Distributed Cooperative Spectrum Sensing Using Weighted Average Consensus

Wenlin Zhang*, Zheng Wang, Hongbo Liu, Yi Guo, Yingying Chen, Joseph Mitola III

Abstract—In this paper, we study the distributed spectrum sensing in cognitive radio networks. Existing distributed consensus-based fusion algorithms only ensure equal gain combining of local measurements, whose performance may be incomparable to various centralized soft combining schemes. Motivated by this fact, we consider practical channel conditions and link failures, and develop new weighted soft measurement combining without a centralized fusion center. Following the measurement by its energy detector, each secondary user exchanges its own measurement statistics with its local one-hop neighbors, and chooses the information exchanging rate according to the measurement channel condition, e.g., the signal-to-noise ratio (SNR). We rigorously prove the convergence of the new consensus algorithm, and show all SUs hold the same global decision statistics from the weighted soft measurement combining throughout the network. We also provide distributed optimal weight design under uncorrelated measurement channels. The convergence rate of the consensus iteration is given under the assumption that each communication link has an independent probability to fail, and the upper bound of the iteration number of the $\epsilon$-convergence is explicitly given as a function of system parameters. Simulation results show significant improvement of the sensing performance compared to existing consensus-based approaches, and the performance of the distributed weighted design is comparable to the centralized weighted combining scheme.

I. INTRODUCTION

Cognitive Radio (CR) [1] aims to improve the spectrum utilization by allowing unlicensed secondary user (SU) to operate in the ‘white spaces’ of the licensed spectrum bands without interfering the licensed primary user (PU). Revealing a future communication paradigm with dramatically enhanced spectrum efficiency, cognitive radio network is also referred as the neXt Generation (XG) or dynamic Spectrum Access (DSA) network [2].

One of the fundamental techniques in cognitive radio is spectrum sensing, which enables the secondary users to detect the presence of a primary user in the spectrum, see [3][4] and the references therein. The main challenge of spectrum sensing is the receiver uncertainty problem [1] such as practical multipath fading and shadowing, which compromise the detecting performance significantly. Recent research progress shows cooperative spectrum sensing [5] is a promising methodology to improve the spectrum sensing performance under shadowing, fading and time-varying wireless channels. In this paper, we study the weighted average consensus-based cooperative spectrum sensing using one-hop local communication to achieve distributed weighted soft combining without a centralized fusion center.

A. Related Work in Cooperative Spectrum Sensing

The main advantage of cooperative spectrum sensing is to enhance the sensing performance by exploiting the observation diversity of spatially located SUs [5]. By cooperation, CR users can share their sensing information to make a combined decision which is more accurate than individual decisions. Cooperative sensing usually contains two stages: sensing and fusion. In the sensing stage, each SU makes the measurement using appropriate detecting techniques. Among all types of detectors, energy detector is widely applied because it requires lower design complexity and no priori knowledge of primary users, compared to other techniques such as matched filter detection or cyclostationary detection [6]. In the fusion stage, the SU network cooperatively combines the detecting statistics throughout the network and the final decision is made using global information. Among the fusion techniques, different measurement combining methods have been considered including hard bit combining [7], soft gain combining [8], to name a few.

The key element of cooperative sensing is the cooperation scheme, which decides the SU network structure and the detecting performance. Centralized cooperative sensing and relay-assisted cooperative sensing are two major schemes in literature [5]. Centralized cooperative sensing [9] lets all SUs report their measurement information to a centralized fusion center, then a global decision is made at the fusion center according to certain measurement combining methods. Relay-assisted cooperative sensing [10] [5] is a multi-hop cooperation scheme which makes use of the strong sensing channels and strong reporting channels among the SU network in order to improve the overall performance. Relay-assisted sensing can be either centralized with a fusion center, or distributed without a fusion center. Centralized cooperative spectrum sensing requires the entire received data be gathered at one place which may be difficult due to communication constraints [11]. The multi-hop communication of the relay-assisted sensing may bring extra power cost and the degradation of sensing data quality during the multi-hop communication paths.

Distributed cooperative sensing first appears in [7] with broadcasting schemes. After measurement, each SU broadcasts its own decision to all SU nodes in the network, and the
final decision is decided by OR rule. In [12], an incremental gossip scheme is proposed for the distributed cooperative spectrum sensing. At each time step, each SU randomly selects neighboring nodes to communicate measurement statistics, and the final convergence value depends on the neighbor selection and the communication protocol. Very recently, bio-inspired consensus scheme is introduced to spectrum sensing in [13][14] for distributed measurement fusion and soft combining. Consensus-based spectrum sensing is a biologically inspired approach learned from swarming behaviors of fish schools and bird flocks. The consensus-based cooperation features self-organizable and scalable network structure and only needs one-hop communication among local neighbors.

The future cognitive radio networks will most probably consist of smart phones, tablets and laptops moving with the swarming behaviors of people. Therefore, consensus-based spectrum sensing reveals great potential for future development of distributed cognitive radio networks. However, the existing consensus-based fusion algorithms [13][14] only ensure equal gain combining of local measurements, which is incomparable with centralized weighted combining approaches [8]. To make the distributed consensus-based spectrum sensing more robust to practical channel conditions and link failures, we need to develop new distributed weighted fusion algorithms which are missing in the current literature.

B. Related Work in Average Consensus Algorithm

The consensus algorithm was first proposed in [15] for modeling decentralized decision making and parallel computing. The main benefit of consensus is ensuring each node to hold the global average of the initial values throughout the network using local communication between one-hop neighboring nodes. Two decades later, consensus algorithm is introduced to multi-agent systems [16][17]. In [16], Jadabaie et al. analyze the convergence conditions of a biologically-rooted discrete time consensus model, but the convergence value is not specified. Olfati-Saber and Murray give the conditions for average consensus convergence of continuous time consensus model in [17]. Since the average consensus problem has strong impact on distributed networked systems, it increasingly attracts research attention on decentralized estimation [18], filtering [19], and detection [20], etc.. For signal processing applications, communication constraints and the convergence rate become crucial for performance improvement. Typical problems include communication topology design and optimization [21], convergence rate analysis and optimization [22]. Interested readers are referred to the review papers [23][24] and the references therein.

