Performance Analysis of Co-Phased Combining for Achieving Binary Consensus Over Fading Wireless Channels With Imperfect CSI

Venugopalakrishna Y. Ramakrishnaiah and Chandra R. Murthy, *Senior Member, IEEE*

***Abstract*—This paper considers the problem of achieving binary consensus among a set of nodes using physical layer communication over noisy wireless links. The channel state information (CSI) available at the nodes is imperfect due to practical estimation errors. Two schemes for updating the majority bit estimates at the nodes are contrasted: a linear minimum mean-squared error (LMMSE) based scheme and a co-phased combining scheme. The evolution of network consensus is modeled as a Markov chain, and the average transition probability matrix (TPM) is analytically derived for the co-phased combining scheme, whereas, for the LMMSE based scheme, the average TPM is computed through Monte Carlo simulations. The co-phased combining scheme is found to perform better at low to intermediate pilot SNRs, in addition to being analytically tractable and having lower computational complexity, compared with the LMMSE-based scheme. Also, to further characterize the consensus behavior, the probability of accurate consensus, the second eigenvalue of the TPM, the average hitting time to the first consensus state, and the average consensus duration are derived for the co-phased combining scheme. The power allocation between the pilot and data symbols is optimized, subject to a total power constraint. It is found that the optimal power allocation can lead to a significant improvement in the consensus performance. Monte Carlo simulation results validate the theoretical results, and provide insights into the complexity and performance tradeoffs involved.**

***Index Terms*—Binary consensus, co-phasing, transition probability matrix, hitting time, consensus duration.**

I. Introduction

THIS work considers the problem of achieving binary consensus among a set of nodes, where, starting with an initial binary value, the nodes exchange messages, with the goal of agreeing upon the majority value among them. Majority consensus has many applications, for example, in cooperative spectrum sensing in cognitive radios [2], [3] and throughput optimization in sensor networks [4]. The classical approach to achieving consensus considers error-free exchange of data between neighboring nodes in the form of packets. In wireless sensor networks (WSNs), one can avoid the control information overhead required for a packet [5] by broadcasting the binary symbols over the wireless medium. This constitutes achieving binary consensus in the physical layer itself. However, the estimation of the fading channel and the errors due to the noise at the receiver lead to a new set of challenges in achieving accurate consensus, due to the unreliable communication between the nodes. In this context, our goal in this paper is to study the performance of binary consensus protocols that involve bit exchanges among the nodes over fading wireless channels, when the channel state information (CSI) at the nodes is imperfect.

Most of the existing literature on consensus problems, starting from the thesis of Tsitsiklis [6] to the more recent studies involving gossip algorithms [7]–[13], is mainly concerned with distributed averaging. Here, the nodes attain consensus by repeatedly computing a weighted average of the values of their neighbors. This requires nodes to exchange real-valued estimates. On the other hand, several recent studies have also considered the transmission of quantized states to the neighbors and then attaining the average quantized consensus state [14]–[23]. However, all of these works consider message exchange over the network layer of the protocol stack. In contrast, the idea of achieving binary consensus over the physical layer itself has only been explored more recently [24]–[29]. The exchange of a real-valued test statistic over fading channels to arrive at consensus on the global average was considered in [24]–[26], with the goal of performing distributed hypothesis testing. In [27]–[29], the authors considered a scheme where the nodes iteratively broadcast their majority-bit estimates over a noisy channel in a round-robin manner and update their majority-bit estimates using the received signals, to achieve consensus among the nodes. Further, this scheme has been extended to networks with link failures in [29]. However, these studies assumed perfect CSI to be available at the nodes, which is impractical in low power sensor network applications.

In [27], the authors consider an AWGN channel, and use a sum of votes estimator at every node to detect the majority bit. The evolution of the network state is modeled as a Markov chain, and the second largest eigenvalue of its transition probability matrix is used to characterize the consensus behavior. A key finding of this work is that the network asymptotically loses memory of the initial state. However, in the transient period, the

Manuscript received May 28, 2015; revised December 03, 2015; accepted January 14, 2016. Date of publication February 08, 2016; date of current version April 27, 2016. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Subhrakanti Dey. This work was financially supported by a research grant from the Aerospace Network Research Consortium. It has appeared in part in the *Proceedings of the IEEE Global Communications Conference*, December 2014 [1].

The authors are with the Department of Electrical Communication Engineering, Indian Institute of Science, Bangalore 560012, India (e-mail: yrv@ece.iisc.ernet.in; cmurthy@ece.iisc.ernet.in).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSP.2016.2526962

network can be in accurate consensus with high probability and can stay in consensus for long duration.

In [29], the authors studied binary consensus with i.i.d. fading channels between the nodes. A linear MMSE (LMMSE) estimator was proposed for updating the sum of votes at the nodes as a weighted combination of the received samples. The weights are computed based on the channel magnitudes, which were assumed to be perfectly known at the nodes. An alternative way to estimate the sum of votes is to simply compute the sum of the co-phased received samples (akin to equal gain combining). To estimate the sum of votes, which is a key quantity in determining the majority bit update at the nodes, the physical layer protocols considered in this work and related literature are based on spatial diversity combining, of which, the LMMSE-based scheme and co-phased combining scheme are robust across a wide range of SNRs. Other techniques, namely, maximum ratio combining and selective gain combining are not suitable under non-identical inputs [30]. Thus, the two schemes considered, the LMMSE-based scheme and the co-phased combining are the relevant protocols in the scope of this work. Furthermore, in the context of decentralized detection, the co-phased combining scheme is known to be robust to channel estimation errors [30].

In this work, we study the efficacy of the two aforementioned options for achieving binary consensus over i.i.d. fading channels. Specifically, and for the first time in the literature to the best of our knowledge, our analysis accounts for the effect of channel estimation errors on physical layer consensus protocols. Further, while the analysis in [29] focused on the second largest eigenvalue of the average transition probability matrix (TPM) of the Markov chain, we analyze metrics that are more directly related to the consensus performance, such as the probability of detecting the correct majority bit, the time to reach first consensus, and the average consensus duration, all for the co-phased combining scheme. In addition, we study the allocation of power between the pilot and data symbols, under a total power constraint, to optimize the consensus performance. We also study the performance of the co-phased combining scheme when the network is not fully connected.

Our key findings are as follows. The co-phased combining scheme is simple from an implementation perspective, and also makes the average probability of detecting the correct majority bit analytically tractable. The latter is necessary for computing the average TPM of the Markov chain. The second largest eigenvalue of the TPM governs the consensus behavior of the underlying protocol. We also characterize the average hitting time and the average consensus duration, which jointly determine the time required for the network to achieve consensus with high probability. We also find that, while the co-phased combining scheme is simpler than the LMMSE scheme, somewhat surprisingly, it offers better performance at low to moderate SNRs. This is due to its lesser dependence on the channel estimates, which makes it a robust scheme in the presence of errors in estimation. Finally, we find that optimizing the power allocation between pilot and data symbols is important, as it can lead to a significant improvement in the consensus performance compared to naïve allocation schemes.

The rest of the paper is organized as follows. We discuss the problem setup and system model in the next section. We present the update rules employed after the bit exchange in Section III. In Section IV, we derive the average probability of majority bit detection, average consensus duration and average hitting time for the proposed co-phased combining based scheme. In Section V, we provide simulation results comparing the proposed and the existing consensus protocols. We conclude the paper in Section VI.

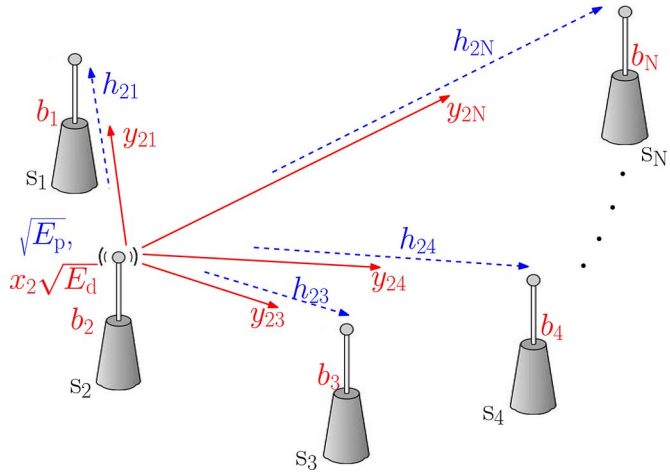

Fig. 1. Illustration of the physical layer consensus protocol. Node s_2 broadcasts a pilot symbol followed by a BPSK data symbol $x_2 \in \{-1, +1\}$. In each bit-exchange cycle, all nodes employ the same mechanism, in a round-robin manner.

Notation: In this paper, we use boldface capital letters to denote matrices and boldface small letters to denote vectors. We use $(\cdot)^T$ to denote the transpose of a matrix. The function $f_H(h)$ represents the probability density function (pdf) of a random variable H, and h denotes its realization.

II. System Model and Problem Set-Up

Our set-up consists of a fully-connected network with N nodes, with the nodes denoted by $\mathcal{S} \triangleq \{s_1, s_2, \ldots, s_N\}$. Each node starts with an initial binary data bit $b_l(0) \in \{0, 1\}$, $l = 1, \ldots, N$. The goal is for the nodes to achieve consensus on the bit value corresponding to the majority of their initial values. To this end, nodes broadcast their data bit in a round-robin manner over noisy fading channels. Then, nodes utilize the received signals to detect and update a local majority bit estimate. This process is repeated over multiple cycles, as consensus may not be attained in a single cycle of bit exchange and bit update due to the noisy communication between the nodes. The physical layer protocol considered in this work is schematically illustrated in Fig. 1.

