldsymbol{\beta}[k] := [\beta_1[k], \ldots, \beta_N[k]]^\top \in \mathbb{R}^N. \tag{12}$$

The pairwise drift estimates at node $i \in \mathcal{I}$ are then used to apply the local correction

$$\begin{aligned}\beta_i[k+1] &= \beta_i[k]\left(1 + \mu \sum_{j\in\mathcal{J}_i} \hat{\beta}_{j,i}[k]\right)\\ &\approx \beta_i[k] + \mu \sum_{j\in\mathcal{J}_i} \hat{\beta}_{j,i}[k]\\ &= \beta_i[k] + \mu\left(\sum_{j\in\mathcal{J}_i} (\boldsymbol{e}_j - \boldsymbol{e}_i)^\top \boldsymbol{\beta}[k] + \zeta_{j,i}[k]\right)\end{aligned}$$

where $\mu > 0$ is a stepsize parameter and the approximation discards the insignificant terms under the assumption that $|\alpha_i[k]| \ll 1$. The drift vector update in timeslot k is then

$$\begin{aligned}\boldsymbol{\beta}[k+1] &= \boldsymbol{\beta}[k] + \mu \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}_i} \boldsymbol{e}_i\left((\boldsymbol{e}_j - \boldsymbol{e}_i)^\top \boldsymbol{\beta}[k] + \zeta_{j,i}[k]\right)\\ &= \left(\boldsymbol{I}_N + \mu \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}_i} \boldsymbol{R}_{i,j}\right)\boldsymbol{\beta}[k] + \mu \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}_i} \boldsymbol{e}_i \zeta_{j,i}[k]\end{aligned}$$

where

$$\boldsymbol{R}_{i,j} := \boldsymbol{e}_i(\boldsymbol{e}_j - \boldsymbol{e}_i)^\top. \tag{13}$$

Since the sets $\mathcal{I}$ and $\mathcal{J}_i$ are randomly generated in each timeslot, the drift update vector can be written as

$$\boldsymbol{\beta}[k+1] = (\boldsymbol{I}_N + \mu \boldsymbol{R}[k])\,\boldsymbol{\beta}[k] + \mu \boldsymbol{\zeta}[k] \tag{14}$$

where $\boldsymbol{R}[k] := \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}_i} \boldsymbol{R}_{i,j}$ is a random matrix drawn from the finite set $\mathcal{R}$ and $\boldsymbol{\zeta}[k] := \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}_i} \boldsymbol{e}_i \zeta_{j,i}[k]$ is the random pairwise drift estimation error vector.

B. Step 2: Offset Compensation

During the offset compensation step, each node $i \in \mathcal{I}$ that initiated a message exchange in timeslot k uses the pairwise offset estimates $\hat{\Delta}_{j,i}[k]$ for $j \in \mathcal{J}_i$ to adjust its local clock offset. We assume that the drifts have previously achieved consensus such that $\boldsymbol{\beta}[k] = \bar{\beta}\mathbf{1}_N$. The offset vector at time k is defined as

$$\boldsymbol{\Delta}[k] := [\Delta_1[k], \ldots, \Delta_N[k]]^\top \in \mathbb{R}^N.$$

Note that, in light of (11), $\boldsymbol{\Delta}[k+1] = \boldsymbol{\Delta}[k] + T\bar{\beta}\mathbf{1}_N$ in the absence of offset corrections. The pairwise offset estimates at node $i \in \mathcal{I}$ are used to apply the local correction

$$\begin{aligned}\Delta_i[k+1] &= \Delta_i[k] + T\bar{\beta} + \mu \sum_{j\in\mathcal{J}_i} \hat{\Delta}_{j,i}[k]\\ &= \Delta_i[k] + T\bar{\beta} + \mu\left(\sum_{j\in\mathcal{J}_i} (\boldsymbol{e}_j - \boldsymbol{e}_i)^\top \boldsymbol{\Delta}[k] + \zeta_{j,i}[k]\right).\end{aligned}$$

Using $\boldsymbol{R}_{i,j}$ as defined in (13), the offset vector update in timeslot k follows as

$$\begin{aligned}\boldsymbol{\Delta}[k+1] &= \boldsymbol{\Delta}[k] + T\bar{\beta}\mathbf{1}_N + \mu\left(\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}_i} \boldsymbol{R}_{i,j}\boldsymbol{\Delta}[k] + \boldsymbol{e}_i \eta_{j,i}[k]\right)\\ &= (\boldsymbol{I}_N + \mu \boldsymbol{R}[k])\,\boldsymbol{\Delta}[k] + T\bar{\beta}\mathbf{1}_N + \mu \boldsymbol{\eta}[k]\end{aligned} \tag{15}$$

where $\boldsymbol{R}[k]$ is a random matrix as defined previously for drift compensation and $\boldsymbol{\eta}[k] := \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}_i} \boldsymbol{e}_i \eta_{j,i}[k]$ is the random pairwise offset estimation error vector. Note that the only difference between offset compensation and drift compensation in (14) is the common term $T\bar{\beta}\mathbf{1}_N$ present in (15). This term, however, has no effect on the distance from consensus metric in (1).

V. Random Asymmetric Messaging Models

This section applies the general convergence results from Section II to the drift and offset consensus synchronization systems developed Section IV under two different probabilistic assumptions on the initiating/responding node sets $\mathcal{I}$ and $\mathcal{J}_i$. The analysis in this section assumes that the estimation errors $\boldsymbol{\zeta}[k]$ and $\boldsymbol{\eta}[k]$ are zero.[1] An analysis of the optimal stepsize μ to maximize the convergence rate is also provided in Section V-C.

A. Random Asymmetric Gossip Consensus

In this section, we consider a scenario with a common wireless channel for all of the nodes in the network. In each timeslot, to avoid interference, only one pair of nodes exchange messages while all other nodes remain silent.

To model the random node interactions in this scenario, we denote the probability that node i initiates the message exchange with node j as $p_{i,j}$ and assume that these probabilities are time-invariant. Note that $p_{i,i} = 0$ and $\sum_i \sum_j p_{i,j} = 1$. We can further form a matrix $\boldsymbol{P} \in \mathbb{R}^{N\times N}$ with element (i,j) equal to $p_{i,j}$. In general we do not assume $p_{i,j} = p_{j,i}$; for example, $p_{i,j} > 0$ and $p_{j,i} = 0$ corresponds to the case where i initiates message exchanges with node j but node j never initiates message exchanges with node i.

The $\boldsymbol{P}$ matrix can be specified to encompass a wide range of network topologies including hierarchical and non-hierarchical structures. An example of a simple non-hierarchical topology is $p_{i,j} \equiv \frac{1}{N(N-1)}$ for all $i \neq j$. This corresponds to the case of a fully-connected network with equiprobable message exchanges between any pair of nodes in the network in each timeslot, similar to [22]. An example of a simple hierarchical topology for $N = 3$ nodes is

$$\boldsymbol{P} = \begin{bmatrix} 0 & 0 & 0 \\ 0.5 & 0 & 0 \\ 0.5 & 0 & 0 \end{bmatrix}. \tag{16}$$

In this example, node 1 never initiates message exchanges with the other nodes in the network. Nodes 2 and 3 initiate message exchanges with node 1 equiprobably, but never exchange messages with each other.

[1]The effect of non-zero estimation error is analyzed in Section VI and simulated in Section VII.