Compared to the extensively studied average consensus, much less research attention is paid to weighted average consensus. As stated in [23], weighted average consensus algorithm is modeled by asymmetric matrices which makes the mathematical tools for average consensus algorithm inapplicable, and it is difficult to predict the convergence value on dynamic communication channels. However, weighted average consensus algorithm in the fusion process of spectrum sensing can achieve weighted gain combining without a fusion center, which advances the consensus-based spectrum sensing significantly. Therefore, it is important to develop solid theoretical analysis of weighted average consensus algorithms on dynamic communication topologies.

C. Contribution

In this paper, we propose a distributed cooperative spectrum sensing scheme based on weighted average consensus algorithm. After the measurement stage by the energy detector, SUs exchange their decision statistics with their local neighbors, and weight the information exchanging rate according to their own measurement channel conditions. We rigorously prove that after the consensus is reached, each SU holds the global decision statistics from the weighted soft combining throughout the network. The proposed scheme is fully distributed since the combining process only involves local one-hop communication without a centralized fusion center. The weighted combining makes the final decision statistic adaptive to the measurement channel conditions and is robust to communication channel failures. We also provide distributed optimal weight design under uncorrelated measurement channels, which offers comparable detecting performance compared to the centralized optimal weight design.

The main contribution of this paper has two folds. First, we provide formal convergence analysis of the weighted average consensus under fixed and dynamic communication channels, which advances the theoretical development of consensus algorithms and encompasses average consensus as a special case. In particular, we rigorously prove that temporary communication link failures do not affect the convergence of the weighted average consensus under the jointly connected condition. Second, to the best of our knowledge, we are the first to propose distributed weighted soft combining method in cooperative spectrum sensing. Based on preliminary results presented in our early conference paper [25], we provide a formal treatment of the distributed sensing algorithm in this paper. We obtain closed-form optimal weight design in the distributed weighted combining scheme for the generic additive Gaussian channel approximation, and estimate the convergence rate of the consensus iteration under the assumption that each communication link has an independent probability to fail. We characterize the upper bound of the iteration number of the $\epsilon$-convergence, which indicates all SUs are $\epsilon$ close to the final convergence value in the probability sense. Simulation results show significant improvement of the sensing performance compared to existing consensus-based approaches, and the performance of the distributed weighted design is comparable to centralized weight combining schemes.

The rest of this paper is organized as follows: In section II, we introduce the notations used in this paper, the consensus related SU network model, and fundamentals about the energy detector. In Section III, we present the main results on weighted average consensus-based spectrum sensing, where we prove the convergence of the proposed weighted average consensus algorithm under fixed and dynamic communication channels, estimate the convergence rate, and derive
optimal weight design. In section V, we discuss simulation results and make comparison with existing approaches. We conclude this study in Section VI.

II. SECONDARY USER NETWORK MODELING AND ENERGY DETECTION

In this section, we introduce the fundamentals on consensus-based cooperative spectrum sensing, including notations, consensus modeling of the SU network and the basics of energy detection.

A. Notations and Symbols

complementary cumulative distribution function of a zero mean unit variance Gaussian variable, i.e., \( Q(x) = \int_x^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2} dt \);

\( N(\mu, \sigma^2) \) Gaussian distribution with mean \( \mu \) and variance \( \sigma^2 \);

\( E(\cdot) \) Expectation of a random variable;

\( \text{Var}(\cdot) \) Variance of a random variable;

\( \| \cdot \|_2 \) \( L_2 \)-norm of a vector;

\( \| \cdot \|_\infty \) \( L_{\infty} \)-norm of a vector;

\( \text{diag}(\cdot) \) Square diagonal matrix with the elements of a given vector on the diagonal;

\( \text{Pr}(\cdot) \) Probability of a specified random variable;

\( | \cdot | \) Denote the number of elements in a finite discrete set or the absolute value of a complex number;

\( I \) denote the \( n \) dimensional identity matrix;

\( \text{Span}(\cdot) \) denote the space spanned a vector;

\( N(\cdot) \) Denote the null space of a matrix;

\( R(\cdot) \) Denote the rank space of a matrix;

B. The SU Network Model and Consensus Notions

In the information fusion stage, SUs communicate with their local neighbors through the SU network and adopt the consensus iteration to obtain the global measurement statistics. For convenience, we assign an index set \( I = \{1, 2, \ldots, n\} \) for the SU network formed by \( n \) SUs.

To model the consensus algorithm, we adopt the standard undirected graph model for the bidirectional SU communication network. The SU network is represented by an undirected graph \( G = (E, V) \), where \( V = \{v_i | i \in I\} \) is a finite nonempty set of nodes. We refer the \( i^{\text{th}} \) node as the \( i^{\text{th}} \) SU. The two names, SU and node, will be used alternatively. The edge set \( E = \{e_{ij} = (v_i, v_j) | i, j \in I\} \). The set of neighbors of node \( i \) is denoted by \( N_i = \{j : e_{ij} \in E\} \). A path in \( G \) consists of a sequence of nodes \( v_1, v_2, \ldots, v_l, l \geq 2 \), satisfying \((e_{m,m+1}) \in E, 1 \leq m \leq l - 1 \). The graph \( G \) is connected if any two distinct nodes in \( G \) are connected by a path. When considering the directed graph (i.e. digraph), we refer to \( v_i \) and \( v_j \), as the tail and head of a directed edge \( e_{ij} = (v_i, v_j) \), which represents the unidirectional communication link between two neighboring SUs. A digraph is called strongly connected if it is possible to reach any node starting from any other node following the edge directions.