In the tth update cycle, node $s_k \in \mathcal{S}$ broadcasts a known pilot symbol followed by its current majority bit estimate, denoted by $b_k(t-1)$. This is received by the other $N-1$ nodes $s_l \in \mathcal{S} \setminus \{s_k\}$. The wireless channel from s_k to s_l, denoted by $h_{kl} \triangleq |h_{kl}| e^{j\theta_{kl}}$, is assumed to be Gaussian $\mathcal{CN}(0, \sigma^2)$ distributed, remain constant over a given update cycle, and vary in an i.i.d. fashion across the nodes and across update cycles. We note that the consensus protocols considered in the sequel are directly applicable to non-i.i.d. fading channels. The performance analysis can also be extended to the non-i.i.d. case, by using the techniques in recent work in the context of the outage

analysis of relay selection schemes [31], [32]. However, in the sequel, we restrict to the i.i.d. model for simplicity of exposition and because it is sufficient to bring out the critical importance of the system parameter settings, for example, the power allocation for the pilot and data symbols, on the performance of physical layer consensus protocols. The node s_l estimates the channel h_{kl} using the received pilot symbol

$$y_{kl}^{(p)} = h_{kl}\sqrt{E_p} + w_{kl}^{(p)}, \tag{1}$$

where E_p is the pilot power and $w_{kl}^{(p)} \sim \mathcal{CN}(0, \sigma_w^2)$ is the circularly symmetric complex additive white Gaussian noise at the receiver, with zero mean and variance σ_w^2. The maximum likelihood estimate of the channel is given by $\hat{h}_{kl} = h_{kl} + w_{kl}^{(p)}/\sqrt{E_p}$. With BPSK signaling employed for broadcasting the data bits, the received signal at s_l is given by

$$y_{kl}^{(d)} = h_{kl}x_k\sqrt{E_d} + w_{kl}^{(d)}, \quad \text{where} \quad x_k = 2b_k(t-1) - 1$$

and E_d is the data power and $w_{kl}^{(d)} \sim \mathcal{CN}(0, \sigma_w^2)$. At the end of a cycle, s_l has $N-1$ data samples $\{y_{kl}^{(d)}\}$ and corresponding channel estimates $\{\hat{h}_{kl}\}$, $k = 1, 2, \ldots, N$, $k \neq l$. These samples are used to find the bit-update of the tth cycle. The majority-bit update rules considered in this work are discussed in the next section. This is followed by a rigorous theoretical analysis of the consensus procedure.

III. Majority Bit Detection

In this section, we discuss the bit-update procedure at each node s_l. First, the received data $y_{kl}^{(d)}$ is pre-processed using an estimate of the phase of the channel h_{kl}, as follows:

$$\begin{aligned} r_{kl} &= \mathrm{Re}\left\{y_{kl}^{(d)}e^{-j\hat{\theta}_{kl}}\right\}, k \in \{1, 2, \ldots, N\}, k \neq l, \\ r_{kl} &= |h_{kl}|\cos\tilde{\theta}_{kl}x_k\sqrt{E_d} + v_{kl}, \end{aligned} \tag{2}$$

where the estimated channel $\hat{h}_{kl}$ is written as $\hat{h}_{kl} \triangleq |\hat{h}_{kl}|e^{j\hat{\theta}_{kl}}$, $\tilde{\theta}_{kl} \triangleq \theta_{kl} - \hat{\theta}_{kl}$ is the phase estimation error, and $v_{kl} \triangleq \mathrm{Re}\{w_{kl}e^{-j\hat{\theta}_{kl}}\} \sim \mathcal{N}(0, \sigma_w^2/2)$ is the AWGN at the receiver. The samples $\{r_{kl}\}$, $k = 1, 2, \ldots, N, k \neq l$ are used to detect and update the majority bit at node s_l.[1] Since the sum of votes, $\Delta_l \triangleq \sum_{k=1,k\neq l}^{N} x_k$ is indicative of the majority bit at node s_l, where $x_k \in \{-1, +1\}$ is the BPSK symbol corresponding to $b_k(t-1)$, we use the majority bit decision rule:

$$g(\hat{\Delta}_l) \triangleq \begin{cases} 1 & \hat{\Delta}_l \geq 0 \\ 0 & \text{otherwise}, \end{cases} \tag{3}$$

where $\hat{\Delta}_l$ is an estimate of sum of votes, computed at node s_l. In this work, two schemes are considered for obtaining $\hat{\Delta}_l$: 1) an LMMSE-based scheme and 2) a co-phased combining scheme.

LMMSE-Based Scheme [29]: Here, a weighted linear combination of the processed samples $\{r_{kl}\}$ is considered as an estimate of Δ_l, i.e., $\hat{\Delta}_l^{(wc)} = \boldsymbol{\alpha}_l^T\mathbf{r}_l$, where $\boldsymbol{\alpha}_l \triangleq [\alpha_{1l}\alpha_{2l}\ldots\alpha_{kl}\ldots\alpha_{Nl}]_{k\neq l}^T$ and $\mathbf{r}_l \triangleq [r_{1l}r_{2l}\ldots r_{kl}\ldots r_{Nl}]_{k\neq l}^T$. Here, the superscript wc stands for weighted combining. The optimal weight vector $\boldsymbol{\alpha}_l^*$ is evaluated by solving the MMSE estimation problem

$$\boldsymbol{\alpha}_l^* = \arg\min_{\boldsymbol{\alpha}_l} \mathbb{E}\left[\left(\hat{\Delta}_l^{(wc)} - \Delta_l\right)^2\right]. \tag{4}$$

On substituting the MMSE solution $\alpha_{kl}^* = \frac{|\hat{h}_{kl}|}{|\hat{h}_{kl}|^2 + \sigma_w^2/2}$, the sum of votes estimate can be expressed as

$$\hat{\Delta}_l^{(wc)} = \sum_{k=1,k\neq l}^{N}\left[\alpha_{kl}^*|h_{kl}|\cos\tilde{\theta}_{kl}x_k\sqrt{E_d} + \alpha_{kl}^*v_{kl}\right]. \tag{5}$$

We note that the above LMMSE-based scheme uses the estimated channels $\hat{h}_{kl}$ to compute the combining weights, in contrast to the scheme in [29], which assumed the availability of perfect channel state information h_{kl} at the nodes.

Co-Phased Combining Scheme: In this scheme, the sum of co-phased samples r_{kl} is used as an estimate of Δ_l, i.e.,

$$\hat{\Delta}_l^{(cc)} = \sum_{k=1,k\neq l}^{N}[|h_{kl}|\cos\tilde{\theta}_{kl}x_k\sqrt{E_d} + v_{kl}]. \tag{6}$$

The sum of votes estimate in (6) only requires the channel phase, while (5) depends on both the gain and the phase. Due to this, the consensus behavior of the two schemes in the face of fading and the imperfect CSI at the nodes can be different. It is of interest to study their relative performance.

Note that, $\hat{\Delta}_l^{(wc)}$ and $\hat{\Delta}_l^{(cc)}$ can be compactly expressed as

$$\hat{\Delta}_l = h\sqrt{E_d} + v, \tag{7}$$

where we term $h \triangleq \sum_{k=1,k\neq l}^{N}\beta_{kl}|h_{kl}|\cos\tilde{\theta}_{kl}x_k$ as the *effective channel-symbol*, and the noise $v \sim \mathcal{N}(0, \sigma_v^2)$ with $\sigma_v^2 = \sum_{k=1,k\neq l}^{N}(\beta_{kl})^2\sigma_w^2/2$. The parameter $\beta_{kl} = \alpha_{kl}^*$ for the LMMSE-based scheme, and $\beta_{kl} = 1$ for the co-phased combining scheme.

IV. Performance Analysis

We define the *state* of the network at time t as an ordered collection of decision bits (majority bit estimates), $[b_1(t)b_2(t) \ \ldots \ b_N(t)]$. After each bit-update, the network can attain any one of the $M = 2^N$ possible states with a probability that depends solely on the previous state and the current received pilot and data samples at the nodes. That is, the network state evolves as a first order discrete-time Markov chain (MC). The accurate consensus state of the MC is either the all-ones state or the all-zeros state, depending on whether the initial majority bit is one or zero. This MC is time inhomogeneous, as the channels are time-varying across cycles. For time inhomogeneous MCs with independent random transition probability matrices across the time steps, the marginal state distribution of the MC, i.e., the state distribution averaged over the randomness of the transition probability matrices, is precisely that of a *time homogeneous MC* with the *average TPM* [33]. Therefore, the average of the TPM over the distribution of the channel states is used to analyze the consensus behavior

[1]For simplicity, we ignore the lth sensor's own data bit, x_l. This is fine when the bit distribution across the nodes is such that any one node's data bit does not alter the majority vote. In Section V, we will show through simulations that ignoring the self-bit only marginally affects the consensus performance.

of the network.[2] The average state distribution vector at time t, denoted by $\bar{\boldsymbol{\pi}}(t) \in [0,1]^N$, is thus given by $\bar{\boldsymbol{\pi}}(t) = (\bar{\mathbf{P}})^t \boldsymbol{\pi}(0)$. Here, $\bar{\mathbf{P}} \triangleq [\bar{P}_{ij}]$ is the average TPM, $\boldsymbol{\pi}(0)$ is the initial state distribution vector, and $\bar{P}_{ij}$ is the average probability of transition (averaging over the channel state distribution) from state $\phi^{(j)}$ to state $\phi^{(i)}$ in one cycle, where $\{\phi^{(i)}, i \in \{1,2,\ldots,M\}\}$ represents the different possible states of the MC.

Due to the additive noise term in (1), $0 < \bar{P}_{ij} < 1 \; \forall i,j \in \{1,2,\ldots,M\}$ (see Appendix F for the proof), and hence, the MC is irreducible and aperiodic. Thus, the stationary state distribution $\bar{\boldsymbol{\pi}}_\infty \triangleq \lim_{t\to\infty} \bar{\boldsymbol{\pi}}(t)$ is independent of the initial state [34], as $\bar{\boldsymbol{\pi}}_\infty = \bar{\mathbf{P}}\bar{\boldsymbol{\pi}}_\infty$ and $\mathbf{1}^T\bar{\boldsymbol{\pi}}_\infty = 1$. However, the network can still achieve accurate consensus with high probability in the transient period of the MC.

From Perron's and Gershgorin's theorems [35], it is known that, for the positive stochastic matrix $\bar{\mathbf{P}}$, the largest eigenvalue λ_1 is 1 and is of multiplicity one, and the absolute value of all the other eigenvalues is strictly less than 1. The second largest eigenvalue λ_2 of $\bar{\mathbf{P}}$ characterizes the transient duration of the MC [36]. The closer λ_2 is to unity, the longer the transient period, and the higher the probability that the network attains and stays in consensus for a long duration. Thus, λ_2, along with the probability of attaining accurate consensus starting from an arbitrary state, are important performance metrics for understanding the consensus behavior of the network.