Under these assumptions, we have $\mathrm{Prob}[\boldsymbol{R}[k] = \boldsymbol{R}_{i,j}] = p_{i,j}$ with $\boldsymbol{R}_{i,j}$ as defined in (13). Hence, to apply Theorem 1, we can compute

$$\begin{aligned}\bar{\boldsymbol{R}} &:= \mathrm{E}\left\{\boldsymbol{R}[k]\right\} \\ &= \sum_{i,j} \boldsymbol{R}_{i,j} p_{i,j} \\ &= \sum_{i,j} \boldsymbol{e}_i (\boldsymbol{e}_j - \boldsymbol{e}_i)^\top p_{i,j} \\ &= \boldsymbol{P} - \mathrm{diag}(\boldsymbol{P}\mathbf{1}_N)\end{aligned} \tag{17}$$

where $\boldsymbol{P}\mathbf{1}_N$ is a column vector containing the row sums of $\boldsymbol{P}$. Given $\boldsymbol{Q}$ as defined in (3), we can also compute

$$\begin{aligned}\bar{\boldsymbol{S}} &:= \mathrm{E}\left\{\boldsymbol{R}^\top[k]\boldsymbol{Q}^\top\boldsymbol{Q}\boldsymbol{R}[k]\right\} \\ &= \sum_{i,j} (\boldsymbol{e}_j - \boldsymbol{e}_i)\boldsymbol{e}_i^\top \left(\boldsymbol{I}_N - \frac{1}{N}\mathbf{1}_N\mathbf{1}_N^\top\right) \boldsymbol{e}_i(\boldsymbol{e}_j - \boldsymbol{e}_i)^\top p_{i,j} \\ &= \sum_{i,j} \left(1 - \frac{1}{N}\right)(\boldsymbol{e}_j - \boldsymbol{e}_i)(\boldsymbol{e}_j - \boldsymbol{e}_i)^\top p_{i,j} \\ &= \left(1 - \frac{1}{N}\right)\left(\mathrm{diag}\left((\boldsymbol{P}+\boldsymbol{P}^\top)\mathbf{1}_N\right) - (\boldsymbol{P}+\boldsymbol{P}^\top)\right)\end{aligned} \tag{18}$$

where the second equality follows from the fact that $\boldsymbol{Q}$ is idempotent and where $(\boldsymbol{P} + \boldsymbol{P}^\top)\mathbf{1}_N$ is a column vector containing the sum of the row sums and column sums of $\boldsymbol{P}$. Note that $\boldsymbol{P}$ is not symmetric in general, but $\bar{\boldsymbol{S}}$ is symmetric.

These results show that, in the context of random asymmetric gossip synchronization, $\bar{\boldsymbol{R}}$ and $\bar{\boldsymbol{S}}$ are simple functions of $\boldsymbol{P}$. In fact, it is straightforward to verify that the $\bar{\boldsymbol{R}}$ and $\bar{\boldsymbol{S}}$ expressions in Example 1 in Section II-A correspond to the $\boldsymbol{P}$ matrix in (16).

Note that $\boldsymbol{P}$ induces a weighted directed graph $\mathcal{G}$ with vertices $V = \{1, \ldots, N\}$ and edges $E \subseteq V \times V$. Edge (j, i) is present, i.e., synchronization information flows from node i to node j, if $p_{i,j} > 0$. If the row sums of $\boldsymbol{P}$ are equal to the column sums of $\boldsymbol{P}$, the corresponding graph is *balanced*. A (weakly) connected graph that is also balanced is known to be strongly connected [17]. The following corollary provides an explicit condition on the stepsize μ when we have a connected network with the balanced property $\boldsymbol{P}\mathbf{1}_N = \boldsymbol{P}^\top\mathbf{1}_N$.

Corollary 3 (Balanced Connected Network Convergence): Given an asymmetric gossip scenario with a connected graph and $\boldsymbol{P}\mathbf{1}_N = \boldsymbol{P}^\top\mathbf{1}_N$, we have

$$\bar{\boldsymbol{R}}^\top + \bar{\boldsymbol{R}} = -\left(\frac{N}{N-1}\right)\bar{\boldsymbol{S}}$$

and $\boldsymbol{U}^\top\bar{\boldsymbol{S}}\boldsymbol{U}$ is positive definite.

Proof: Given a connected network with $\boldsymbol{P}\mathbf{1}_N = \boldsymbol{P}^\top\mathbf{1}_N$, we have $\mathrm{diag}(\boldsymbol{P}\mathbf{1}_N)^\top = \mathrm{diag}(\boldsymbol{P}^\top\mathbf{1}_N)$. From (17) we can write

$$\bar{\boldsymbol{R}}^\top + \bar{\boldsymbol{R}} = \boldsymbol{P} + \boldsymbol{P}^\top - \mathrm{diag}\left((\boldsymbol{P}+\boldsymbol{P}^\top)\mathbf{1}_N\right) = -\left(\frac{N}{N-1}\right)\bar{\boldsymbol{S}}$$

where the second equality uses (18).

We now establish that $\boldsymbol{U}^\top\bar{\boldsymbol{S}}\boldsymbol{U}$ is strictly positive definite when the graph is connected. For a weighted directed graph $\mathcal{G}$ with adjacency matrix $\boldsymbol{P}^\top$, the *degree matrix* of $\mathcal{G}$ is given by $\mathrm{diag}(\boldsymbol{P}^\top\mathbf{1}_N)$, and the *Laplacian matrix* is given by $\mathrm{diag}(\boldsymbol{P}^\top\mathbf{1}_N) - \boldsymbol{P}^\top$ [17]. Hence, $\bar{\boldsymbol{S}}$ is a symmetric Laplacian matrix for the weighted graph $\mathcal{G}'$ having adjacency matrix $\boldsymbol{P} + \boldsymbol{P}^\top$. Symmetric Laplacian matrices are positive semidefinite, with at least one zero eigenvalue corresponding to eigenvector $\mathbf{1}_N$. The zero eigenvalue has multiplicity m where m is the number of connected components in the graph [17]. Due to the assumption of connectedness there is $m = 1$ connected component, hence $\bar{\boldsymbol{S}}$ is positive semidefinite with a single zero eigenvalue corresponding to eigenvector $\mathbf{1}_N$. The columns of $\boldsymbol{U}$ are orthogonal to $\mathbf{1}_N$, so $\boldsymbol{U}^\top\bar{\boldsymbol{S}}\boldsymbol{U}$ is strictly positive definite since the single eigenvalue equal to zero in $\bar{\boldsymbol{S}}$ is annihilated by $\boldsymbol{U}$. ∎

Corollary 3 combined with Corollary 2 implies that, in the context of random asymmetric gossip, balanced connected networks exhibit monotonic mean squared convergence if and only if

$$0 < \mu < \frac{N}{N-1}. \tag{19}$$

This is a wider range of stepsizes than is usually considered for random asymmetric gossip [30], [31]. A consequence of this result is that μ can take on values larger than one (but less than $\frac{N}{N-1}$) and the system will still exhibit monotonic mean squared convergence. In fact, when $1 < \mu < \frac{N}{N-1}$, the matrix $\boldsymbol{W}[k] = \boldsymbol{I}_N + \mu\boldsymbol{R}[k]$ becomes non-stochastic since it has at least one diagonal element less than zero. This case is generally not considered in the consensus literature but, as shown in Corollary 3, still results in monotonic mean squared convergence of the distance from consensus metric.

Corollary 3 encompasses a wide range of common network structures including certain ring networks, networks with symmetric $\boldsymbol{P}$, and fully-connected networks with equiprobable edges with $p_{i,j} = \frac{1}{N(N-1)}$ for all $i \neq j$.

B. Fully-Connected Random Broadcast Consensus

In this section, we consider a random broadcast consensus scenario similar to [25]. While [25] specifies unidirectional messages, the resulting consensus dynamics described here are identical to those in [25] with the only difference being that our bidirectional message exchanges allow for the removal of propagation delays prior to local clock correction. In the random broadcast scenario with bidirectional message exchanges, each node in the network equiprobably and independently decides whether to initiate a message exchange or to respond to messages from other initiating nodes. Messages are orthogonalized in each timeslot to avoid interference. To facilitate analysis, we assume a fully-connected network such that every responding node responds to every initiating node, i.e.,

$$\mathcal{J}_i = \{1, \ldots, N\} \setminus \mathcal{I} = \mathcal{J}$$

for all $i \in \mathcal{I}$. All nodes either initiate or respond to messages. Note that, unlike the random asymmetric gossip scenario where one pairwise message occurs in each timeslot, the number of pairwise message exchanges for fully-connected random broadcast consensus is equal to $N_{\mathcal{I}}(N - N_{\mathcal{I}})$ where $N_{\mathcal{I}}$ is the number of initiating nodes in the timeslot and is binomially distributed with parameter $p = \frac{1}{2}$. The average number of pairwise message exchanges per timeslot is $\mathrm{E}\{N_{\mathcal{I}}(N - N_{\mathcal{I}})\} = \frac{N(N-1)}{4}$ for fully-connected random broadcast consensus. Hence, as demonstrated in Section VII, we can expect random broadcast consensus synchronization to converge more quickly than asymmetric gossip at the cost of significantly increased network resource requirements.