In the case of the time-varying communication links, we model the SU network by \( G(k) = (E(k), V) \), where \( E(k) \) is the set of active edges at time \( k \). Let \( N_i(k) = \{j \in V | (i, j \in E(k))\} \), and \( d_i(k) = |N_i(k)| \) denote the degree (number of neighbors) of node \( i \) at time \( k \).

Let \( G_i = (E_i, V), i = 1, \ldots, r, \) denote a finite collection of graphs with common vertex set \( V \). Their union is a graph \( G \) with the same vertex set and an edge set that is the union of the \( E_i \)'s, i.e., \( G = \bigcup_{i=1}^{r} G_i = \bigcup_{i=1}^{r} E_i, V \). The set of undirected graphs \( \{G_1, \ldots, G_r\} \) is called jointly connected if their union is a connected graph.

In consensus network modeling, the Laplacian matrix \( L \in \mathbb{R}^{n \times n} \) of the communication graph \( G \) formed by the secondary user nodes is defined as

\[
L_{ij} = \begin{cases} 
0, & \text{if } i = j, \\
d_i, & \text{if } i \neq j, \\
-1, & j \in N_i,
\end{cases}
\]

where \( d_i = |N_i| \) is the degree of node \( i \). The maximum node degree is denoted as

\[
d_{\text{max}} = \max_i |N_i|.
\]

It’s easy to see, the undirected graph Laplacian matrix \( L \) is symmetric and has the left and right eigenvector \( \mathbf{1} \) and \( \mathbf{1} \) associated with the eigenvalue 0, respectively.

In the context of consensus-based spectrum sensing, for the \( n \) SUs modeled by the graph \( G \), the \( i^{\text{th}} \) SU is assigned a state variable \( x_i, i \in I \). The \( i^{\text{th}} \) SU uses \( x_i \) for representing its measurement statistics of the energy detection. By reaching consensus, we mean the individual state \( x_i \) asymptotically converge to a common value \( x^* \), i.e.,

\[
x_i(k) \to x^* \text{ as } k \to \infty, \forall i \in I,
\]

where \( k \) is the discrete time step, \( k = 0, 1, 2, \ldots, \) and \( x_i(k) \) is updated based on the previous states of node \( i \) and its neighbors.

C. Sensing and Measurement Stage

In the sensing stage, we adopt the energy detector [26] because it requires lower design complexity and no priori knowledge of primary users. For the \( i^{\text{th}} \) SU, the received signal \( y_i(t) \) is modeled as

\[
y_i(t) = \begin{cases} 
N_0(t), & H_0, \\
h_i s_i(t) + N_0(t), & H_1,
\end{cases}
\]

where \( h_i \) represents gain of the channel, \( s_i(t) \) is the signal from PU, \( N_0(t) \) is the additive white Gaussian noise.
AWGN), i.e., $n_i(t) \sim \mathcal{N}(0, \sigma_i^2)$. We call $n_i(t)$ the sensing noises and collect their variances into a vector

$$\sigma = [\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2]^T.$$  

(5)

Without loss of generality, $s(t)$ and $\{n_i(t)\}$ are assumed to be independent of each other.

According to [27], each secondary user calculates a summary statistic $Y_i$ over a detection interval of $m$ samples, i.e.,

$$Y_i = \sum_{t=0}^{m-1} |y_i(t)|^2 \quad i \in \mathcal{I}$$  

(6)

where $m$ is determined from the time-bandwidth product $TW$.

Under AWGN measurement channels, the test statistic of the $i$th secondary user using energy detection is given by (6). Since $Y_i$ is the sum of the squares of $m$ Gaussian random variables, it can be shown that $Y_i/\sigma_i^2$ follows a central chi-square distribution with $m$ degrees of freedom if $\mathcal{H}$ is true; otherwise, it would follow a non-central chi-square distribution with $m$ degrees of freedom and parameter $\eta_i$. That is,

$$Y_i/\sigma_i^2 = \begin{cases} 
\chi^2_m, & \mathcal{H}_0 \\
\chi^2_m(\eta_i), & \mathcal{H}_1 
\end{cases}$$  

(7)

where

$$\eta_i = E_s|h_i|^2/\sigma_i^2$$  

(8)

is the local SNR at the $i$th SU and the quantity

$$E_s = \sum_{t=0}^{m-1} |s(t)|^2$$  

(9)

represents the transmitted signal energy over a sequence of $m$ samples during each detection interval. Note that the so-defined local SNR is $m$ times the average SNR at the output of the energy detector, which should be equal to $E_s|h_i|^2/m\sigma_i^2$. For convenience, we put all $\eta_i$ into the following vector:

$$\eta = [\eta_1, \eta_2, \ldots, \eta_n]^T$$  

(10)

According to the central limit theorem, if the number of samples $m$ is large enough (e.g., $\geq 10$ in practice), the test statistics $Y_i$ are asymptotically normally distributed with mean

$$E(Y_i) = \begin{cases} 
\frac{m\sigma_i^2}{(m + \eta_i)} \mathcal{H}_0 \\
\frac{m\sigma_i^2}{(m + \eta_i)} \mathcal{H}_1 
\end{cases}$$  

(11)

and variance

$$\text{Var}(Y_i) = \begin{cases} 
\frac{2m\sigma_i^4}{(m + \eta_i)^2} \mathcal{H}_0 \\
\frac{2m\sigma_i^4}{(m + 2\eta_i)\sigma_i^4} \mathcal{H}_1 
\end{cases}$$  

(12)

Gaussian distribution approximation will facilitate the optimal weights design of the soft weighted combining.