In this work, two more quantities, the average hitting time and the average consensus duration, are analyzed. There are no absorbing states due to the noisy channels, and hence the average hitting time and average consensus duration are also important metrics to study the consensus behavior. The average hitting time is the average number of cycles required to attain the consensus for the first time starting from an initial state $\phi^{(j)}$. The average consensus duration is the average number of cycles for which the network will stay in consensus once it is attained. These metrics are used to compare the performance of the two bit-update procedures and quantify the effect of channel estimation errors on the consensus performance. Finally, the tradeoff between the power allowed for pilot and data symbols on the consensus behavior is studied.

As discussed above, the consensus behavior of the network is determined by the average TPM of the MC. However, in the presence of channel estimation errors, the average TPM is analytically intractable for the LMMSE-based scheme because of its dependence on the channel gain estimates. The analysis for the co-phased combining scheme is presented below.

A. Probability of Detecting the Majority Bit

The probability of correctly detecting the majority bit from the received data symbols is necessary to determine the TPM of the MC, and is critical to the consensus behavior of the network. Since the noise is Gaussian distributed, conditioned on K, the number of sensors transmitting a $+1$, the average probability $\bar{p}_l$ of detecting the majority bit as '1' at node s_l is given by

$$\begin{aligned}\bar{p}_l &= \mathbb{E}[\Pr\{\hat{\Delta}_l \geq 0 | H = h\}] \\ &= \int_{-\infty}^{\infty} \mathcal{Q}\left(\frac{-h\sqrt{E_d}}{\sigma_v}\right) f_H(h)\,\mathrm{d}h, \end{aligned} \tag{8}$$

where $f_H(h)$ is the pdf of the effective channel-symbol H and $\mathcal{Q}(\cdot)$ is the Gaussian Q-function.

Suppose a subset of the nodes, $\mathcal{K} \subseteq \mathcal{S}\backslash\{s_l\}, |\mathcal{K}| = K$, transmit $+1$, and the remaining $N-K-1$ nodes in $\mathcal{K}^c \triangleq \mathcal{S}\backslash\{\mathcal{K}, s_l\}$ transmit -1. Then, the effective channel-symbol h in (7) can be expressed as $h = h_p - h_n$, where $h_p \triangleq \sum_{s_k\in\mathcal{K}} |h_{kl}| \cos\tilde{\theta}_{kl}$ and $h_n \triangleq \sum_{s_k\in\mathcal{K}^c} |h_{kl}| \cos\tilde{\theta}_{kl}$. Let H_p and H_n be random variables corresponding to the realizations h_p and h_n, respectively. The pdf of the sum of K weighted i.i.d. Rayleigh r.v.s, H_p, is not available in closed-form. However, it is well approximated by the pdf of a Nakagami r.v. with shape parameter, $m_1 = (\mathbb{E}[H_p^2])^2/\mathrm{Var}[H_p^2]$, and spread parameter, $\Omega_1 = \mathbb{E}[H_p^2]$. Similarly, h_n can be approximated as a Nakagami r.v. with shape and spread parameters m_2 and Ω_2, respectively. At high pilot SNRs, the weights are of nearly unit magnitude, and approximation error turns out to be particularly small. Also, as found through simulations, the approximation error reduces with increasing K. The derivation of $\mathbb{E}[H_p^2]$ and $\mathrm{Var}[H_p^2]$, required for evaluating the parameters m_1, m_2, Ω_1 and Ω_2, is presented in Lemma 1. The resulting analytical expression of the approximated pdf $f_H(h)$ is presented in Lemma 2.

Lemma 1: For a given pilot SNR, $\mathrm{SNR}_p \triangleq \sigma^2 E_p/\sigma_w^2$, and with the second moment of i.i.d. Rayleigh r.v.s H_{kl}, $\mathbb{E}[H_{kl}^2] = \sigma^2$, the mean and variance of the r.v. H_p^2 defined above are given by

$$\mathbb{E}\left[H_p^2\right] = \frac{K\sigma^2(2+(4+(K-1)\pi)\gamma_p\sigma^2)}{1+\gamma_p\sigma^2} \tag{9}$$

$$\begin{aligned}\mathrm{Var}\left[H_p^2\right] &= K\mathbb{E}\left[G_{kl}^4\right] + 3K(K-1)\left(\mathbb{E}\left[G_{kl}^2\right]\right)^2 \\ &\quad + K(K-1)(K-2)(K-3)(\mathbb{E}[G_{kl}])^4 \\ &\quad + 6K(K-1)(K-2)(\mathbb{E}[G_{kl}])^2\mathbb{E}\left[G_{kl}^2\right] \\ &\quad + 4K(K-1)\mathbb{E}\left[G_{kl}^3\right]\mathbb{E}[G_{kl}] - \left(\mathbb{E}\left[H_p^2\right]\right)^2, \end{aligned} \tag{10}$$

where $\gamma_p \triangleq E_p/\sigma_w^2$ and $G_{kl} \triangleq |H_{kl}|\cos\tilde{\theta}_{kl}$; and closed form expressions for $\mathbb{E}[G_{kl}], \mathbb{E}[G_{kl}^2], \mathbb{E}[G_{kl}^3]$ and $\mathbb{E}[G_{kl}^4]$ are provided in the proof.

Proof: See Appendix A. ∎

Lemma 2: The pdf of the effective channel-symbol is given by

$$\begin{aligned} f_H(h) &= \frac{2\left(\frac{m_1}{\Omega_1}\right)^{m_1}\left(\frac{m_2}{\Omega_2}\right)^{m_2}\mathrm{e}^{-\frac{h^2 m_1 m_2}{m}}}{\Gamma(m_1)\Gamma(m_2)\left(\frac{m}{\Omega}\right)^{m_1+m_2-\frac{1}{2}}} \\ &\quad \times \sum_{i=0}^{2m_1-1}\sum_{j=0}^{2m_2-1}\binom{2m_1-1}{i}\binom{2m_2-1}{j} \\ &\quad \times \left(\frac{m_2\Omega_1 h}{\sqrt{m\Omega}}\right)^{2m_1-1-i}\left(\frac{-m_1\Omega_2 h}{\sqrt{m\Omega}}\right)^{2m_2-1-j} \\ &\quad \times \Gamma\left(\frac{i+j+1}{2}, \frac{(m_1\Omega_2 h)^2}{m\Omega}\right), \end{aligned} \tag{11}$$

[2] An alternative model, under slowly varying channels, is to assume that the channel stays fixed for the duration of the consensus cycles. In this case, we would work with the TPM conditioned on the joint channel states between the nodes. The resulting MC is time homogeneous, making the analysis easier than the case considered in this paper.

for $h \geq 0$, where $m \triangleq m_1\Omega_2 + m_2\Omega_1$, $\Omega \triangleq \Omega_1\Omega_2$ and $\Gamma(\cdot,\cdot)$ is the upper incomplete Gamma function. For $h < 0$, $f_H(h)$ can be evaluated by swapping the parameters m_1, Ω_1 with m_2, Ω_2, respectively.

Proof: See Appendix B. ∎

To obtain the result in Lemma 2, the parameters $2m_1$ and $2m_2$ are rounded-off to the nearest integers. The $f_H(h)$ given by Lemma 2 can now be substituted into (8) to evaluate $\bar{p}_l$ as a single integral, this is discussed in Appendix C. Note that, the average probability $\bar{p}_l$ is conditioned on K, the number of sensors transmitting $+1$, but this dependence is not explicitly indicated to keep the notation light.

B. Average Transition Probability Matrix

Suppose the network is in a state $\phi^{(j)}$ at time $t-1$ and $\phi^{(i)} \triangleq [b_1^{(i)} b_2^{(i)} \ldots b_N^{(i)}]$ at time t. Let us denote the average probability of node s_l detecting bit '1' at time t, conditioned on $\phi^{(j)}$, by $\bar{p}_l^{(j)}$. At node s_l, the average probability of the bit being updated to $b_l^{(i)}$ is given by $b_l^{(i)}\bar{p}_l^{(j)} + (1-b_l^{(i)})(1-\bar{p}_l^{(j)})$. The update decisions $b_l^{(i)}$ at each of the nodes $s_l, l = 1, 2, \ldots, N$, are independent, as the receiver thermal noise and the wireless channels between the nodes are independent. Thus, the average probability of going from $\phi^{(j)}$ to $\phi^{(i)}$ in one cycle is given by

$$\bar{P}_{ij} = \prod_{l=1}^{N} \left[b_l^{(i)}\bar{p}_l^{(j)} + \left(1 - b_l^{(i)}\right)\left(1 - \bar{p}_l^{(j)}\right)\right], \tag{12}$$

for $i, j \in \{1, 2, \ldots, M\}$. The average probability of going from $\phi^{(j)}$ to an all-one state $\phi^{(M)}$ or to an all-zero state $\phi^{(1)}$ is

$$\bar{P}_{Mj} = \prod_{l=1}^{N} \bar{p}_l^{(j)} \tag{13}$$

$$\bar{P}_{1j} = \prod_{l=1}^{N} \left(1 - \bar{p}_l^{(j)}\right). \tag{14}$$

We have thus determined the average TPM $\bar{\mathbf{P}}$ of the MC. Note that, the average probability of detecting bit '1' conditioned on the all-zero state, $\bar{p}_l^{(1)}$, is the same for all the nodes. Hence, hereafter, $\bar{p}_l^{(1)}$ is denoted as $\bar{p}^{(1)}$, and similarly, $\bar{p}_l^{(M)}$ is denoted as $\bar{p}^{(M)}$. In [29], a simple approximation to the second largest eigenvalue λ_2 of the average TPM $\bar{\mathbf{P}}$ is shown to be $1 - 2\bar{p}^{(1)}$. This approximation is derived for the Rayleigh fading channel by linearizing the Q-function. Moreover, for $N = 2$ or 3 sensors, it can be shown that λ_2 is exactly $1 - 2\bar{p}^{(1)}$. The derivation of the second largest eigenvalue for $N = 3$ sensors is discussed in Appendix D. With larger N, the second eigen value would be much closer to 1. This is because, the probability of detecting bit 1 when all the nodes have bit 0, $\bar{p}^{(1)}$, tends to zero roughly exponentially with increasing N, and consequently, the second largest eigen value tends to 1. As mentioned earlier, the closer the second largest eigenvalue to unity, the better the consensus behavior of the network.