To analyze the convergence behavior of fully-connected random broadcast consensus, let $\boldsymbol{f}[k] \in \mathbb{R}^N$ be a vector with ith element equal to one if node i is an initiator in timeslot k and equal to zero otherwise. Further denote $N_{\mathcal{J}} = |\mathcal{J}| = N - N_{\mathcal{I}}$ as the number of responding nodes. It follows that

$$\begin{aligned}\boldsymbol{R}[k] &= \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}} \boldsymbol{e}_i(\boldsymbol{e}_j - \boldsymbol{e}_i)^\top \\ &= \boldsymbol{f}[k]\left(\mathbf{1}_N - \boldsymbol{f}[k]\right)^\top - N_{\mathcal{J}}\mathrm{diag}\left(\boldsymbol{f}[k]\right) \\ &= \boldsymbol{f}[k]\left(\mathbf{1}_N - \boldsymbol{f}[k]\right)^\top - (N - N_{\mathcal{I}})\mathrm{diag}\left(\boldsymbol{f}[k]\right).\end{aligned}$$

To compute $\bar{\boldsymbol{R}} = \mathrm{E}\{\boldsymbol{R}[k]\}$, it is straightforward to calculate

$$\mathrm{E}\left\{\boldsymbol{f}[k]\left(\mathbf{1}_N - \boldsymbol{f}[k]\right)^\top\right\} = \frac{1}{4}\left(\mathbf{1}_N\mathbf{1}_N^\top - \boldsymbol{I}_N\right).$$

The remaining term can be calculated by first calculating the conditional expectation

$$\mathrm{E}\left\{(N - N_{\mathcal{I}})\mathrm{diag}\left(\boldsymbol{f}[k]\right) \mid N_{\mathcal{I}}\right\} = (N - N_{\mathcal{I}})\frac{N_{\mathcal{I}}}{N}\boldsymbol{I}_N.$$

The unconditional expectation then follows as

$$\begin{aligned}\mathrm{E}\{(N - N_{\mathcal{I}})\mathrm{diag}\left(\boldsymbol{f}[k]\right)\} &= \left(\mathrm{E}\{N_{\mathcal{I}}\} - \frac{1}{N}\mathrm{E}\left\{N_{\mathcal{I}}^2\right\}\right)\boldsymbol{I}_N \\ &= \left(\frac{N}{2} - \frac{1}{N}\left(\left(\frac{N}{2}\right)^2 + \frac{N}{4}\right)\right)\boldsymbol{I}_N \\ &= \left(\frac{N-1}{4}\right)\boldsymbol{I}_N\end{aligned}$$

where the second equality uses the result that $N_{\mathcal{I}}$ is binomially distributed with parameter $p = \frac{1}{2}$. Putting these results together, we have

$$\begin{aligned}\bar{\boldsymbol{R}} &= \frac{1}{4}\left(\mathbf{1}_N\mathbf{1}_N^\top - \boldsymbol{I}_N\right) - \left(\frac{N-1}{4}\right)\boldsymbol{I}_N \\ &= \frac{1}{4}\left(\mathbf{1}_N\mathbf{1}_N^\top - N\boldsymbol{I}_N\right) \\ &= \frac{-N}{4}\boldsymbol{Q}.\end{aligned}$$

Using similar methods, it can be shown that

$$\bar{\boldsymbol{S}} = \frac{N}{8}\left(N\boldsymbol{I}_N - \mathbf{1}_N\mathbf{1}_N^\top\right) = \frac{N^2}{8}\boldsymbol{Q}.$$

It is straightforward to see that the conditions of Corollary 2 are satisfied in the fully-connected random broadcast consensus scenario with $\theta = \frac{4}{N}$. Hence, monotonic mean squared convergence occurs in this scenario if and only if

$$0 < \mu < \frac{4}{N}. \tag{20}$$

In [25], the maximum step size for asymptotic (not necessarily monotonic) convergence is specified as $\mu < \frac{4}{\max(|C_i|)}$ where $|C_i|$ is the number of neighbors of node i. In a fully-connected network, $|C_i| = N - 1$ and the asymptotic convergence condition in [25] reduces to $\mu < \frac{4}{N-1}$. Our analysis shows the slightly tighter condition in (20) is necessary and sufficient to guarantee monotonic mean squared convergence in the fully-connected random broadcast consensus scenario.

C. Optimal Step Size Selection

In the general case, choosing a stepsize μ to achieve monotonic mean squared convergence requires knowledge of $\bar{\boldsymbol{R}}$ and $\bar{\boldsymbol{S}}$ to calculate the maximum eigenvalue in (9). In the unbalanced random asymmetric gossip scenario, full knowledge of $\boldsymbol{P}$ is required to calculate $\bar{\boldsymbol{R}}$ and $\bar{\boldsymbol{S}}$. In practice, since we consider a scenario where the synchronization function is embedded in existing network traffic, nodes can estimate $\boldsymbol{P}$ (or the locally observable portions of $\boldsymbol{P}$) and disseminate these estimates to facilitate calculation of (9) prior to commencement of the consensus synchronization process. The use of network knowledge to select an appropriate stepsize parameter is common in consensus techniques. For example, selecting a stepsize to achieve asymptotic consensus in [15], [25] requires knowledge of the maximum number of neighbors among all nodes in the network.

Interestingly, for balanced connected random asymmetric gossip networks, Corollary 3 implies that the stepsize parameter μ can be chosen with no knowledge of $\boldsymbol{P}$. The range of μ leading to monotonic mean squared convergence is only a function of the network size N. This is also the case for fully-connected random broadcast consensus as shown in Section V-B. For balanced connected random asymmetric gossip, it is in fact sufficient to simply select $0 < \mu \leq 1$ to ensure monotonic mean squared convergence, i.e., the step size can be selected without even knowing the number of nodes in the network.

In networks that admit monotonic mean squared convergence, it is possible to derive from (28) in Appendix A the value of μ that provides the largest expected convergence step in timeslot k. Using the general state notation $\boldsymbol{x}[k]$ and denoting $\boldsymbol{v}[k] \in \mathbb{R}^{N-1}$ as the unique solution to $\boldsymbol{U}\boldsymbol{v}[k] = \boldsymbol{Q}\boldsymbol{x}[k]$, we can rewrite the numerator of (28) as

$$g(\mu) := \mu\boldsymbol{v}^\top[k]\boldsymbol{U}^\top(\bar{\boldsymbol{R}}^\top + \bar{\boldsymbol{R}})\boldsymbol{U}\boldsymbol{v}[k] + \mu^2\boldsymbol{v}^\top[k]\boldsymbol{U}^\top\bar{\boldsymbol{S}}\boldsymbol{U}\boldsymbol{v}[k].$$

The value of μ that minimizes $g(\mu)$ is

$$\mu_{\mathrm{opt}} = \frac{-\boldsymbol{v}^\top[k]\boldsymbol{U}^\top(\bar{\boldsymbol{R}}^\top + \bar{\boldsymbol{R}})\boldsymbol{U}\boldsymbol{v}[k]}{2\boldsymbol{v}^\top[k]\boldsymbol{U}^\top\bar{\boldsymbol{S}}\boldsymbol{U}\boldsymbol{v}[k]} \tag{21}$$

which reduces to $\mu_{\mathrm{opt}} = \frac{N}{2(N-1)}$ in the balanced connected random asymmetric gossip scenario since $\bar{\boldsymbol{R}}^\top + \bar{\boldsymbol{R}} = -\left(\frac{N}{N-1}\right)\bar{\boldsymbol{S}}$ as shown in Corollary 3. Similarly, for fully-connected broadcast consensus, the optimum stepsize can be calculated as $\mu_{\mathrm{opt}} = \frac{2}{N}$ since $\bar{\boldsymbol{R}}^\top + \bar{\boldsymbol{R}} = -\left(\frac{4}{N}\right)\bar{\boldsymbol{S}}$ as shown in Section V-B.

D. Discussion

It is worth mentioning that, even with perfect estimates, random consensus systems do not necessarily improve the *actual* distance from consensus metric in each timeslot even if the conditions for monotonic mean squared convergence are satisfied. For example, consider a fully-connected random asymmetric gossip scenario with $N = 3$ nodes and equiprobable edge activations such that $p_{i,j} = \frac{1}{6}$ for all $i \neq j$. Corollary 3 applies in this case and the range of μ necessary and sufficient for monotonic mean squared convergence is $0 < \mu < \frac{3}{2}$. Now suppose $\boldsymbol{\beta}[k] = [0, 2, 4]^\top$ with $d[k] = \frac{8}{3}$. If edge (3,2) is activated and $\mu = \frac{1}{2}$, the resulting drift vector at time $k+1$ is then $\boldsymbol{\beta}[k+1] = [0, 3, 4]^\top$ and $d[k+1] = \frac{26}{9} > d[k]$. In this example, even though the stepsize satisfies the requirements of Corollary 3 and node 2 has become more closely synchronized with node 3, the actual distance from consensus metric has become worse. We emphasize that the theorem and corollaries provide necessary and sufficient conditions for monotonic convergence of the *average* distance from consensus.