D. Centralized Weighted Combining

Centralized cooperative spectrum sensing combines the measurements of the SUs at a fusion center as [28]

$$Y_g = \sum_{i=1}^{n} \omega_i Y_i = \omega^T Y_i$$  

(13)

where $\omega = [\omega_1, \omega_2, \ldots, \omega_n]^T, \omega_i \geq 0$ is the weighting ratio, $Y = [Y_1, Y_2, \ldots, Y_n]^T$ is the measurement of the CR network.

Assume the reporting channel is noise free and all $\{Y_i\}$ are assumed to be normal random variables, $Y_g$ is also normally distributed and has mean

$$\bar{Y}_g = EY_g = \begin{cases} 
m\sigma^T \omega \mathcal{H}_0 \\
(m\sigma + E_s g)^T \omega \mathcal{H}_1 
\end{cases}$$  

(14)

where

$$\sigma = [\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2]^T,$$  

(15)

$$g = ||h_1|^2, |h_2|^2, \ldots, |h_n|^2||^T,$$  

(16)

and the variances under different hypotheses are respectively given by

$$\text{Var}(Y_g|\mathcal{H}_1) = \text{E}(Y_g - \bar{Y}_g)^2$$  

(17)

$$= \omega^T \text{E}[(Y - \bar{Y}_{\mathcal{H}_0})(Y - \bar{Y}_{\mathcal{H}_0})^T|\mathcal{H}_0] \omega$$  

(18)

where $\bar{Y} = EY$ and $l \in \{0, 1\}$, specifically, we have

$$\text{Var}(Y_g|\mathcal{H}_1) = \begin{cases} 
\omega^T \Sigma_{\mathcal{H}_0} \omega & \text{under } \mathcal{H}_0 \\
\omega^T \Sigma_{\mathcal{H}_1} \omega & \text{under } \mathcal{H}_1. 
\end{cases}$$  

(19)

With a test threshold $\lambda$, we have

$$Y_g \gtrless \lambda$$  

(20)

and, the performance of the proposed cooperative spectrum detection scheme can be evaluated as

$$P_f = Q \left[ \frac{\lambda - N\sigma^T \omega}{\sqrt{\omega^T \Sigma_{\mathcal{H}_0} \omega}} \right]$$  

(21)

and

$$P_d = Q \left[ \frac{Q^{-1}(P_f) \sqrt{\omega^T \Sigma_{\mathcal{H}_0} \omega - E_s g^T \omega}}{\sqrt{\omega^T \Sigma_{\mathcal{H}_1} \omega}} \right]$$  

(22)

Given a fixed false alarm $P_f$, maximizing $P_d$ in Eqn. (39) will yield the optimal weights $\omega$, see [29], where an optimal solution based on modified deflection coefficient are discussed on centralized soft combining.

III. SPECTRUM SENSING USING WEIGHTED AVERAGE CONSENSUS

In this section, we present our new consensus-based distributed scheme to achieve the weighted measurement combining through local interactions among SUs, instead of processing the measurements in a centralized fusion center.
A. Weighted Average Consensus Algorithm

Following the notation of consensus [23], we denote the value of the $i^{th}$ agent’s measurement $Y_i$ as $x_i$, and the proposed weighted average consensus-based combining scheme is shown as

$$x_i(k + 1) = x_i(k) + \frac{\alpha}{\delta_i} \sum_{j \in N_i(k)} (x_j(k) - x_i(k)), i \in \mathcal{I}$$

(25)

where $\alpha$ is the iteration step size satisfying the maximum node degree constraint [23], $N_i(k)$ denotes neighboring node of the $i^{th}$ SU at time step $k$, and $\delta_i \geq 1$ is the weighting ratio.

If the communication topologies formed by the SU network are jointly connected, all the SUs' decision statistics are jointly connected, all the SUs' decision statistics will reach consensus. The final convergence value is:

$$x_i(k) \to x^* = \sum_{i=1}^{\infty} \delta_i x_i(0)$$

(26)

The rigorous proof of the convergence is given in Section IV. By comparing the decision value $x^*$ with a pre-defined threshold $\lambda$, every SU locally obtains the global decision as:

$$\text{Decision} \ H = \begin{cases} 1, & x^* > \lambda \\ 0, & \text{otherwise} \end{cases}$$

(27)

Remark 1: An important feature of the algorithm is the choice of the weighting factor $\delta_i$. If the $i^{th}$ SU has better measurement channel condition, it sets a larger $\delta_i$, which makes the iteration (25) rely less on the local information exchange. On the contrary, an SU with poor measurement channel sets a smaller $\delta_i$ and relies more on the information from the network in order to improve the overall performance. Therefore, the statistics value across the whole SU network will be dominated by the SUs with better performance. Therefore, the statistics value across the whole SU network will be dominated by the SUs with better performance.

Remark 2: The average consensus-based combining approach [13][14][20] is a special case in our proposed algorithm when $\delta_i = 1, \forall i \in \mathcal{I}$. Compared to the centralized soft combining in Eqn. (13), the distributed consensus iteration in Eqn.(25) achieves an equivalent combining as

$$\omega_i = \frac{\delta_i}{\sum_{j=1}^{N_i} \delta_j}.$$ 

(28)

After the convergence, the final decision statistic $x^*$ equals the global combing $Y'_{\mathcal{I}}$, and every SU holds a weighted global decision consensus only through local information exchange.