C. Average Hitting Time

Starting from a state $\phi^{(j)}$, the average number of cycles taken to reach a consensus state for the first time is termed as the average hitting time. The probability of reaching state $\phi^{(i)}$ for the first time starting from state $\phi^{(j)}$ after exactly n cycles, denoted by $f_{ij}^{(n)}$, can be recursively expressed as

$$f_{ij}^{(n)} = \sum_{\substack{k=1 \\ k \neq i}}^{M} \bar{P}_{kj} f_{ik}^{(n-1)}. \tag{15}$$

The vector of probabilities $\mathbf{f}_i^{(n)} \triangleq [f_{i1}^{(n)}\ f_{i2}^{(n)}\ \ldots\ f_{ij}^{(n)}\ \ldots\ f_{iM}^{(n)}]^T$, $j \neq i$ can be expressed as

$$\mathbf{f}_i^{(n)} = \mathbf{Q}^T \mathbf{f}_i^{(n-1)}, \tag{16}$$

where $\mathbf{Q}$ is the TPM obtained by removing the ith row and ith column of average TPM $\bar{\mathbf{P}}$. Simplification of (16) leads to

$$\mathbf{f}_i^{(n)} = (\mathbf{Q}^T)^{n-1}\mathbf{f}_i^{(1)} = (\mathbf{Q}^T)^{n-1}\mathbf{p}_i^T, \tag{17}$$

where $\mathbf{p}_i$ is ith row of the average TPM $\bar{\mathbf{P}}$ with the (i,i)th entry removed. Then, the average hitting time is given by

$$\tau_h = \sum_{n=1}^{\infty} n f_{ij}^{(n)}. \tag{18}$$

Setting $\phi^{(i)}$ to be the appropriate consensus state, i.e., either the all zeros state or the all ones state depending on the initial distribution of the data bits, we can compute the average hitting time using the above equation.

D. Average Consensus Duration

The average consensus duration, τ_c, is the average number of cycles for which the network stays in the same consensus state once it has reached consensus. Suppose the network is in consensus at the end of t_0th cycle. Let the random variable T_c represent the number of consecutive cycles for which the network stays in consensus. Assuming channel independence between any two cycles, the average probability (here, the averaging is over the channel states) that the network is in consensus for n consecutive cycles is simply given by

$$Pr\{T_c = n\} = \left(\prod_{k=1}^{n} \bar{P}_c\right)(1 - \bar{P}_c), \tag{19}$$

where $\bar{P}_c$ is the average probability of being in consensus after the next cycle once the network is already in consensus. The expected number of cycles for which the network stays in consensus is thus

$$\tau_c = \sum_{n=1}^{\infty} n(\bar{P}_c)^n(1 - \bar{P}_c) = \frac{\bar{P}_c}{1 - \bar{P}_c}. \tag{20}$$

The average probability $\bar{P}_c$ for an all-one consensus state is given by $\bar{P}_{MM} = (\bar{p}^{(M)})^N$ (see (13)). Similarly, for the all-zero consensus state, $\bar{P}_c = \bar{P}_{11} = (1 - \bar{p}^{(1)})^N$. The average probabilities $\bar{p}^{(1)}$ and $\bar{p}^{(M)}$ are derived in Appendix E. Thus, we can use the above equation to obtain the average consensus duration of the network. Note that, the average hitting time decreases and average consensus duration increases with increase in data SNR, $\text{SNR}_d \triangleq \frac{E_d\sigma^2}{\sigma_w^2/2}$, and pilot SNR, $\text{SNR}_p \triangleq \frac{E_p\sigma^2}{\sigma_w^2}$. The SNR values determine the entries of the TPM via the Q-function, which determines the average probabilities of the different states. The average probabilities, in turn, determine the average hitting time

and the average consensus duration via the analysis presented in the paper. Due to the intricate nature of the analysis, it is unfortunately difficult to directly relate the performance of the protocol to the data and pilot SNRs. Hence, we study the performance via simulations.

E. Data Power and Pilot Power Allocation

In a given consensus cycle, under a total power constraint, the optimization of the power allotted to the data and pilot transmission can be carried out, for example, by maximizing the average consensus duration, or, by minimizing the average hitting time. To do this, the average probability for the all-ones consensus state, $\bar{P}_{MM}$, equivalently, $\bar{p}^{(M)}$, can be used as the cost function. Therefore, the optimal data and pilot powers are obtained by solving

$$E_d^*, E_p^* = \arg \max_{E_d, E_p} \bar{p}^{(M)} \quad \text{subject to } E_d + E_p = E. \quad (21)$$

As shown in Appendix E, the average probability $\bar{p}^{(M)}$ can be obtained in closed form as

$$\bar{p}^{(M)} = 0.5 + \sqrt{\frac{\gamma_d}{2\pi}} \frac{\Gamma(m_1 + 0.5)}{\Gamma(m_1)\sqrt{m_1/\Omega_1}} \times {}_2F_1\left(0.5, m_1 + 0.5; 1.5; -\frac{\gamma_d}{2m_1/\Omega_1}\right), \quad (22)$$

where $\gamma_d = E_d/\sigma_v^2$, $\sigma_v^2 = \sigma_w^2(N-1)/2$ and ${}_2F_1$ is the Gaussian hypergeometric function. Note that the parameters of the Nakagami r.v. H_p, namely, $m_1 = (\mathbb{E}[H_p^2])^2/\mathrm{Var}[H_p^2]$ and $\Omega_1 = \mathbb{E}[H_p^2]$, are functions of the pilot power E_p. Substituting (22) in (21) leads to a one dimensional optimization problem, which can be solved numerically to obtain the optimal values of E_d and E_p. The importance of optimizing the power allocation to training and data symbols is illustrated in the next section.

V. Simulation Results

The simulation set-up consists of $N = 8$ sensor nodes. The receiver noise and the channels between the sensors are drawn i.i.d. from $\mathcal{CN}(0,1)$. The average TPM is evaluated by averaging over 20,000 channel instantiations.

We start with presenting results on the second largest eigenvalue, λ_2, of the average TPM of the state space of the Markov chain. In Fig. 2, the value of λ_2 obtained by the two schemes is compared, along with their approximation, $1 - 2\bar{p}^{(1)}$, presented in Section IV.B. It can be seen that λ_2 gets closer to 1 with increasing data and pilot SNRs. At lower pilot SNRs, the co-phased combining scheme outperforms the LMMSE-based scheme from [29], i.e., λ_2 of the former is closer to 1 than the latter. This is because, the co-phased combining depends only on the channel phase estimates, whereas the LMMSE-based scheme depends on both magnitude and phase estimates of the channel. This leads to larger errors in the sum of votes estimate for the LMMSE-based scheme. However, with increasing pilot power, the two schemes result in nearly the same value of λ_2. Also, we see that the approximate expression for λ_2 is a lower bound on the actual eigenvalue when $N = 8$, although it is accurate for the $N = 2$ and the $N = 3$ case. The approximation captures the relative behavior of the LMMSE and co-phasing

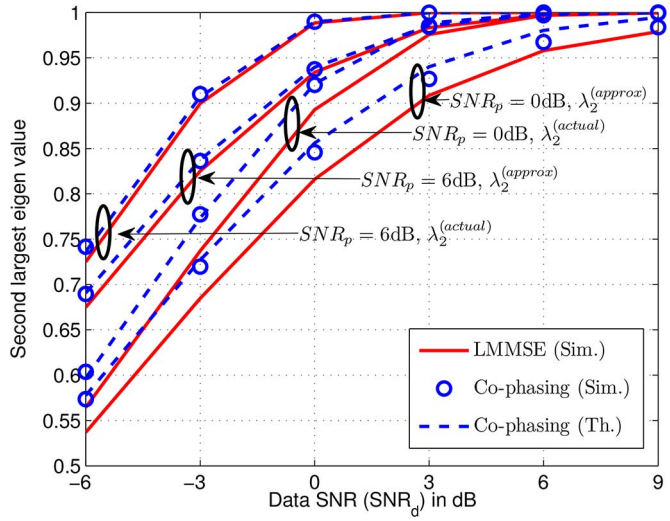

Fig. 2. Second largest eigenvalue vs. Data SNR (denoted by $\mathrm{SNR_d}$) for different pilot SNRs (denoted by $\mathrm{SNR_p}$), with $N = 8$ nodes.

schemes well. Finally, the excellent agreement between the theoretical and experimental values of λ_2 for the co-phased combining scheme is clear from the plot.

Next, we evaluate the average consensus duration performance of the two schemes, in Fig. 3. At low SNRs, the average consensus duration increases linearly, whereas at high SNRs, the increase in the average consensus duration is roughly quadratic. At low to intermediate pilot SNRs, the co-phased combining scheme stays longer in consensus state than the LMMSE-based scheme, as expected. The overall consensus performance depends both on the pilot SNR and data SNR, and therefore, even with accurate channel estimates (high pilot SNRs), the co-phased combining scheme outperforms the LMMSE scheme at low to moderate data SNRs. The LMMSE-based scheme starts outperforming the co-phased combining scheme only at high pilot SNR and high data SNR. Also, for the co-phased combining scheme, the average consensus duration obtained through simulations matches the theoretical results obtained through (20).

The average hitting time and the average probability of accurate consensus performance of the two schemes is plotted in Figs. 4 and 5, respectively. In Fig. 4, the average hitting time is evaluated by averaging the time to hit the all-ones consensus state over the initial states '11111110', '11111100', and '11111000'. In Fig. 5, the average probability of accurate consensus is plotted as a function of the number of cycles of the update procedure, when the initial state across the nodes is '00011111'. The conclusions from the figure are similar to that of the average consensus duration: the performance improves with increasing data and pilot SNRs; at low to intermediate pilot SNRs, the co-phased combining scheme outperforms the LMMSE-based scheme; and, for the co-phased combining scheme, the theoretical expressions in (18) match the simulation results. Also, the importance of accounting for the channel estimation errors in evaluating the relative performance of different consensus protocols is clear from the plot. At high pilot SNRs, the channel estimates are accurate, and the two schemes offer similar performance, while at intermediate or lower pilot

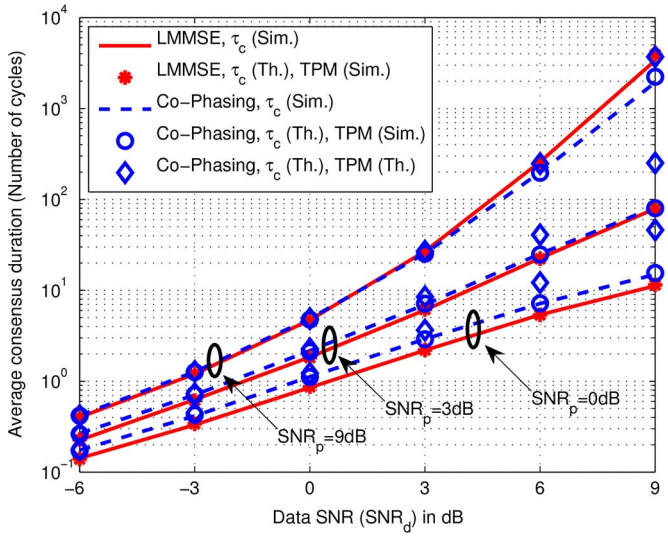

Fig. 3. Average consensus duration vs. Data SNR, with $N = 8$, and for different pilot SNRs. The curves labeled τ_c (Th.), TPM (Sim.) correspond to the theoretical value of the average consensus duration computed using the TPM obtained from the Monte Carlo averaging, while the curves labeled τ_c (Th.), TPM (Th.) correspond to the theoretical value of the average consensus duration computed using the theoretical TPM.