Although pairwise drift and offset compensation allow the nodes in a wireless network to achieve consensus on the drifts and offsets such that $\beta_i \to \bar{\beta}$ and $\Delta_i \to \bar{\Delta}$ for all i, consensus techniques can also be used to synchronize to an external source of reference time if one or more nodes in the network have access to reference time. The nodes that have access to an external source of reference time simply do not adapt their clocks. An example of this case is the asymmetric random gossip scenario with $\boldsymbol{P}$ matrix given in (16). This forces the other nodes in the network to adapt to the reference time. Hence, if the conditions of Theorem 1 are satisfied, the network will exhibit monotonic mean squared convergence toward synchronization with an external reference.

VI. Estimation Error

This section develops a bound for the distance from consensus metric for large k in the presence of non-zero drift and offset estimation errors. In general, with a fixed stepsize $\mu > 0$, non-zero drift and/or offset estimation errors prevent convergence of $d[k]$ to zero [19], [27], [32].

Since the distance from consensus dynamics are identical for drifts and offsets, our analysis in this section will focus on the drifts. We assume the drift estimation errors $\boldsymbol{\zeta}[k]$ are zero-mean and independent of $\boldsymbol{\beta}[k]$ and $d[k]$. If μ is chosen to satisfy the conditions of Theorem 1, we can write

$$\begin{aligned} \mathrm{E}\{d[k+1]|\boldsymbol{\beta}[k]\} &= \frac{1}{N}\mathrm{E}\left\{\|\boldsymbol{Q}\boldsymbol{\beta}[k+1]\|_2^2 \mid \boldsymbol{\beta}[k]\right\} \\ &= \frac{1}{N}\mathrm{E}\left\{\|\boldsymbol{Q}\boldsymbol{W}[k]\boldsymbol{\beta}[k] + \mu\boldsymbol{Q}\boldsymbol{\zeta}[k]\|_2^2 \mid \boldsymbol{\beta}[k]\right\} \\ &< (1-\epsilon)d[k] + \frac{\mu^2}{N}\mathrm{E}\left\{\boldsymbol{\zeta}^\top[k]\boldsymbol{Q}^\top\boldsymbol{Q}\boldsymbol{\zeta}[k]\right\} \end{aligned}$$

where $\boldsymbol{W}[k] := \boldsymbol{I}_N + \mu\boldsymbol{R}[k]$ and ϵ is the convergence rate parameter according to (10).

For large k, if the system has converged, we have $\mathrm{E}\{d[k+1]|\boldsymbol{\beta}[k]\} = d[k]$. This implies

$$\lim_{k\to\infty} d[k] < \frac{\mu^2}{N\epsilon}\mathrm{E}\left\{\boldsymbol{\zeta}^\top[k]\boldsymbol{Q}\boldsymbol{\zeta}[k]\right\} \tag{22}$$

where we have also used the fact that $\boldsymbol{Q}$ is idempotent.

This expression can be further simplified in the particular case of random asymmetric gossip consensus if we assume the drift estimation errors are i.i.d. with variance σ_β^2. In this case, since $\boldsymbol{\zeta}[k] = \boldsymbol{e}_i\zeta_{j,i}[k]$, we can write

$$\mathrm{E}\left\{\boldsymbol{\zeta}^\top[k]\boldsymbol{Q}\boldsymbol{\zeta}[k]\right\} = \sigma_\beta^2\mathrm{E}\left\{\boldsymbol{e}_i^\top\boldsymbol{Q}\boldsymbol{e}_i\right\} = \sigma_\beta^2\frac{N-1}{N}$$

and it follows from (22) that

$$\lim_{k\to\infty} d[k] < \frac{(N-1)\mu^2\sigma_\beta^2}{N^2\epsilon} \tag{23}$$

in the random asymmetric gossip scenario. If the conditions for Corollary 3 hold, we can set $\mu = c\frac{N}{N-1}$ with $0 < c < 1$ and write

$$\lim_{k\to\infty} d[k] < \frac{c^2\sigma_\beta^2}{(N-1)\epsilon}. \tag{24}$$

This result suggests that the steady-state distance from consensus metric is proportional to c^2 and inversely proportional to N. It is important to note, however, that the convergence rate parameter ϵ is an implicit function of μ and N such that $\epsilon \to 0$ as $\mu \to 0$ and/or $N \to \infty$. The precise effect of μ and N on the steady-state distance from consensus metric is left as a potential extension to this work. Also, it may be of interest to consider the use of a time-varying stepsize, e.g., [27], [32], to potentially drive the steady-state distance from consensus metric to zero.

For broadcast consensus, note that the ith element of $\boldsymbol{\zeta}[k]$ is zero when node $i \in \mathcal{J}$ is a responder and is equal to $\sum_{j\in\mathcal{J}}\zeta_{j,i}[k]$ when node $i \in \mathcal{I}$ is an initiator. Conditioning on $N_\mathcal{I} = |\mathcal{I}|$, the total number of i.i.d. noise terms in this scenario is then $N_\mathcal{J}N_\mathcal{I} = (N - N_\mathcal{I})N_\mathcal{I}$. Since $N_\mathcal{I}$ is binomially distributed with parameter $p = \frac{1}{2}$, we can calculate the unconditional expectation

$$\begin{aligned} \mathrm{E}\left\{\boldsymbol{\zeta}^\top[k]\boldsymbol{Q}\boldsymbol{\zeta}[k]\right\} &= \frac{(N-1)\sigma_\beta^2}{N}\mathrm{E}\left\{(N-N_\mathcal{I})N_\mathcal{I}\right\} \\ &= \frac{(N-1)^2\sigma_\beta^2}{4} \end{aligned}$$

and it follows from (22) that

$$\lim_{k\to\infty} d[k] < \frac{(N-1)^2\mu^2\sigma_\beta^2}{4N\epsilon} \tag{25}$$

in the fully-connected random broadcast scenario. A direct comparison between (25) and (23) can be misleading since the stepsize μ in the case of random broadcast is upper bounded by $\frac{4}{N}$. If we set $\mu = c\frac{4}{N}$ with $0 < c < 1$, we can write

$$\lim_{k\to\infty} d[k] < \frac{4(N-1)^2c^2\sigma_\beta^2}{N^3\epsilon} < \frac{4c^2\sigma_\beta^2}{N\epsilon}$$

which is similar to (24).

Similar results can also be derived for the offset distance from consensus metric by replacing σ_β^2 with σ_Δ^2. The presence of significant non-zero residual drift, however, may cause the offset distance from consensus bound to exceed the upper bound. These results are numerically verified and the effect of residual drift is numerically demonstrated in Section VII.

VII. Numerical Results

This section presents numerical results demonstrating convergence of drifts and offsets in networks with random pairwise message exchanges. The distance from consensus metric(s) are computed for the drifts and offsets in each timeslot and ensemble averaged over 5000 runs. In all of the results in this section, the initial offsets $\Delta_i[0]$ for $i = 1, \ldots, N$ were randomly generated as i.i.d. zero mean Gaussian random variables with standard deviation 5 ms. To generate a "worst-case" initialization for the drifts, the initial drifts were determined by first computing the eigenvector corresponding to $\lambda_{\max}(\boldsymbol{U}^\top(\bar{\boldsymbol{R}}^\top + \bar{\boldsymbol{R}} + \mu\bar{\boldsymbol{S}})\boldsymbol{U})$. Denoting this eigenvector as $\boldsymbol{v}$, we then set $\boldsymbol{\beta}[0] = \gamma\boldsymbol{U}\boldsymbol{v}$

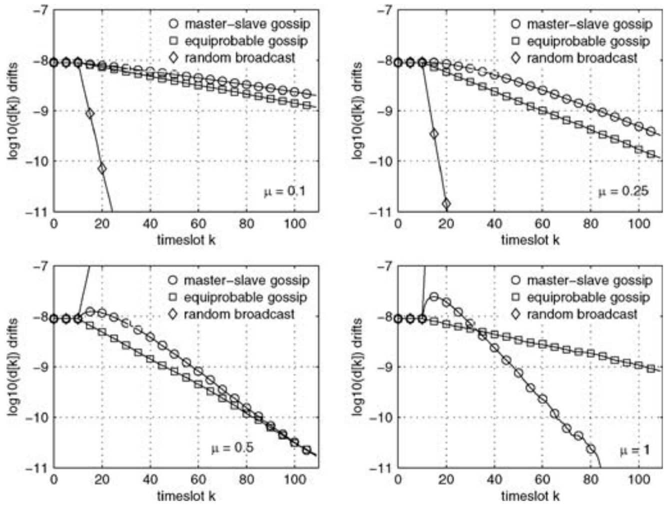


Fig. 2. Average drift distance from consensus metrics for three different $N = 10$ node network scenarios and $\mu \in \{0.1, 0.25, 0.5, 1\}$. Drift compensation begins at timeslot $k = 10$.