B. Transient Performance of the Distributed Weighted Combining

If we write the algorithm of Eqn.(29) in the compact form:

$$x(k + 1) = Wx(k),$$

(29)

where $x = [x_1, \ldots, x_n]^T$, and $W = [w_{ij}]$ is the Perron matrix defined in [23]. We will prove in Section IV that

$$\lim_{k \to \infty} W^k = \frac{1}{\delta^T I}$$

(30)

where

$$\delta = [\delta_1, \delta_2, \ldots, \delta_n]^T,$$

(31)

where $\delta_i$ is the weighting ratio in the algorithm (29).

If we assume that $x_i(0)$ follows a normal distribution as discussed in Section II-C, we have

$$x_i(k) = [W]x(0)$$

(32)

where $[W]_i$ denotes the $i^{th}$ row of the matrix $W$. Therefore, $x_i(k)$ is a weighted average of Gaussian distributed random variables, which is also Gaussian distributed, i.e.,

$$\mathcal{H}_0 : x(k)_i \sim \mathcal{N}\left(n \sum_{j=1}^{\infty} w_{ij} \sigma_i^2, \sqrt{2m \sum_{j=1}^{\infty} w_{ij}^2 \sigma_i^4}\right)$$

$$\mathcal{H}_1 : x(k)_i \sim \mathcal{N}\left(n \sum_{j=1}^{\infty} w_{ij}(m + \eta_i) \sigma_i^2, \sqrt{m \sum_{j=1}^{\infty} w_{ij}^2(2m + 2\eta_i) \sigma_i^4}\right)$$

where $w_{ij}$ is the element of matrix $W^k$ at the $i$th row and $j$th column. Thus, the probability of detection at the $i$th SU at time $k$ is given by

$$P_d(k)_i = Q(\lambda; \mu_i, \nu_i)$$

(33)

$$P_d(k)_i = Q(\lambda; \mu_i, \nu_i)$$

(34)

where $Q(\cdot)$ is the complementary cumulative distribution function of Gaussian variable defined in Section II-A, $\lambda$ is the decision threshold, and

$$\{\mu_i, \nu_i\} = \left\{ \sum_{j=1}^{\infty} w_{ij} \sigma_i^2, \sqrt{2m \sum_{j=1}^{\infty} w_{ij}^2 \sigma_i^4}\right\}$$

From Eqn. (30), we have

$$\lim_{k \to \infty} w_{ij} = \delta_j,$$

(35)

where $\delta_j$ is the weighting ratio in the algorithm (25).

Practically, it’s unnecessary to process the algorithm for the infinite iteration. We can use Eqn. (34) with respect to the time step $k$ as an evaluation for the transient performance in finite steps of the consensus based cooperative spectrum sensing schemes.

C. Distributed Optimal Weight Design

In this subsection, we describe the distributed optimal weight design based on the proposed weighted average consensus algorithm. First, we consider the combined global statistic in Eqn. (13) and obtain an optimized centralized weights $\omega_{\mathcal{I}}$. Then, using Eqn. (28), we obtain the distributed optimal weights $\delta_{\mathcal{I}}$. 
Given a false alarm constraint $P_f$, the optimal weights can be obtained by maximizing $P_d$. Generally speaking, the closed form solution does not exist for maximizing $P_d$ in Eqn. (24). To give an optimal weights design, we consider maximizing the deflection coefficient [30] for a near optimal weight design. Specifically, the centralized optimal solution is

$$\omega_{oi} = \frac{\eta_i}{\sum_{i=1}^{n} \sigma_i^2}, \quad (36)$$

where $\omega_{oi}$ denotes the optimal value of $\omega_i$, $\eta_i$ is the local SNR defined in Eqn.(8) and $\sigma_i^2$ is the variance of the Gaussian noise in the measurement channel. Comparing Eqn. (28) and (36), we obtain the distributed weights as

$$\delta_{oi} = \frac{\eta_i}{\sigma_i^2}. \quad \text{(37)}$$

**Remark 3:** Because the weighted average consensus ensures the linear combining, the uniformed weights should be in a linear form as in Eqn. (36). All the $\delta_{oi}$ need to be scaled or saturated to be larger than 1 without affecting the convergence of the consensus iteration under i.i.d. AWGN channel.

To show the optimality of the weight in Eqn. (37), we define the deflection coefficient based on the cooperative spectrum sensing settings, as

$$d^2(\omega) = \frac{\left[ E(Y_g|H_1) - E(Y_g|H_0) \right]^2}{\text{Var}(Y_g|H_0)} = \frac{(E_d g^T \omega)^2}{\omega^T \Sigma_{H_0} \omega}. \quad \text{(38)}$$

Rewriting Eqn. (24) as

$$P_d = \frac{Q^{-1}(P_f) - \frac{E_d g^T \omega}{\sqrt{\omega^T \Sigma_{H_0} \omega}}}{\sqrt{\frac{\omega^T \Sigma_{H_0} \omega}{\omega^T \Sigma_{H_0} \omega}}} \quad \text{(39)}$$

$$= Q^{-1}(P_f) - \frac{E_d g^T \omega}{\sqrt{\omega^T \Sigma_{H_0} \omega}} \sqrt{1 + \frac{4 \omega^T \Sigma_{H_0} \omega}{\omega^T \Sigma_{H_0} \omega}} \quad \text{(40)}$$

where $Q(\cdot)$ denotes the complementary cumulative distribution function. From Eqn. (40), we can see that in low SNR channel condition when $\frac{E_d g^T \omega}{\sqrt{\omega^T \Sigma_{H_0} \omega}} \ll 1$, maximizing $d^2(\omega)$ will yield a near optimal weights design. We formulate the problem as

$$\max \limits_{\omega} d^2(\omega) \quad \text{(41)}$$

subject to

$$\sum_{i=1}^{N} \omega_i = 1, \quad \text{st.} \quad \omega_i > 0, \forall i \in \mathcal{I}. \quad \text{(42)}$$

Solving (41), we can obtain optimal distributed solution using Eqn. (28).