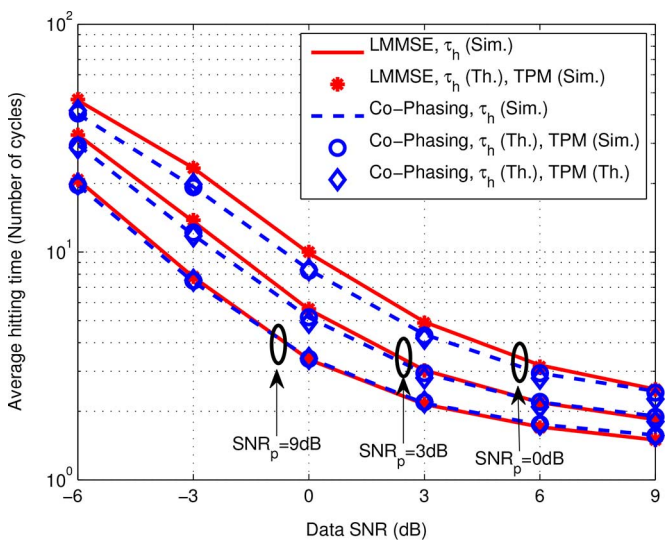

Fig. 4. Average hitting time Vs. Data SNR, with $N = 8$, and for different pilot SNRs.

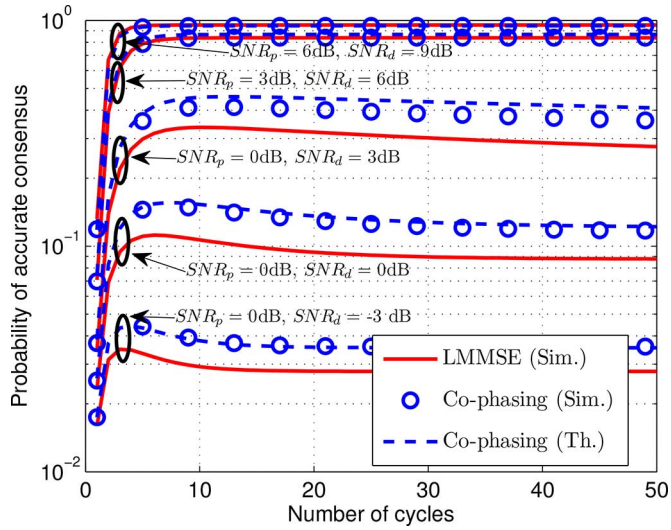

Fig. 5. Probability of accurate consensus vs. number of consensus cycles, starting from an initial state of '00011111'.

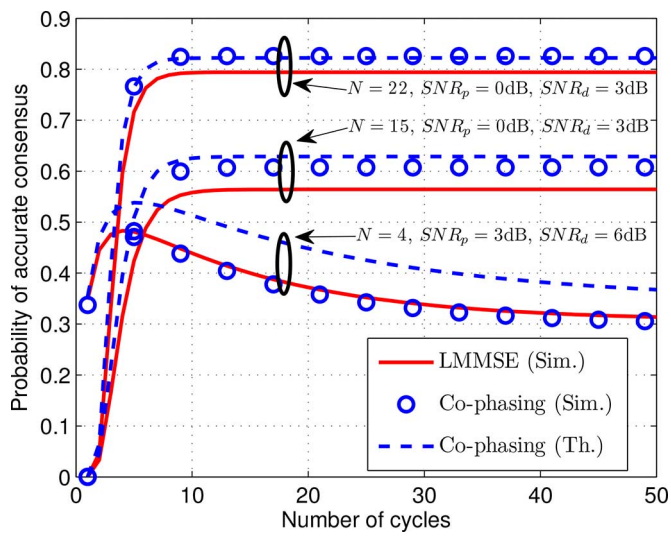

Fig. 6. Probability of accurate consensus vs. number of consensus cycles, for number of nodes, $N = 4$, 15, and 22. The initial state of the network is the simple majority state, for example, the initial state of the 4 nodes network is '0111'.

SNRs, the co-phasing scheme outperforms the LMMSE-based scheme.

Figure 6 shows the average probability of accurate consensus for various numbers of nodes, as a function of the number of consensus cycles. The probability of accurate consensus is much higher for $N = 22$ sensors with $\text{SNR}_p = 0$ dB and $\text{SNR}_d = 3$ dB compared to $N = 4$ sensors and a higher receive SNR, $\text{SNR}_p = 3$ dB and $\text{SNR}_d = 6$ dB. This is because of the linear scaling of the average effective SNR with the number of sensors. Also, for the co-phased combining scheme, with increase in the number of nodes, the theoretical curves match well with simulations.

In developing the consensus protocol, the self-bit was ignored for simplicity of presentation. Figure 7 shows the second largest eigenvalue for the co-phased combining scheme when the sensors' own data bit is considered. The sum of votes estimate is computed by adding the self-bit with a scaling of $\sqrt{E_d}$ to (6). Accounting for the self-bit clearly improves the performance, but the loss due to the approximation gets smaller as the number of sensors is increased. Also, as data SNR increases, the sum of votes estimate improves. This leads to the second largest eigenvalue without the self-bit to be close to that with self-bit. Hence, when the SNR and the number of nodes are moderately large, it is reasonable to ignore the self-bit in evaluating the consensus performance.

Next, we consider the allocation of the pilot and data power to optimize the consensus performance. In Figs. 8 and 9, the average consensus duration and average hitting time of the co-phased combining scheme are plotted as a function of the

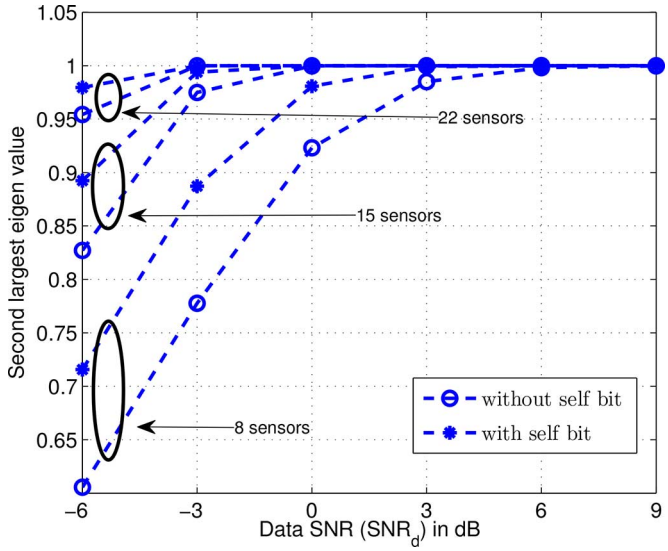

Fig. 7. Second largest eigenvalue vs. Data SNR for the co-phased combining scheme at a pilot SNR of 0 dB, when the sensors' self-bit is taken into account.

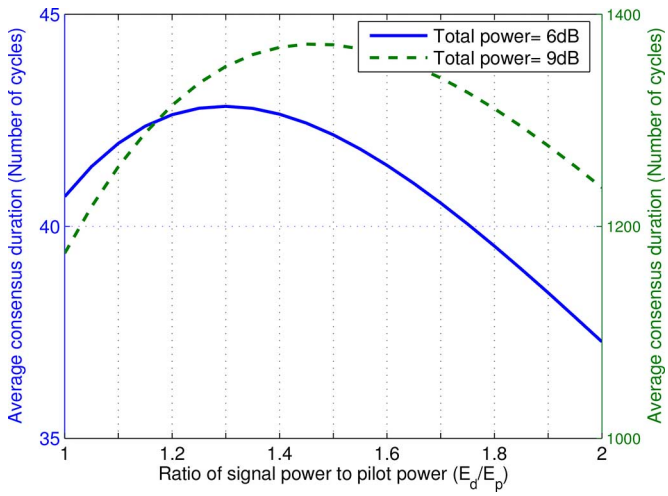

Fig. 8. Average consensus duration vs. ratio of data power to pilot power under a total power constraint.

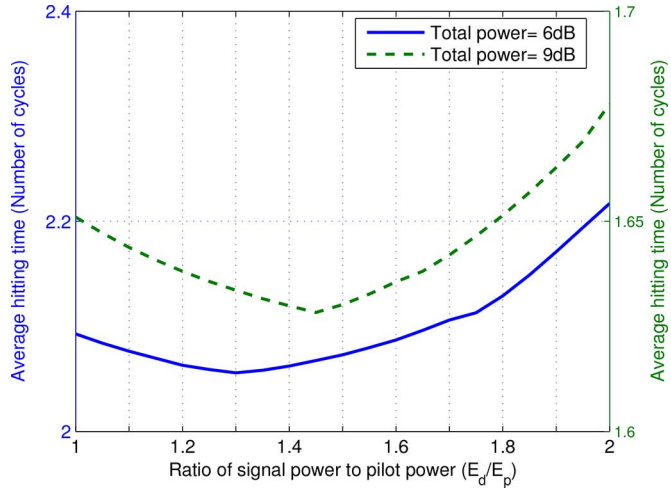

Fig. 9. Average hitting time vs. ratio of data power to pilot power under a total power constraint.