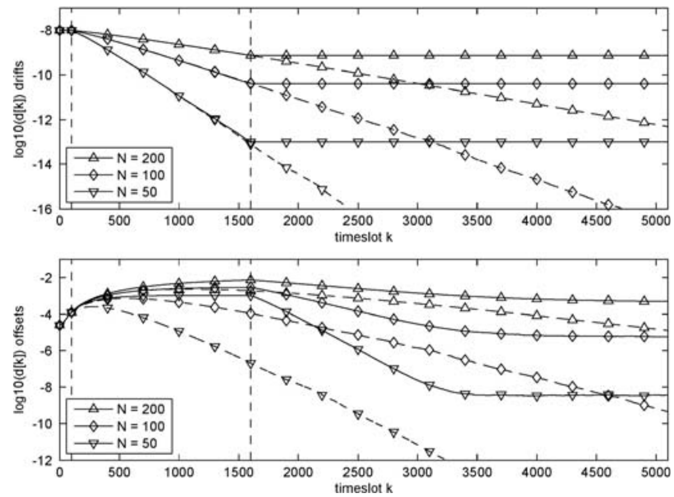


Fig. 3. Average distance from consensus metrics for $N = 50$, 100, 200 node asymmetric gossip networks with randomly deployed nodes and fixed stepsize $\mu = 0.5$. Solid lines correspond to separate drift and offset compensation. Dashed lines correspond to simultaneous drift and offset compensation.

where γ is a scale factor selected such that the root mean squared value $\boldsymbol{\beta}[0]$ is 100 $\mu\text{s}/\text{timeslot}$.

The results in Fig. 2 show the ensemble averaged drift distance from consensus metric in three different $N = 10$ node network scenarios:

1) Master-slave random asymmetric gossip network with $p_{i,1} = \frac{1}{9}$ for $i = 2, \ldots, 10$.
2) Fully-connected equiprobable random asymmetric gossip with $p_{i,j} = \frac{1}{90}$ for all $i \neq j$.
3) Fully-connected random broadcast network.

Note that the first scenario requires the use of Theorem 1 whereas the second scenario satisfies the conditions of Corollaries 2 and 3. Straightforward calculations result in stepsize bounds for monotonic mean squared convergence of

$$\mu \in \begin{cases} \left(0, \frac{2}{9}\right) & \text{scenario 1} \\ \left(0, \frac{10}{9}\right) & \text{scenario 2} \\ \left(0, \frac{4}{10}\right) & \text{scenario 3.} \end{cases}$$

Drift compensation begins at timeslot $k = 10$.

For the master-slave random asymmetric gossip case, we see that monotonic mean squared convergence occurs only for $\mu = 0.2$, in agreement with Theorem 1. The master-slave network exhibits general asymptotic mean squared convergence for all values of μ tested but convergence is non-monotonic when $\mu \geq \frac{2}{9}$. The fully-connected equiprobable random asymmetric case exhibits monotonic mean squared convergence for all tested values of μ, also in agreement with the analysis. The fastest rate of convergence is achieved in this case when $\mu = 0.5$, which is close to the optimum stepsize $\mu_{\text{opt}} = \frac{10}{18}$ as discussed in Section V-C. The random broadcast network exhibits rapid monotonic mean squared convergence for $\mu = 0.1$ and $\mu = 0.25$ and diverges for $\mu = 0.5$ and $\mu = 1$, which also agrees with the analysis. The rapid convergence of the fully-connected random broadcast consensus system is a consequence of the fact that there are on average $\mathrm{E}\{N_{\mathcal{I}}(N - N_{\mathcal{I}})\} = \frac{N(N-1)}{4} = 22.5$ pairwise message exchanges in each timeslot. This represents significantly more network resources than required by the single pairwise message exchange per timeslot of random asymmetric gossip.

For the particular stepsize value $\mu = 0.1$ where all three scenarios exhibit monotonic mean squared convergence, we have

$$(1-\epsilon)^k \approx \begin{cases} 0.9878^k & N = 10 \text{ master-slave gossip} \\ 0.9798^k & N = 10 \text{ equiprobable gossip} \\ 0.6250^k & N = 10 \text{ random broadcast} \end{cases}$$

which shows the worst-case convergence rates are in agreement with the $\mu = 0.1$ results shown in Fig. 2.

As an example of drift and offset consensus behavior for larger-scale networks, Fig. 3 shows the ensemble averaged distance from consensus metrics for a randomly deployed random asymmetric gossip network on a two-dimensional surface with $N = 50$, 100, 200 nodes and a fixed step size $\mu = 0.5$. The x, y coordinates of each node were uniformly generated on a 100 $\times$ 100 square and the probability $p_{i,j}$ was set to be inversely proportional to the Euclidean distance between nodes i and j, normalized to satisfy $\sum_i \sum_j p_{i,j} = 1$. The results in Fig. 3 are typical for one realization of the node deployments and are not averaged over the node positions. Since $\boldsymbol{P} = \boldsymbol{P}^\top$ in this example, Corollary 3 implies that monotonic mean squared convergence is achieved if and only if $0 < \mu < \frac{N}{N-1}$. The stepsize value of $\mu = 0.5$ satisfies this requirement and results in worst-case convergence rates of

$$(1-\epsilon)^k \approx \begin{cases} 0.9987^k & N = 200 \text{ randomly-deployed network} \\ 0.9975^k & N = 100 \text{ randomly-deployed network} \\ 0.9950^k & N = 50 \text{ randomly-deployed network.} \end{cases}$$

These results are consistent with the results in Fig. 3 where convergence times increase with the size of the network.

The solid lines in Fig. 3 correspond to the scenario considered in the analytical results with separate drift and offset compensation. In this case, drift compensation occurs in timeslots $k \in \{100, \ldots, 1599\}$ and offset compensation occurs in timeslots $k \geq 1600$. Since drift compensation ceases in timeslot $k = 1600$, there is some non-zero residual drift present when offset compensation commences. In this example, the presence of non-zero residual drifts for $k \geq 1600$ does not cause divergence of the offsets. Rather, the offsets converge to a floor established by the presence of the non-zero residual drifts.

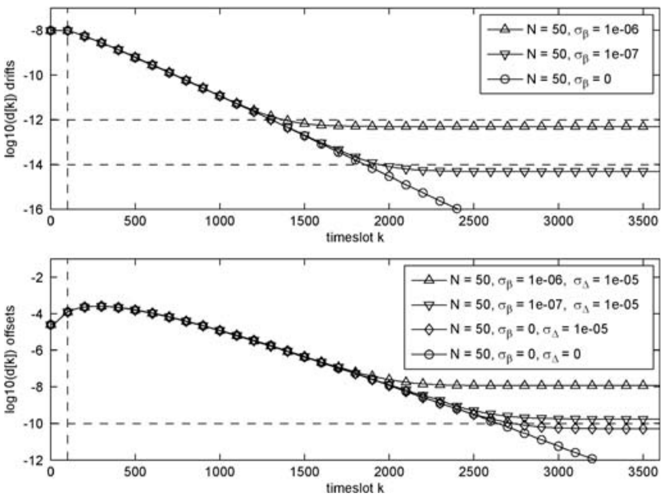


Fig. 4. Average distance from consensus metrics for a randomly deployed $N = 50$ node asymmetric gossip network with drift and offset estimation errors, fixed stepsize $\mu = 0.5$, and simultaneous drift and offset compensation for $k \geq 100$. Horizontal dashed lines correspond to analytical approximations from Section VI.