Substituting $\omega' = \Sigma_{H_0}^{-1/2}$ into (38) yields

$$d^2_m(\omega) = \frac{E^2_d \omega'^T \Sigma_{H_0}^{-1/2} g g^T \Sigma_{H_0}^{-1/2} \omega'}{\omega'^T \omega'} \leq \frac{E^2_d \lambda_{max}}{\Sigma_{H_0}^{-1/2} g g^T \Sigma_{H_0}^{-1/2}} \quad \text{(44)}$$

We can see that the matrix $\Sigma_{H_0}^{-T/2} g g^T \Sigma_{H_0}^{-1/2}$ is a rank one matrix having the nonzero eigenvalue $\lambda_{max} = \|\Sigma_{H_0}^{-T/2} g\|^2_2$, and the associated eigenvector $\Sigma_{H_0}^{-T/2} g$. Let

$$\omega' = \Sigma_{H_0}^{-T/2} g, \quad \text{(47)}$$

$d^2(\omega)$ will achieve the maximum value $E^2_d \|\Sigma_{H_0}^{-T/2} g\|^2_2$. Therefore, the uniformed optimal weight is

$$\omega_0 = \frac{\Sigma_{H_0}^{-1/2} \omega'}{1^T \Sigma_{H_0}^{-1/2} \omega'} = \frac{\Sigma_{H_0}^{-1/2} g}{1^T \Sigma_{H_0}^{-1/2} g}. \quad \text{(48)}$$

Because $\Sigma_{H_0}$ defined in Eqn.(20) is a diagonal matrix, we have

$$\omega_{oi} = \frac{\eta_i}{\sum_{i=1}^{n} \eta_i \sigma_i^2} = \frac{\eta_i}{\sum_{i=1}^{n} \eta_i \sigma_i^2}. \quad \text{(49)}$$

Using Eqn. (28), we can choose

$$\delta_{oi} = \frac{\eta_i}{\sigma_i^2}. \quad \text{(50)}$$

as a distributed optimal design. Thus, the final consensus value is the near optimal soft weighted combining.

**IV. CONVERGENCE ANALYSIS OF THE WEIGHTED AVERAGE CONSENSUS ALGORITHM**

In this subsection, we rigorously prove the convergence of the consensus-based combining algorithm in Eqn. (25) under fixed and dynamic communication channel conditions. We further characterize the convergence rate assuming each communication link has a failure probability.

1) **Fixed communication channel:** For convenience, we re-write the algorithm (25) in the following compact form:

$$x(k + 1) = W x(k), \quad \text{(51)}$$

where $x = [x_1, \ldots, x_n]^T$, and $W$ is defined as

$$W = I - \alpha \Delta^{-1} L, \quad \text{(52)}$$

where $\Delta = \text{diag}[\delta_1, \ldots, \delta_n]$, $L \in \mathbb{R}^{n \times n}$ is the Laplacian matrix defined in Eqn. (19). The stepsize $\alpha$ satisfies

$$0 < \alpha < \frac{1}{d_{max}}, i \in \mathcal{I}. \quad \text{(53)}$$

The convergence of Eqn. (51) depends on the convergence of the infinite matrix product

$$\lim_{k \to \infty} W^k = \frac{1}{\delta^T} \text{I}, \quad \text{(54)}$$

where

$$\delta = [\delta_1, \delta_2, \ldots, \delta_n]^T, \quad \text{(55)}$$

and $\delta^T$ is the left eigenvector of $W$ associated with the eigenvalue 1. We have the following theorem.

**Theorem 1:** For the iteration process (25), if the stepsize $\alpha$ satisfies maximum node degree constraint (53), and the elements of matrix $\Delta = \text{diag}[\delta_1, \ldots, \delta_n]$ satisfy $\delta_i \geq \delta_{max}$
1, ∀i ∈ I, and the communication graph is fixed, then the iteration exponentially converges to
\[
\lim_{k \to \infty} W^k x(0) = \frac{\sum_{i=1}^{n} \delta_i x_i(0)}{\sum_{i=1}^{n} \delta_i}. \tag{56}
\]
That is,
\[
x^* = \lim_{k \to \infty} x(k) = \frac{\sum_{i=1}^{n} \delta_i x_i(0)}{\sum_{i=1}^{n} \delta_i}. \tag{57}
\]

Proof: The proof mainly follows from the spectral decomposition and Perron Frobenius Theorem [31]. For the main proof, we need the following lemma:

Lemma 1: [23] Let \( G \) be a strongly connected digraph with \( n \) nodes and the maximum node degree \( d_{\text{max}} \). Then, the associated Perron matrix \( W \) defined as \( W = I - \alpha L \) with parameter \( 0 < \alpha < \frac{1}{d_{\text{max}}} \) satisfies the following properties:

i) \( W \) is a row stochastic nonnegative matrix with a trivial eigenvalue of 1; ii) \( W \) has the simple eigenvalue \( \lambda_1 = 1 \) as the spectral radius \( \rho(W) \); iii) All eigenvalues of \( W \) are in a unit circle \( |\lambda_i| < 1, i = 2, \ldots, n \).

To characterize the convergence of the algorithm (51), we make the spectral decomposition [31] of \( W \), shown as
\[
W = \lambda_1 J_1 + \lambda_2 J_2 + \ldots + \lambda_n J_n, \tag{58}
\]
where \( \lambda_i \in \mathbb{C}, i \in I \) is the spectrum of the \( W \), \( J_i, i \in I \) are the spectral projectors of \( W \). \( J_i \)'s satisfy

1) \( J_i \) is the projector onto \( N(W-\lambda_i I) \) along \( R(W-\lambda_i I) \), \( J_i J_j = 0 \) whenever \( i \neq j \),
2) \( J_i J_j = 0 \) whenever \( i \neq j \),
3) \( J_1 + J_2 + \ldots + J_n = I \).