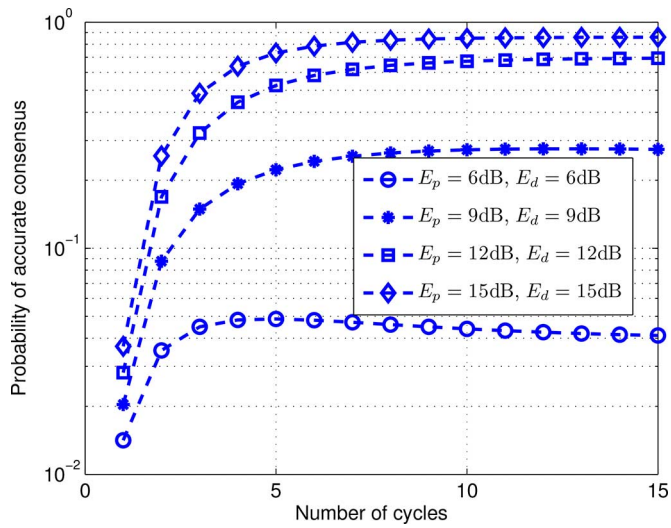

Fig. 10. Probability of accurate consensus vs. number of consensus cycles for the co-phased combining scheme when the path-loss between sensors is taken into account. The probability of accurate consensus is averaged over all the initial states corresponding to 5 sensors with bit '1' and 3 sensors with bit '0'.

ratio of data power to pilot power, under a total power constraint. The average consensus duration is the highest and the average hitting time is the lowest for $\frac{E_d}{E_p} = 1.3$ and $\frac{E_d}{E_p} = 1.45$, when the total power available is 6 dB and 9 dB, respectively. This indicates that, at higher total powers, the data symbol has to be given a larger share of the total power compared to the pilot symbol. At lower total powers, the best average consensus performance is obtained for nearly equal sharing of the available power. Similar conclusions are obtained from optimizing the other metrics such as the second largest eigenvalue or probability of accurate consensus; they are not presented here to avoid repetition.

Next, Fig. 10 shows the average probability of accurate consensus for the co-phased combining scheme when the path-loss (with path-loss exponent of 2) is also considered, along with the Rayleigh fading. A linear arrangement of 8 sensors that are uniformly spaced on a stretch of 10 m and a reference distance of 1 m is considered for the simulations. At higher transmit powers for pilot and data symbol, $E_p = 15$ dB and $E_d = 15$ dB, the probability of accurate consensus is close to 1. Also, the network attains its consensus state within about 5 cycles.

Finally, in Fig. 11, we show the probability of accurate consensus performance of the co-phased combining scheme when the network is not fully connected. While estimating the sum of votes in (7), the weights β_{kl} are set to 0 for all those links which have a link gain less than a threshold η, where η is determined for a given link failure probability, $\Pr\{|h_{kl}| \leq \eta\} = q$. At high data and pilot SNRs, the performance deteriorates as the link failure probability increases, as expected. At lower SNRs, as the channels are more noisy, the performance is relatively robust to link failures.

VI. Conclusion

In this paper, we considered a physical layer protocol for achieving majority-bit consensus, where a set of nodes exchange their current majority-bit estimates over multiple cycles.

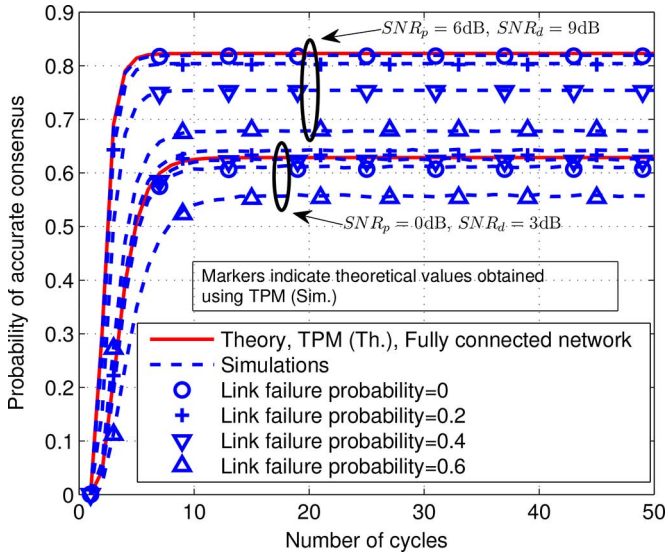

Fig. 11. Probability of accurate consensus vs. number of consensus cycles for the co-phased combining scheme for an $N = 15$ sensor network which is not fully connected.

We contrasted two bit-update schemes: 1) LMMSE-based scheme and 2) Co-phased combining scheme, when the available CSI is estimated using pilot symbols sent from the nodes. We analytically evaluated several metrics that determine the average consensus performance, such as the average probability of detecting the correct majority bit, the average hitting time, and the average consensus duration, by employing a difference-of-Nakagami approximation for a combined effective channel. The results highlighted the importance of accounting for the effect of channel estimation errors on the performance. For example, at low to intermediate SNRs, the co-phased combining scheme outperforms the LMMSE-based scheme, although they offer very comparable performance under perfect channel state information. This is because of the lesser dependence of the co-phased combining scheme on channel estimates. We showed that optimizing the power allocation between the pilot and data symbols can result in a significant improvement in the consensus performance. We also presented extensive simulation results to validate the theoretical expressions and illustrate the various tradeoffs involved. Future work could consider extension of this study to non-binary (e.g., average) consensus problems.

Appendix A
Proof of Lemma 1

Note that, $\mathbb{E}[H_p^2] = \mathbb{E}[(\sum_{s_k \in \mathcal{K}} |H_{kl}| \cos\tilde{\theta}_{kl})^2]$ has i.i.d. summand terms. Hence, it can be expressed as

$$\mathbb{E}\left[H_p^2\right] = K\mathbb{E}\left[G_{kl}^2\right] + K(K-1)(\mathbb{E}[G_{kl}])^2. \quad (23)$$

In [37], the first and second moments of G_{kl} are evaluated as

$$\mathbb{E}[G_{kl}] = \sqrt{\frac{\pi\sigma^2}{4}}\sqrt{\frac{\gamma_p\sigma^2}{1+\gamma_p\sigma^2}},$$
$$\mathbb{E}\left[G_{kl}^2\right] = \frac{\sigma^2(1+2\gamma_p\sigma^2)}{2+2\gamma_p\sigma^2}. \quad (24)$$

Substituting the above expressions in (23) leads to (9).

Next, the variance of H_p^2 is given by

$$\mathrm{Var}\left[H_p^2\right] = \mathbb{E}\left[H_p^4\right] - \left(\mathbb{E}\left[H_p^2\right]\right)^2. \quad (25)$$

As the terms G_{kl} are i.i.d., we have

$$\begin{aligned}\mathbb{E}\left[H_p^4\right] &= K\mathbb{E}\left[G_{kl}^4\right] + 3K(K-1)\left(\mathbb{E}\left[G_{kl}^2\right]\right)^2 \\ &\quad + K(K-1)(K-2)(K-3)(\mathbb{E}[G_{kl}])^4 \\ &\quad + 6K(K-1)(K-2)(\mathbb{E}[G_{kl}])^2\mathbb{E}\left[G_{kl}^2\right] \\ &\quad + 4K(K-1)\mathbb{E}\left[G_{kl}^3\right]\mathbb{E}[G_{kl}]. \quad (26)\end{aligned}$$

The third and fourth moment of G_{kl},

$$\mathbb{E}\left[G_{kl}^3\right] = \frac{3\sigma^4\sqrt{\gamma_p\pi}}{4\sqrt{1+\gamma_p\sigma^2}}, \quad \text{and} \quad (27)$$
$$\mathbb{E}\left[G_{kl}^4\right] = 2\sigma^4 - \frac{\sigma^4}{4}\frac{5+4\gamma_p\sigma^2}{(1+\gamma_p\sigma^2)^2} \quad (28)$$

are derived in [38], [39]. The expressions in (27), (28) along with (24) and (9) are used to obtain the expression for $\mathrm{Var}[H_p^2]$.

Appendix B
The Probability Density Function of the Difference of Two Nakagami Random Variables

The pdf of the difference of two independent Nakagami r.v.s is

$$f_H(h) = \int_0^\infty f_{H_p}(h+h_n)\, f_{H_n}(h_n)\, \mathrm{d}h_n, \quad \text{for } h \geq 0. \quad (29)$$

Substituting for the pdf of Nakagami r.v.s and then by completing the squares, (29) can be expressed as

$$\begin{aligned}f_H(h) &= \frac{4\left(\frac{m_1}{\Omega_1}\right)^{m_1}\left(\frac{m_2}{\Omega_2}\right)^{m_2}\mathrm{e}^{-\frac{h^2 m_1 m_2}{m}}}{\Gamma(m_1)\Gamma(m_2)} \\ &\quad \times \int_0^\infty (h+h_n)^{2m_1-1}(h_n)^{2m_2-1} \\ &\quad \times \mathrm{e}^{-\frac{m}{\Omega}\left(h_n+\frac{m_1\Omega_2 h}{m}\right)^2}\mathrm{d}h_n, \quad (30)\end{aligned}$$

where $m \triangleq m_1\Omega_2 + m_2\Omega_1$ and $\Omega \triangleq \Omega_1\Omega_2$. By changing the variable of integration to $u \triangleq \sqrt{\frac{m}{\Omega}}(h_n + \frac{m_1\Omega_2 h}{m})$, (30) is expressed as

$$\begin{aligned}f_H(h) &= \frac{4\left(\frac{m_1}{\Omega_1}\right)^{m_1}\left(\frac{m_2}{\Omega_2}\right)^{m_2}\mathrm{e}^{-\frac{h^2 m_1 m_2}{m}}}{\Gamma(m_1)\Gamma(m_2)\left(\frac{m}{\Omega}\right)^{m_1+m_2-\frac{1}{2}}} \\ &\quad \times \int_{\frac{m_1\Omega_2 h}{\sqrt{m\Omega}}}^\infty \left(u + \frac{m_2\Omega_1 h}{\sqrt{m\Omega}}\right)^{2m_1-1} \\ &\quad \times \left(u - \frac{m_1\Omega_2 h}{\sqrt{m\Omega}}\right)^{2m_2-1}\mathrm{e}^{-u^2}\mathrm{d}u. \quad (31)\end{aligned}$$

The parameters $2m_1$ and $2m_2$ are rounded-off to their corresponding nearest integers. Then, using the binomial expansion for power terms, the pdf of the effective channel can be simplified as given in (11).