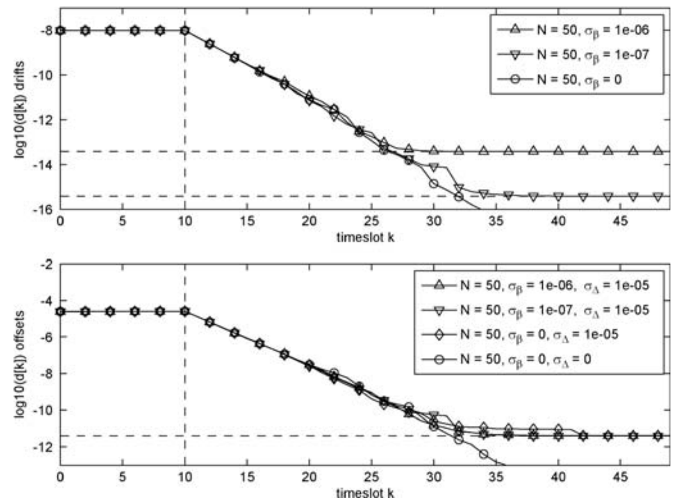


Fig. 5. Average distance from consensus metrics for a $N = 50$ node fully-connected random broadcast network with drift and offset estimation errors, fixed stepsize $\mu = \frac{2}{50}$, and simultaneous drift and offset compensation for $k \geq 10$. Horizontal dashed lines correspond to analytical approximations from Section VI.

This floor can be lowered by allowing the drift compensation to achieve a better distance from consensus metric prior to commencing offset compensation.

The dashed lines in Fig. 3 correspond to a scenario with *simultaneous drift and offset compensation* in timeslots $k \geq 100$. While analytical results in Section V were based on the assumption that drift compensation is performed prior to offset compensation, this example shows that simultaneous drift and offset compensation can improve the performance of the synchronization system. In particular, by performing drift compensation for all $k \geq 100$, the drift distance from consensus metric for continues to monotonically converge to zero, consistent with the analytical results, thus improving the drift synchronization and lowering the floor on the offset distance from consensus metric. Additionally, the offset distance from consensus metric is improved due to the fact that offset compensation is performed earlier in the synchronization process. The offset distance from consensus metric, however, is clearly non-monotonic in the initial adaptation. A convergence analysis of the offset distance from consensus metric for simultaneous drift and offset compensation is left as a potential extension to this work.

Fig. 4 shows an example of consensus behavior for an asymmetric gossip system with drift and offset *estimation error* and simultaneous drift and offset compensation in timeslots $k \geq 100$. Similar to Fig. 3, this example assumes a randomly deployed network on a two-dimensional surface with $N = 50$ nodes, fixed step size $\mu = 0.5$, and edge activation probability $p_{i,j}$ inversely proportional to the Euclidean distance between nodes i and j. The drift and offset estimation errors were modeled as independent zero-mean Gaussian random variables with $\zeta_{j,i}[k] \stackrel{\text{i.i.d.}}{\sim} \mathcal{N}(0, \sigma_\beta^2)$ and $\eta_{j,i}[k] \stackrel{\text{i.i.d.}}{\sim} \mathcal{N}(0, \sigma_\Delta^2)$ with standard deviations $\sigma_\beta \in \{0, 10^{-7}, 10^{-6}\}$ and $\sigma_\Delta \in \{0, 10^{-5}\}$.

The results in Fig. 4 show that the estimation errors do not have a significant effect on the initial convergence of the drifts or offsets but, as discussed in Section VI, do establish a non-zero floor on the achievable distance from consensus. The horizontal dashed lines in Fig. 4 correspond to the analytical bounds on the steady-state distance from consensus as derived in Section VI. In this example, the bounds are close to the actual steady-state performance for the drift distance from consensus metric and also for the offset distance from consensus metric when the residual drift distance from consensus is relatively small, e.g., $\sigma_\beta \leq 10^{-7}$. The offset distance from consensus metric does not satisfy the bound when the residual drift distance from consensus is relatively large, e.g., $\sigma_\beta = 10^{-6}$. In this regime, the residual drift errors prevent the offset distance from consensus from converging to values below the bound.

Fig. 5 shows a simulation similar to that in Fig. 4 except for a fully-connected random broadcast system with stepsize $\mu = \frac{2}{50}$. In this example, simultaneous drift and offset compensation occurs in timeslots $k \geq 10$. As with asymmetric gossip, we see that the estimation errors do not affect the initial convergence behavior and the steady-state behavior agrees closely with the bounds developed in Section VI. As shown previously, convergence occurs much more quickly in this case since the fully-connected random broadcast system exchanges many more messages and uses more network resources in each timeslot than random asymmetric gossip.

VIII. Conclusions and Extensions

This paper develops necessary and sufficient conditions for monotonic mean squared convergence of clock states in random consensus synchronization systems. Our results suggest that nonhierarchical embedded synchronization techniques which glean information from existing network traffic can be effective for low-overhead network synchronization. Simple explicit stepsize bounds are developed for random asymmetric gossip and fully-connected random broadcast scenarios. Numerical examples showing convergence and divergence of the consensus clock under different network topologies are also provided.

Potential extensions of this work include: (i) convergence analysis for simultaneous drift and offset compensation, (ii) the development of explicit bounds on the stepsize μ for unbalanced networks, (iii) an analysis of the precise effects of the stepsize μ, network size N, and/or the use of time-varying stepsizes on the steady-state distance from consensus metric for systems with drift and offset estimation error, and (iv) an analysis of asymmetric consensus techniques for nodes with stochastic clock dynamics.

Appendix A
Proof of Theorem 1

Proof: Given the state vector $\boldsymbol{x}[k]$ with dynamics in (5) and the distance from consensus metric defined in (1) with stepsize μ. We will show (9) is necessary and sufficient for (8) to hold for some $\epsilon > 0$ and all $\boldsymbol{x}[k]$ such that $d[k] > 0$. We denote $\boldsymbol{W}[k] = \boldsymbol{I}_N + \mu \boldsymbol{R}[k]$. Since $\boldsymbol{x}[k+1] = \boldsymbol{W}[k]\boldsymbol{x}[k]$, we have

$$d[k+1] = \frac{1}{N} \left\| \boldsymbol{Q}\boldsymbol{W}[k]\boldsymbol{x}[k] \right\|_2^2$$

and we can write (8) equivalently as

$$\frac{\boldsymbol{x}^\top[k] \mathrm{E}\left\{\boldsymbol{W}^\top[k]\boldsymbol{Q}^\top\boldsymbol{Q}\boldsymbol{W}[k]\right\}\boldsymbol{x}[k]}{\boldsymbol{x}^\top[k]\boldsymbol{Q}^\top\boldsymbol{Q}\boldsymbol{x}[k]} < 1 - \epsilon \qquad (26)$$

for some $\epsilon > 0$ and all $\boldsymbol{x}[k]$ such that $d[k] > 0$. Now, since $d[k] > 0$, we have $\boldsymbol{x}[k] \neq \boldsymbol{1}_N \bar{x}[k]$, and we can write

$$\begin{aligned}\boldsymbol{x}[k] &= \boldsymbol{1}_N \bar{x}[k] + \boldsymbol{x}[k] \\ &= \boldsymbol{1}_N \bar{x}[k] + \boldsymbol{U}\boldsymbol{v}[k]\end{aligned}$$

with $\boldsymbol{x}[k] \neq 0$ orthogonal to $\boldsymbol{1}_N$, $\boldsymbol{U} \in \mathbb{R}^{N \times (N-1)}$ composed of orthonormal columns all orthogonal to $\boldsymbol{1}_N$, and $\boldsymbol{v}[k] \in \mathbb{R}^{N-1}$ with $\boldsymbol{v}[k] \neq 0$.