Then, we obtain
\[
\lim_{k \to \infty} W^k = \lim_{k \to \infty} \left( \lambda_1^k J_1 + \lambda_2^k J_2 + \ldots + \lambda_n^k J_n \right). \tag{59}
\]
From Eqn. (59), we can see that \( \rho(W) \leq 1 \) is necessary for the convergence of the algorithm (51), and the eigenvalue \( \lambda_1 = 1 \) will decide the final convergence value. Note that the spectral decomposition and projection are in the \( W \)'s generalized eigenspace \( E(W) \subset \mathbb{C}^{n \times n} \), since \( W \) is not symmetric.

For a connected undirected graph \( G \), if we see each undirected link of the graph \( G \) as two directed links with opposite directions and different weights, as shown in Fig. 1, then \( L = \Delta^{-1}L \) is the Laplacian matrix of a strongly connected digraph \( G \). We choose \( \delta_i > 1, \forall i \in I \) to ensure \( d_{\text{max}}(G) \leq d_{\text{max}}(\Delta^{-1}L) \), so that \( \alpha \) will satisfy the maximum degree constraint of graph \( G \). For a strongly connected graph, according to Lemma 3 in [23], the Perron matrix \( W \) has the simple eigenvalue \( \lambda_1 = 1 \) as the spectral radius \( \rho(W) \); all the other eigenvalues of \( W \) are in a unit circle \( |\lambda_i| < 1, i = 2, \ldots, n \). Therefore, from Eqn. (59), we obtain
\[
\lim_{k \to \infty} W^k = J_1. \tag{60}
\]
Since \( \lambda_1 = 1 \) is the simple eigenvalue, the associated projector \( J_1 \) is given by [31]
\[
J_1 = \frac{1}{\delta^T} \delta. \tag{61}
\]
where \( \delta \) defined in Eqn. (55). It means
\[
\lim_{k \to \infty} W^k = \frac{1}{\delta^T} \delta^T \mathbf{1}, \tag{62}
\]
which leads to Eqn. (56) and Eqn. (57).

Remark 4: Theorem 1 is a direct application of the famous Perron Frobenius Theorem [31]. Slightly different versions of this theorem are presented in [17] and [32]. Setting \( \delta_i \geq 1, \forall i \in I \), is to ensure convergence of the consensus algorithm.

Remark 5: Setting weights \( \delta_i \) in the consensus algorithm makes the information flow rate balance between any pair of SU nodes. For any pair of neighboring SUs \( (v_i, v_j) \), the \( i^{th} \) SU has the stepsize \( \frac{1}{\delta_i} \), while the \( j^{th} \) SU has the stepsize \( \frac{1}{\delta_j} \). This makes the network matrix, Laplacian matrix and Perron matrix, asymmetric, and the final convergence value deviates from the average consensus. Intuitively speaking, in Fig. 1, the pair of SUs on the left has equal link weights and same stepsize \( \alpha \). In the iteration process, \( \forall i, j \), the statistics held by the \( i^{th} \) SU and \( j^{th} \) SU converge to each other in the same rate, and the iteration stops when they meet at the average of both initial values, which leads to the average consensus convergence of the whole network. In contrast, the pair of SUs on the right has different link weights and different stepsize in the iteration (25). In the iteration process, statistics held by the \( i^{th} \) SU and \( j^{th} \) SU converge to each other in different rates, which makes the final consensus value deviate from the average consensus and converge to a linearly weighted average of the initial measurement statistics of the whole SU network.

Fig. 1. Information flow with weights

Remark 6: Setting \( \Delta = I \), all the weight \( \delta_i = 1, \forall i \), we have \( W \) as a symmetric matrix with real eigen spectrum and eigen space. \( I \) is the simple and largest eigenvalue of \( W \), the vector \( 1 \) and \( 1^T \) are the associated left and right eigenvectors respectively. The convergence of the consensus iteration is given as
\[
\lim_{k \to \infty} W^k x(0) = \frac{1}{\delta^T} \delta^T x(0) = \frac{\sum_{i=1}^{n} x_i(0)}{n} \mathbf{1}, \tag{63}
\]
which is the average consensus algorithm extensively studied in the literature [18][22][33], to name a few.

2) Dynamic communication channel: Realistic SU networks suffer from noise and error interruption or power use constraints. Link failures and dynamic switching communication channels should be considered. In this subsection, we
characterize the conditions for the weighted average consensus convergence on the dynamic communication channels.

For a network of $n$ secondary users, there are a finite number, say a total of $r$, of possible communication graphs. We denote the set of all possible graphs by $\{G_1, \ldots, G_r\}$, and the set of corresponding Perron matrices by $\{W_1, \ldots, W_r\}$. The weighted average consensus algorithm is given by

$$x(k+1) = W_s(k)x(k),$$

where the indices $s(k)$ are integers and satisfy $1 \leq s(k) \leq r$ for all $k > 0$.

**Theorem 2:** For the iteration process (64), if the step size $\alpha$ satisfies

$$0 < \alpha < \frac{1}{n},$$

where $n$ is the number of the SU nodes in the network, and the elements of matrix $\Delta = \text{diag}\{\delta_1, \ldots, \delta_r\}$ satisfy $\delta_i \geq 1, \forall i \in \mathcal{I}$, and the collection of bidirectional communication graphs that occur infinitely often are jointly connected, then the iteration converges to

$$\lim_{k \to \infty} x_i(k) = \frac{\sum_{i=1}^{n} \delta_i x_i(0)}{\sum_{i=1}^{n} \delta_i}, \forall i \in \mathcal{I}. \quad (66)$$

**Proof:** We show that consensus iteration (64) is actually a paracontraction process under the $\ell_\infty$ norm and its fixed point is decided by the eigenspaces of the related Perron matrices.