APPENDIX C
DERIVATION OF $\bar{p}_l$ FOR THE CO-PHASED COMBINING SCHEME

The $\bar{p}_l$ in (8) can be expressed as

$$\bar{p}_l = \int_{-\infty}^{\infty} \int_{-\frac{h\sqrt{E_d}}{\sigma_v}}^{\infty} \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-v^2/2} f_H(h)\, \mathrm{d}v\, \mathrm{d}h. \quad (32)$$

By setting the variable of integration as $x \triangleq v\sigma_v/h\sqrt{E_d}$ and changing the order of integration, (32) can be expressed as

$$\bar{p}_l = \frac{\sqrt{E_d}}{\sqrt{2\pi}\sigma_v} \int_{-1}^{\infty} \underbrace{\int_{-\infty}^{\infty} h\, \mathrm{e}^{-h^2x^2E_d/2\sigma_v^2} f_H(h)\, \mathrm{d}h}_{\triangleq I_1}\, \mathrm{d}x. \quad (33)$$

By substituting (30) for $f_H(h)$, $h \geq 0$, and then using (6.455) from [40], the integral I_1 between the limits 0 and ∞ can be evaluated as

$$\begin{aligned} &\frac{2\left(\frac{m_1}{\Omega_1}\right)^{m_1}\left(\frac{m_2}{\Omega_2}\right)^{m_2}\left(\frac{\Omega}{m}\right)^{m_1+m_2-\frac{1}{2}}}{\Gamma(m_1)\Gamma(m_2)} \\ &\quad\times \sum_{i=0}^{2m_1-1}\sum_{j=0}^{2m_2-1} \binom{2m_1-1}{i}\binom{2m_2-1}{j} \\ &\quad\times \left(\frac{m_2\Omega_1}{\sqrt{m\Omega}}\right)^{2m_1-1-i} (-1)^{2m_2-1-j} \\ &\quad\times \frac{\left(\frac{m_1\Omega_2}{\sqrt{m\Omega}}\right)^{2m_2-i}\Gamma\left(m_1+m_2+\frac{1}{2}\right)}{(2m_1+2m_2-i-j)\left(\frac{x^2\gamma_d}{2}+\frac{m_1}{\Omega_1}\right)^{m_1+m_2+\frac{1}{2}}} \\ &\times {}_2F_1\left(1, m_1+m_2+\frac{1}{2}; m_1+m_2 \right. \\ &\qquad \left. -\frac{i+j-2}{2}; \frac{\frac{x^2\gamma_d}{2}+\frac{m_1m_2}{m}}{\frac{x^2\gamma_d}{2}+\frac{m_1}{\Omega_1}}\right), \end{aligned} \quad (34)$$

where ${}_2F_1$ is the Gauss hypergeometric function and $\gamma_d = E_d/\sigma_v^2$. Following the similar procedure, I_1 can be evaluated between the limits $-\infty$ and 0. The two expressions can now be substituted in (33) to evaluate $\bar{p}_l$.

APPENDIX D
SECOND LARGEST EIGENVALUE OF THE AVERAGE TPM FOR $N = 3$ SENSORS

It can be shown that the second largest eigenvalue of the average TPM is $1 - 2\bar{p}^{(1)}$ when $N = 3$ sensors are involved in achieving consensus. Let, p_0, p_1 and p_2 denote the probability of detecting bit '1' at a sensor when two other sensors transmit bits '(0, 0)', '(0, 1)', and '(1, 1)', respectively. Then, by combining equivalent states (for e.g., state '001' is equivalent to '100' and '010'), the average TPM is given by

$$\bar{\mathbf{P}} = \begin{bmatrix} p_2^3 & 3p_0p_2^2 & 3p_0^2p_2 & p_0^3 \\ p_1^2p_2 & p_1^2p_0+2p_1^2p_2 & p_1^2p_2+2p_1^2p_0 & p_1^2p_0 \\ p_1^2p_0 & p_1^2p_2+2p_1^2p_0 & p_1^2p_0+2p_1^2p_2 & p_1^2p_2 \\ p_0^3 & 3p_0^2p_2 & 3p_0p_2^2 & p_2^3 \end{bmatrix}. \quad (35)$$

The average TPM, $\bar{\mathbf{P}}$ is a centro-symmetric matrix [41], i.e.,

$$\bar{\mathbf{P}} = \begin{bmatrix} A & B \\ JBJ & JAJ \end{bmatrix}, \quad (36)$$

where J is a counter-identity matrix, and

$$A = \begin{bmatrix} p_2^3 & 3p_0p_2^2 \\ p_1^2p_2 & p_1^2p_0+2p_1^2p_2 \end{bmatrix}, \quad \text{and}$$
$$B = \begin{bmatrix} 3p_0^2p_2 & p_0^3 \\ p_1^2p_2+2p_1^2p_0 & p_1^2p_0 \end{bmatrix}. \quad (37)$$

For such an average TPM, it is known that $\bar{\mathbf{P}}$ is similar to

$$C = \begin{bmatrix} A-BJ & 0 \\ 0 & A+BJ \end{bmatrix}. \quad (38)$$

Thus, the eigenvalues of $\bar{\mathbf{P}}$ are same as that of $A-BJ$ and $A+BJ$. The eigenvalues of $A-BJ$ are found to be 1 and $\frac{3}{4}-3p_0p_2$ and eigenvalues of $A+BJ$ are $1-2p_0$ and $(1-2p_0)(\frac{1}{4}-p_0p_2)$. The second largest eigenvalue is $1-2p_0$, which is the same as $1-2\bar{p}^{(1)}$.

APPENDIX E
DERIVATION OF $\bar{p}^{(M)}$ AND $\bar{p}^{(1)}$ FOR THE CO-PHASED COMBINING SCHEME

When the bits b_1 to b_N are $+1$, $|\mathcal{K}| \triangleq |\mathcal{S}\setminus\{s_l\}| = N-1$ and the effective channel H can be approximated by a Nakagami r.v. with parameters m_1 and Ω_1. So, the average probability $\bar{p}^{(M)}$ in (8) can be expressed as

$$\bar{p}^{(M)} = \int_0^{\infty}\int_{-\frac{h\sqrt{E_d}}{\sigma_v}}^{\infty} \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-v^2/2}\frac{2}{\Gamma(m_1)} \times \left(\frac{m_1}{\Omega_1}\right)^{m_1} h^{2m_1-1}e^{-m_1h^2/\Omega}\, \mathrm{d}v\, \mathrm{d}h. \quad (39)$$

By setting the variable of integration as $x \triangleq v\sigma_v/h\sqrt{E_d}$ and changing the order of integration, (39) can be expressed as

$$\bar{p}^{(M)} = \sqrt{\frac{2\gamma_d}{\pi}}\frac{1}{\Gamma(m_1)}\left(\frac{m_1}{\Omega_1}\right)^{m_1} \times \int_{-1}^{\infty} \underbrace{\int_0^{\infty} h^{2m_1}\, \mathrm{e}^{-h^2\left(\frac{x^2\gamma_d}{2}+\frac{m_1}{\Omega_1}\right)}\, \mathrm{d}h}_{\triangleq I_2}\, \mathrm{d}x. \quad (40)$$

The integral I_2 can be solved using (3.326) from [40] and substituted in (40) to obtain

$$\bar{p}^{(M)} = \sqrt{\frac{\gamma_d}{2\pi}}\frac{1}{\Gamma(m_1)}\left(\frac{m_1}{\Omega_1}\right)^{m_1} \times \int_{-1}^{\infty} \frac{\Gamma(m_1+0.5)}{\left(\frac{x^2\gamma_d}{2}+\frac{m_1}{\Omega_1}\right)^{m_1+0.5}}\, \mathrm{d}x. \quad (41)$$

Further, the integral in (41) can be expressed as

$$\bar{p}^{(M)} = \sqrt{\frac{\gamma_d}{2\pi}}\frac{\Gamma(m_1+0.5)}{\Gamma(m_1)}\left(\frac{m_1}{\Omega_1}\right)^{m_1} \left[\frac{x}{(m_1/\Omega_1)^{m_1+0.5}} \times {}_2F_1\left(0.5, m_1+0.5; 1.5; -\frac{\gamma_d x^2}{2m_1/\Omega_1}\right)\right]_{-1}^{\infty}. \quad (42)$$

This leads to

$$\bar{p}^{(M)} = 0.5 + \sqrt{\frac{\gamma_d}{2\pi}} \frac{\Gamma(m_1+0.5)}{\Gamma(m_1)\sqrt{m_1/\Omega_1}} \times {}_2F_1\left(0.5, m_1+0.5; 1.5; -\frac{\gamma_d}{2m_1/\Omega_1}\right). \quad (43)$$

When the bits b_1 to b_N are 0, $|\mathcal{K}^c| \triangleq |\mathcal{S}\setminus\{\mathcal{K}, s_l\}| = N-1$ and the effective channel H can be approximated by the negative of a Nakagami r.v. with parameters m_2 and Ω_2. The average probability $\bar{p}^{(1)}$ is given by

$$\bar{p}^{(1)} = 1 - \bar{p}^{(M)}. \quad (44)$$

Therefore, replacing the parameters m_1 and Ω_1 by m_2 and Ω_2, respectively, $\bar{p}^{(1)}$ can be expressed as

$$\bar{p}^{(1)} = 0.5 - \sqrt{\frac{\gamma_d}{2\pi}} \frac{\Gamma(m_2+0.5)}{\Gamma(m_2)\sqrt{m_2/\Omega_2}} \times {}_2F_1\left(0.5, m_2+0.5; 1.5; -\frac{\gamma_d}{2m_2/\Omega_2}\right). \quad (45)$$

Appendix F
Proof That the Average Probability $\bar{P}_{ij}$ Is Strictly Positive

The probability that the system changes from ϕ_i to ϕ_i at an arbitrary time t is given by

$$\bar{P}_{ij} = \prod_{l=1}^{N} \left[b_l^{(i)} \bar{p}_l^{(j)} + \left(1 - b_l^{(i)}\right)\left(1 - \bar{p}_l^{(j)}\right)\right], \quad (46)$$

i.e., it is the product of the N probability terms. We will consider the least of these probabilities, the probability that a node s_l detects bit '1' in the current cycle when all other nodes have bit '0' in the previous cycle. This probability, $\bar{p}^{(1)}$, is derived in (45), and it can be verified that, the second term of (45) is strictly smaller than 0.5. Thus, the least of the probabilities, $\bar{p}^{(1)}$ is strictly greater than zero, and therefore, $\bar{P}_{ij} > 0$.