Using the facts that $\boldsymbol{Q}\boldsymbol{U} = \boldsymbol{U}, \boldsymbol{Q}\boldsymbol{1}_N = \boldsymbol{0}, \boldsymbol{U}^\top\boldsymbol{U} = \boldsymbol{I}_{N-1}$, and $\boldsymbol{W}[k]\boldsymbol{1}_N = \boldsymbol{1}_N$ for all $\boldsymbol{W}[k]$, we can write (26) equivalently as

$$\frac{\boldsymbol{v}^\top[k]\boldsymbol{U}^\top \mathrm{E}\left\{\boldsymbol{W}^\top[k]\boldsymbol{Q}^\top\boldsymbol{Q}\boldsymbol{W}[k]\right\}\boldsymbol{U}v[k]}{\boldsymbol{v}^\top[k]\boldsymbol{v}[k]} < 1 - \epsilon \qquad (27)$$

for some $\epsilon > 0$ and all $\boldsymbol{v}[k] \in \mathbb{R}^{N-1}$. Focusing on the quadratic form in the numerator, we can define $\boldsymbol{A} := \boldsymbol{U}^\top \mathrm{E}\{\boldsymbol{W}^\top[k]\boldsymbol{Q}^\top\boldsymbol{Q}\boldsymbol{W}[k]\}\boldsymbol{U}$ and write

$$\begin{aligned}\boldsymbol{A} &= \boldsymbol{U}^\top \mathrm{E}\left\{\left(\boldsymbol{I}_N + \mu\boldsymbol{R}[k]\right)^\top \boldsymbol{Q}^\top\boldsymbol{Q}\left(\boldsymbol{I}_N + \mu\boldsymbol{R}[k]\right)\right\}\boldsymbol{U} \\ &= \boldsymbol{I}_{N-1} + \mu\boldsymbol{U}^\top\bar{\boldsymbol{R}}^\top\boldsymbol{U} + \mu\boldsymbol{U}^\top\bar{\boldsymbol{R}}\boldsymbol{U} + \mu^2\boldsymbol{U}^\top\bar{\boldsymbol{S}}\boldsymbol{U}\end{aligned}$$

where $\bar{\boldsymbol{R}} := \mathrm{E}\{\boldsymbol{R}[k]\}$ and $\bar{\boldsymbol{S}} := \mathrm{E}\{\boldsymbol{R}^\top[k]\boldsymbol{Q}^\top\boldsymbol{Q}\boldsymbol{R}[k]\}$. Hence, after cancellation of common terms on both sides of the inequality in (27), we have that (27) is equivalent to

$$\frac{\boldsymbol{v}^\top[k]\boldsymbol{U}^\top(\mu\bar{\boldsymbol{R}}^\top + \mu\bar{\boldsymbol{R}} + \mu^2\bar{\boldsymbol{S}})\boldsymbol{U}\boldsymbol{v}[k]}{\boldsymbol{v}^\top[k]\boldsymbol{v}[k]} < -\epsilon \qquad (28)$$

for some $\epsilon > 0$ and all $\boldsymbol{v}[k] \in \mathbb{R}^{N-1}$. Through a sequence of equivalencies, we have shown that (28) is equivalent to (8).

To show (9) is sufficient such that (28) holds for some $\epsilon > 0$ and all $\boldsymbol{v}[k] \in \mathbb{R}^{N-1}$, define

$$\boldsymbol{B} := \boldsymbol{U}^\top(\bar{\boldsymbol{R}}^\top + \bar{\boldsymbol{R}} + \mu\bar{\boldsymbol{S}})\boldsymbol{U}$$

and observe that the left hand side of (28) is a Rayleigh quotient and takes on values between the minimum and maximum eigenvalues of $\mu\boldsymbol{B}$ for all $\boldsymbol{v}[k] \in \mathbb{R}^{N-1}$. Hence $\mu\lambda_{\max}(\boldsymbol{B}) < 0$ implies the left hand side of (28) is strictly negative. Therefore, there exists $\epsilon > 0$ such that (28) is true for all $\boldsymbol{v}[k] \in \mathbb{R}^{N-1}$.

To show the necessity of (9), suppose $\mu\lambda_{\max}(\boldsymbol{B}) = c \geq 0$. Selecting $\boldsymbol{v}[k]$ to be equal to an eigenvector corresponding to this eigenvalue causes the left hand side of (28) to be equal to $c \geq 0$, hence (28) does not hold for any $\epsilon > 0$ for at least one $\boldsymbol{v}[k] \in \mathbb{R}^{N-1}$. ■

Acknowledgment

The authors would like to thank Prof. S. Dasgupta from the University of Iowa as well as the anonymous reviewers for their valuable comments and suggestions to improve the quality and clarity of this manuscript.

References

[1] D. Mills, "Internet time synchronization: The network time protocol," *IEEE Trans. Commun.*, vol. 39, no. 10, pp. 1482–1493, Oct. 1991.

[2] *IEEE Standard for a Precision Clock Synchronization Protocol for Networked Measurement and Control Systems*, IEEE Std. 1588, 2008.

[3] W. Lewandowski, J. Azoubib, and W. Klepczynski, "GPS: Primary tool for time transfer," *Proc. IEEE*, vol. 87, no. 1, pp. 163–172, Jan. 1999.

[4] J. Hahn and E. Powers, "Implementation of the GPS to Galileo time offset (GGTO)," in *Proc. Joint IEEE Int. Freq. Control Symp. Precise Time Time Interval (PTTI) Syst. Appl. Meet.*, Aug. 2005, pp. 33–212.

[5] P. Veríssimo, L. Rodrigues, and A. Casimiro, "Cesiumspray: A precise and accurate global clock service for large-scale systems," *J. Real-Time Syst.*, vol. 3, pp. 241–294, 1997.

[6] J. Elson, L. Girod, and D. Estrin, "Fine-grained network time synchronization using reference broadcasts," *ACM SIGOPS Oper. Syst. Rev.*, vol. 36, pp. 147–163, 2002.

[7] S. Ganeriwal, R. Kumar, and M. Srivastava, "Timing-sync protocol for sensor networks," in *Proc. ACM SenSys 2003*, New York, NY, USA, Nov. 2003, pp. 138–149.

[8] J. van Greunen and J. Rabaey, "Lightweight time synchronization for sensor networks," in *Proc. 2nd ACM Int. Workshop Wireless Sensor Netw. Appl. (WSNA)*, Sep. 2003, pp. 11–19.

[9] M. Sichitiu and C. Veerarittiphan, "Simple, accurate time synchronization for wireless sensor networks," in *Proc. IEEE Wireless Commun. Netw. (WCNC)*, Mar. 2003, vol. 2, pp. 1266–1273.

[10] F. Sivrikaya and B. Yener, "Time synchronization in sensor networks: A survey," *IEEE Netw.*, vol. 18, no. 4, pp. 45–50, Jul.–Aug. 2004.

[11] M. Maroti, B. Kusy, G. Simon, and A. Ledeczi, "The flooding time synchronization protocol," in *Proc. 2nd Int. Conf. Embedded Netw. Sensor Syst.*, Nov. 2004.

[12] W. Su and I. Akyildiz, "Time-diffusion synchronization protocol for wireless sensor networks," *IEEE/ACM Trans. Netw.*, vol. 13, no. 2, pp. 384–397, Apr. 2005.

[13] R. Solis, V. Borkar, and P. Kumar, "A new distributed time synchronization protocol for multihop wireless networks," in *Proc. IEEE Conf. Decision Control*, Dec. 2006, pp. 2734–2739.

[14] Q. Li and D. Rus, "Global clock synchronization in sensor networks," *IEEE Trans. Comput.*, vol. 55, no. 2, pp. 214–226, Feb. 2006.

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

[16] L. Schenato and G. Gamba, "A distributed consensus protocol for clock synchronization in wireless sensor network," in *Proc. IEEE Conf. Decision Control*, Dec. 2007, pp. 2289–2294.

[17] C. Wu, *Synchronization in Complex Networks of Nonlinear Dynamical Systems*. Hackensack, NJ, USA: World Scientific, 2007.

[18] G. Scutari, S. Barbarossa, and L. Pescosolido, "Distributed decision through self-synchronizing sensor networks in the presence of propagation delays and asymmetric channels," *IEEE Trans. Signal Process.*, vol. 56, no. 4, pp. 1667–1684, April 2008.

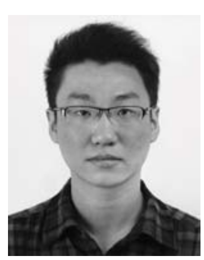
[19] R. Carli, A. Chiuso, L. Schenato, and S. Zampieri, "A PI consensus controller for networked clocks synchronization," in *Proc. IFAC World Congr.*, 2008.

[20] S. Bolognani, R. Carli, and S. Zampieri, "A PI consensus controller with gossip communication for clock synchronization in wireless sensors networks," in *Proc. 1st IFAC Workshop Estimat. Control Netw. Syst.*, 2009.

[21] Z. Li, Z. Duan, G. Chen, and L. Huang, “Consensus of multiagent systems and synchronization of complex networks: A unified viewpoint,” *IEEE Trans. Circuits Syst. I*, vol. 57, no. 1, pp. 213–224, Jan. 2010.
[22] R. Carli, E. D’Elia, and S. Zampieri, “A PI controller based on asymmetric gossip communications for clocks synchronization in wireless sensors networks,” in *Proc. IEEE Conf. Decision Control (CDC-ECC)*, 2011, pp. 7512–7517.
[23] J. He, P. Cheng, L. Shi, and J. Chen, “Time synchronization in WSNs: A maximum value based consensus approach,” in *Proc. IEEE Conf. Decision Control (CDC-ECC)*, 2011, pp. 7882–7887.
[24] M. Maggs, S. O’Keefe, and D. Thiel, “Consensus clock synchronization for wireless sensor networks,” *IEEE Sensors J.*, vol. 12, no. 6, pp. 2269–2277, Jun. 2012.
[25] N. Gresset and J. Letessier, “A random broadcast consensus synchronization algorithm for large scale wireless mesh networks,” in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, 2012, pp. 1573–1577.
[26] L. Kocarev, *Consensus and Synchronization in Complex Networks*. New York, NY, USA: Springer, 2013.
[27] Y. Hatano, A. Das, and M. Mesbahi, “Agreement in presence of noise: Pseudogradients on random geometric networks,” in *Proc. 44th IEEE Conf. Decision Control/Eur. Control Conf. (CDC-ECC)*, Dec. 2005, pp. 6382–6387.
[28] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah, “Randomized gossip algorithms,” *IEEE Trans. Inf. Theory*, vol. 52, no. 6, pp. 2508–2530, Jun. 2006.
[29] G. Picci and T. Taylor, “Almost sure convergence of random gossip algorithms,” in *Proc. IEEE Conf. Decision Control*, Dec. 2007, pp. 282–287.
[30] F. Fagnani and S. Zampieri, “Asymmetric randomized gossip algorithms for consensus,” in *Proc. IFAC World Congr.*, 2008, pp. 9051–9056.
[31] F. Fagnani and S. Zampieri, “Randomized consensus algorithms over large scale networks,” *IEEE J. Sel. Areas Commun.*, vol. 26, no. 4, pp. 634–649, May 2008.
[32] S. Kar and J. Moura, “Distributed consensus algorithms in sensor networks with imperfect communication: Link failures and channel noise,” *IEEE Trans. Signal Process.*, vol. 57, no. 1, pp. 355–369, Jan. 2009.
[33] A. Tahbaz-Salehi and A. Jadbabaie, “Consensus over ergodic stationary graph processes,” *IEEE Trans. Autom. Control*, vol. 55, no. 1, pp. 225–230, Jan. 2010.
[34] A. Jadbabaie, J. Lin, and A. Morse, “Coordination of groups of mobile autonomous agents using nearest neighbor rules,” *IEEE Trans. Autom. Control*, vol. 48, no. 6, pp. 988–1001, June 2003.
[35] L. Xiao, S. Boyd, and S. Lall, “A scheme for robust distributed sensor fusion based on average consensus,” in *Proc. IEEE Int. Symp. Inf. Process. Sensor Netw. (IPSN)*, 2005, pp. 63–70.
[36] L. Moreau, “Stability of multiagent systems with time-dependent communication links,” *IEEE Trans. Autom. Control*, vol. 50, no. 2, pp. 169–182, Feb. 2005.
[37] T. Aysal, M. Yildiz, A. Sarwate, and A. Scaglione, “Broadcast gossip algorithms for consensus,” *IEEE Trans. Signal Process.*, vol. 57, no. 7, pp. 2748–2761, July 2009.
[38] J.-W. Lee, “Uniform consensus among self-driven particles,” in *Hybrid Systems: Computation and Control*, ser. Lecture Notes in Computer Science, R. Majumdar and P. Tabuada, Eds. Berlin, Germany: Springer, 2009, vol. 5469, pp. 252–261.
[39] S. Ghosh and J.-W. Lee, “Equivalent conditions for uniform asymptotic consensus among distributed agents,” in *Proc. Amer. Control Conf. (ACC)*, Jun. 2010, pp. 4821–4826.
[40] T. Li, M. Fu, L. Xie, and J.-F. Zhang, “Distributed consensus with limited communication data rate,” *IEEE Trans. Autom. Control*, vol. 56, no. 2, pp. 279–292, Feb. 2011.
[41] E. Baghdady, R. Lincoln, and B. Nelin, “Short-term frequency stability: Characterization, theory, measurement,” *Proc. IEEE*, vol. 53, no. 7, pp. 704–722, July 1965.
[42] I. S. Ahn and R. G. Brown, “Assessing the validity of suboptimal 2-state clock models,” DTIC Document, Tech. Rep., 1986.
[43] L. Galleani, “A tutorial on the 2-state model of the atomic clock noise,” *Metrologia*, vol. 45, no. 6, pp. S175–S182, Dec. 2008.
[44] M. Koskenvuori, T. Mattila, A. Haara, J. Kiihamaki, I. Tittonen, A. Oja, and H. Seppa, “Long-term stability of single-crystal silicon microresonators,” *Sens. Actuators A, Phys.*, vol. 115, no. 1, pp. 23–27, 2004.
[45] D. Rife and R. Boorstyn, “Single-tone parameter estimation from discrete-time observations,” *IEEE Trans. Inf. Theory*, vol. 20, no. 5, pp. 591–598, Sep. 1974.
[46] R. Krishna and C. Hoene, “Calculating relative clock drifts using IEEE 802.11 beacons,” in *Proc. IEEE India Conf. (INDICON)*, 2009, pp. 1–4.
[47] G. Fettweis, E. Zimmermann, V. Jungnickel, and E. Jorswieck, “Challenges in future short range wireless systems,” *IEEE Veh. Technol. Mag.*, vol. 1, no. 2, pp. 24–31, June 2006.
[48] D. Parish, F. Farzaneh, and C. Barrat, “Methods and apparatus for calibrating radio frequency base stations using antenna arrays,” U.S. Patent 6 037 898, Oct. 1997.
[49] A. Bourdoux, B. Come, and N. Khaled, “Non-reciprocal transceivers in OFDM/SDMA systems: Impact and mitigation,” in *Proc. Radio Wireless Conf. (RAWCON)*, Aug. 2003, pp. 183–186.
[50] M. Guillaud, D. Slock, and R. Knopp, “A practical method for wireless channel reciprocity exploitation through relative calibration,” in *Proc. Int. Symp. Signal Process. Its Appl.*, Aug. 2005, vol. 1, pp. 403–406.
[51] D. R. Brown, III and A. G. Klein, “Precise timestamp-free network synchronization,” in *Proc. Conf. Inf. Sci. Syst. (CISS)*, Mar. 2013.

D. Richard Brown, III (S’97–M’00–SM’09) received the B.S. and M.S. degrees in Electrical Engineering from The University of Connecticut in 1992 and 1996, respectively, and received the Ph.D. degree in Electrical Engineering from Cornell University in 2000. From 1992–1997, he was with General Electric Electrical Distribution and Control. He joined the faculty at Worcester Polytechnic Institute (WPI) in Worcester, Massachusetts in 2000 and currently is an Associate Professor. He also held an appointment as a Visiting Associate Professor at Princeton University from August 2007 to June 2008. His research interests are currently in coordinated wireless transmission and reception, synchronization, distributed computing, and game-theoretic analysis of communication networks.

Andrew G. Klein (S’95–M’98–SM’12) received the B.S. degree from Cornell University, Ithaca, NY, the M.S. degree from the University of California, Berkeley, and the Ph.D. degree from Cornell University in 2006, all in electrical engineering. He joined the Department of Engineering and Design, Western Washington University (WWU), Bellingham, WA, in 2014 as an Assistant Professor. Previously, he was a member of the faculty at Worcester Polytechnic Institute (WPI), Worcester, MA, from 2007–2014, and he was a Post-Doctoral Researcher with Supélec/LSS, Paris, France, from 2006 to 2007. His current research interests include distributed communications and stochastic signal processing.



Rui Wang (S’14) received the B.S. degree in Electrical Engineering from Nanjing University of Aeronuatics and Astronautics (NUAA), China, in 2011. He received the M.S. degree in Electrical and Computer Engineering from Worcester Polytechnic Institute in 2013. He is currently a Ph.D. candidate at Worcester Polytechnic Institute. His research interests are currently in coordinated wireless transmission and reception, synchronization, and wireless power transfer.