References

[1] V. Y. Ramakrishnaiah and C. R. Murthy, "Physical layer binary consensus over fading wireless channels and with imperfect CSI," in *Proc. IEEE Global Commun. Conf.*, Dec. 2014, pp. 3383–3388.

[2] W. Zhang, R. Mallik, and K. Letaief, "Optimization of cooperative spectrum sensing with energy detection in cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 8, no. 12, pp. 5761–5766, Dec. 2009.

[3] S. Ashrafi, M. Malmirchegini, and Y. Mostofi, "Binary consensus for cooperative spectrum sensing in cognitive radio networks," in *Proc. IEEE Global Commun. Conf.*, Dec. 2011, pp. 1–6.

[4] S. Maleki, S. P. Chepuri, and G. Leus, "Optimization of hard fusion based spectrum sensing for energy-constrained cognitive radio networks," *Phys. Commun.*, vol. 9, pp. 193–198, Dec. 2013.

[5] P. Grover, A. Goldsmith, and A. Sahai, "Fundamental limits on the power consumption of encoding and decoding," in *Proc. IEEE Int. Symp. Inf. Theory*, Jul. 2012, pp. 2716–2720.

[6] J. N. Tsitsiklis, "Problems in decentralized decision making and computation," Ph.D. dissertation, Electr. Eng. Comput. Sci. Dept., Massachusetts Inst. of Technol., Cambridge, MA, USA, Nov. 1984.

[7] D. Kempe, A. Dobra, and J. Gehrke, "Gossip-based computation of aggregate information," in *Proc. IEEE Symp. Found. Comput. Sci.*, Oct. 2003, pp. 482–491.

[8] L. Xiao and S. Boyd, "Fast linear iterations for distributed averaging," *Syst. Control Lett.*, vol. 53, no. 1, pp. 65–78, Sep. 2004.

[9] R. Olfati-Saber and R. M. Murray, "Consensus problems in networks of agents with switching topology and time-delays," *IEEE Trans. Autom. Control*, vol. 49, no. 9, pp. 1520–1533, Sep. 2004.

[10] V. D. Blondel, J. M. Hendrickx, A. Olshevsky, and J. N. Tsitsiklis, "Convergence in multiagent coordination, consensus, and flocking," in *Proc. IEEE Conf. Decision Control*, Dec. 2005, pp. 2996–3000.

[11] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah, "Gossip algorithms: Design, analysis and applications," in *Proc. IEEE Int. Conf. Comput. Commun.*, Mar. 2005, vol. 3, pp. 1653–1664.

[12] 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.

[13] B. Nazer, A. Dimakis, and M. Gastpar, "Local interference can accelerate gossip algorithms," *IEEE J. Sel. Topics Signal Process.*, vol. 5, no. 4, pp. 876–887, Aug. 2011.

[14] A. Kashyap, T. Başar, and R. Srikant, "Quantized consensus," *Automatica*, vol. 43, no. 7, pp. 1192–1203, July 2007.

[15] T. C. Aysal, M. Coates, and M. Rabbat, "Distributed average consensus using probabilistic quantization," in *Proc. IEEE Workshop Statist. Signal Process.*, Aug. 2007, pp. 640–644.

[16] S. Kar and J. M. Moura, "Distributed consensus algorithms in sensor networks: Quantized data and random link failures," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1383–1400, Mar. 2010.

[17] F. Bénézit, P. Thiran, and M. Vetterli, "Interval consensus: From quantized gossip to voting," in *Proc. IEEE Int. Conf. Acoust. Speech., Signal Process.*, Apr. 2009, pp. 3661–3664.

[18] M. Draief and M. Vojnovic, "Convergence speed of binary interval consensus," in *Proc. IEEE Int. Conf. Comput. Commun.*, Mar. 2010, pp. 1–9.

[19] F. Bénézit, P. Thiran, and M. Vetterli, "The distributed multiple voting problem," *IEEE J. Sel. Topics Signal Process.*, vol. 5, no. 4, pp. 791–804, Aug. 2011.

[20] Y. Wang and P. M. Djuric, "Reaching consensus on a binary state by exchanging binary actions," in *Proc. IEEE Int. Conf. Acoust. Speech, Signal Process.*, Mar. 2012, pp. 3297–3300.

[21] Y. Wang and P. M. Djuric, "A gossip method for optimal consensus on a binary state from binary actions," *IEEE J. Sel. Topics Signal Process.*, vol. 7, no. 2, pp. 274–283, Apr. 2013.

[22] A. Abdaoui and T. Elfouly, "Distributed binary consensus algorithm in wireless sensor networks with faulty nodes," in *Proc. IEEE GCC Conf. Exhib.*, Nov. 2013, pp. 495–500.

[23] J. George and A. Swami, "Binary consensus through binary communication," in *Proc. IEEE Conf. Decision Control*, Dec. 2014, pp. 721–726.

[24] R. Olfati-Saber, E. Franco, E. Frazzoli, and J. S. Shamma, "Belief consensus and distributed hypothesis testing in sensor networks," in *Proc. Networked Embedded, Sens. Control*. New York, NY, USA: Springer, Jul. 2006, vol. 331, pp. 169–182.

[25] Z. Li, F. R. Yu, and M. Huang, "A distributed consensus-based cooperative spectrum-sensing scheme in cognitive radios," *IEEE Trans. Veh. Technol.*, vol. 59, no. 1, pp. 383–393, Jan. 2010.

[26] K. Chan, A. Swami, Q. Zhao, and A. Scaglione, "Consensus algorithms over fading channels," in *Proc. IEEE Military Commun. Conf.*, Nov. 2010, pp. 549–554.

[27] Y. Mostofi, "Binary consensus with Gaussian communication noise: A probabilistic approach," in *Proc. IEEE Conf. Decision Control*, Dec. 2007, pp. 2528–2533.

[28] Y. Ruan and Y. Mostofi, "Binary consensus with soft information processing in cooperative networks," in *Proc. IEEE Conf. Decision Control*, Dec. 2008, pp. 3613–3619.

[29] Y. Mostofi and M. Malmirchegini, "Binary consensus over fading channels," *IEEE Trans. Signal Process.*, vol. 58, no. 12, pp. 6340–6354, Dec. 2010.

[30] B. Chen, R. Jiang, T. Kasetkasem, and P. Varshney, "Channel aware decision fusion in wireless sensor networks," *IEEE Trans. Signal Process.*, vol. 52, no. 12, pp. 3454–3458, Dec. 2004.

[31] Y. Zou, J. Zhu, B. Zheng, and Y.-D. Yao, "An adaptive cooperation diversity scheme with best-relay selection in cognitive radio networks," *IEEE Trans. Signal Process.*, vol. 58, no. 10, pp. 5438–5445, Oct. 2010.

[32] Y. Zou, Y.-D. Yao, and B. Zheng, "Cognitive transmissions with multiple relays in cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 10, no. 2, pp. 648–659, Feb. 2011.

[33] R. Cogburn, "The ergodic theory of markov chains in random environments," *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, vol. 66, no. 1, pp. 109–128, May 1984.

[34] S. M. Ross, *Introduction to Probability Models*. New York, NY, USA: Academic, 2006.
[35] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1990.
[36] D. Aldous and J. A. Fill, *Reversible Markov Chains and Random Walks on Graphs,* 2002 [Online]. Available: http://www.stat.berkeley.edu/~aldous/RWG/book.html, unfinished monograph, recompiled 2014.
[37] K. V. K. Chaythanya, R. Annavajjala, and C. R. Murthy, "Comparative analysis of pilot-assisted distributed cophasing approaches in wireless sensor networks," *IEEE Trans. Signal Process.*, vol. 59, no. 8, pp. 3722–3737, Aug. 2011.
[38] A. Manesh, C. R. Murthy, and R. Annavajjala, "Design and analysis of distributed co-phasing with arbitrary constellations," in *Proc. IEEE Int. Conf. Commun.*, Jun. 2013, pp. 5780–5785.
[39] A. Manesh, C. R. Murthy, and R. Annavajjala, "Physical layer data fusion via distributed co-phasing with general signal constellations," *IEEE Trans. Signal Process.*, vol. 63, no. 17, pp. 4660–4672, Sept. 2015.
[40] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*. New York, NY, USA: Academic, 2007.
[41] J. R. Weaver, "Centrosymmetric (cross-symmetric) matrices, their basic properties, eigenvalues, and eigenvectors," *Amer. Math. Month.*, vol. 92, no. 10, pp. 711–717, Dec. 1985.

Venugopalakrishna Y. Ramakrishnaiah received B.E. in ECE from UVCE, Bangalore, in the year 2001. He obtained the M.S. (research) from the E.E. Department of IIT Madras, Chennai, in 2009. Currently, he is pursuing the Ph.D. degree at the ECE Department of IISc, Bangalore. His research focus is on data fusion based physical layer protocols for cognitive radio applications. From April 2002 to November 2004, he worked for CMC Limited, Bangalore as an IT engineer. During 2005–2008, in conjunction with his M.S., he worked as a project associate in a DST funded project at IIT Madras, where he developed a syllable-based speech synthesizer for Indian languages.

Chandra R. Murthy (S'03–M'06–SM'11) received the B.Tech. degree in electrical engineering from the Indian Institute of Technology, Madras, in 1998, the M.S. and Ph.D. degrees in electrical and computer engineering from Purdue University and the University of California, San Diego, in 2000 and 2006, respectively. From 2000 to 2002, he worked as an engineer for Qualcomm Inc., where he worked on WCDMA baseband transceiver design and 802.11b baseband receivers. From August 2006 to August 2007, he worked as a staff engineer at Beceem Communications Inc. on advanced receiver architectures for the 802.16e Mobile WiMAX standard. In September 2007, he joined the Department of Electrical Communication Engineering at the Indian Institute of Science, Bangalore, India, where he is currently working as an Associate Professor.

His research interests are in the areas of Cognitive Radio, Energy Harvesting Wireless Sensors, MIMO systems with channel-state feedback, and sparse signal recovery techniques applied to wireless communications. His paper won the best paper award in the Communications Track in the National Conference on Communications 2014. He is currently serving as the Chair of the IEEE Signal Processing Society, Bangalore Chapter, and as an associate editor for the IEEE SIGNAL PROCESSING LETTERS. He is an elected member of the IEEE SPCOM Technical Committee for the years 2014–2016.