A matrix $M \in \mathbb{R}^{n \times n}$ is called paracontracting [34] with respect to a vector norm $\| \cdot \|$ if

$$Mx \neq x \Leftrightarrow \|Mx\| < \|x\|. \quad (67)$$

For a matrix $M$, we denote $\mathcal{H}(M)$ as its fixed-point subspace, i.e., $\mathcal{H}(M) = \{x | x \in \mathbb{R}^n | Mx = x\}$. Apparently, $\mathcal{H}(M)$ is $M$’s eigenspace associated with the eigenvalue 1.

Before presenting the main proof, we give two related lemmas as following:

**Lemma 2:** [18] If a collection of graphs $\{G_1, \ldots, G_p\}$ are jointly connected, then their corresponding Perron matrices satisfy

$$\bigcap_{i=1}^p \mathcal{H}(W_i) = \mathcal{H} \left( \frac{1}{p} \sum_{i=1}^p W_i \right) = \text{span}(1). \quad (68)$$

The proof of Lemma 2 follows the same procedure in the proof of Lemma 2 in [18]. For the jointly connected collection of possible graphs $\{G_1, \ldots, G_r\}$, $r \geq p$, we have

$$\bigcap_{i=1}^r \mathcal{H}(W_i) = \bigcap_{i=1}^p \mathcal{H}(W_i) = \text{span}\{1\}. \quad (69)$$

**Lemma 3:** For any possible graph $G$, the associated graph Perron matrix is $\hat{W} = I - \alpha \Delta^{-1}L$, we have $\|\hat{W}\|_\infty \leq 1$. For any jointly connected graph sequence $\{G_1, \ldots, G_p\}$, then the matrix

$$\hat{W} = \bigcap_{i=1}^p W_i \quad (70)$$

is a paracontracting matrix having 1 as the right eigenvector associated with the simple eigenvalue 1.

To prove Lemma 3, we firstly show that $\|\hat{W}\|_\infty \leq 1$, which is equivalent to the fact that the maximum value in the network is non-increasing and the minimum value in the network is non-decreasing. Under any possible undirected graph $G$ and the associated Perron matrix $\hat{W} = I - \alpha \Delta^{-1}L$ defined in (52), if we assume the $\epsilon$th SU holds the maximum value in the network, we have the algorithm (25) as

$$x_{max}(k+1) = x_{max}(k) + \frac{\alpha}{\delta_i} \sum_{j \in N_i} (x_j(k) - x_{max}(k)) = (1 - \frac{\alpha}{\delta_i}) x_{max}(k) + \frac{\alpha}{\delta_i} \sum_{j \in N_i} x_j(k)$$

because $0 < \alpha < \frac{1}{n}, n \geq N_i$ and $\delta_i > 1$, we have $0 < \|N_i\|_{\max} < 1$, which means $x_{max}$ is non-increasing in every step of the iteration and $x_{max}$ always stays in the convex hull formed by $x_{max}$ and its local neighbors, no matter how the graphs are sequenced. Following the same procedure, we can prove $x_{min}$ is non-decreasing in every step of the iteration and $x_{min}$ always stays in the convex hull formed by $x_{min}$ and its local neighbors. Therefore, we have $\|\hat{W}\|_\infty \leq 1$ which leads to $\|\hat{W}\|_\infty \leq \prod_{i=1}^p \|W_i\|_\infty \leq 1$.

Secondly, we prove that $\hat{W}$ is paracontracting. According to Lemma 2, we know $\hat{W}$ has a simple eigenvalue 1 associated with the eigenvector 1. Therefore, in the iteration $x(k+p) = \prod_{i=1}^p \hat{W}_i x(k) = \hat{W} x(k)$, for $x(k) \notin \text{span}(1)$, then $\max\{x_{max}, |x_{min}|\}$ is strictly decreasing, which means $\|\hat{W}_q x\|_\infty < \|x\|_\infty$. If, say, $|x_{max}| > |x_{min}|$ and $x_{max}$ remains the same after the iteration, then the vector $[0, \ldots, 0, x_{max}, 0, \ldots, 0]^T$ is also an eigenvector associated with the eigenvalue 1, which means 1 is not a simple eigenvalue of $\hat{W}$ and it contradicts with Lemma 2. Thus, we have $\hat{W}$ is paracontracting according to the definition Eqn.(67). This finishes the proof of Lemma 3.

Under the condition that the collection of the jointly connected graphs occurs infinitely, for any graph sequence $\{G_1, \ldots, G_k\}, k > 0$, we can divide it into $h$ jointly connected graph subsequences $\{G_1, \ldots, G_k\} \sqcup \sum_{j=1}^h \tilde{G}_j = k, 0 < h \leq k$. Then we can rewrite Perron matrix sequence as

$$\prod_{s=1}^k W_s = \prod_{j=1}^h \left( \prod_{i=1}^{k_i} W_i \right) = \prod_{j=1}^h \tilde{W}_j, \quad (71)$$

and $\tilde{W}_j = \prod_{i=1}^{k_i} W_i$ are paracontracting for all $1 \leq j \leq h$ according to Lemma 3. So $\prod_{s=1}^k W_s(k)$ is an infinite paracontracting process as $k \to \infty$.

To prove the final convergence value, we decompose the generalized eigenspace $\mathbb{E}_\alpha$ of $\hat{W}$ as $E_1 \oplus \mathbb{E}_\alpha$, where $\hat{W}$ is defined in Eqn. (70) and $E_1 = \text{span}\{1\}$ associated with the simple eigenvalue 1, and $E_\alpha$ is the collection of all the other eigenspaces. For any initial value $x(0)$, we decompose $x(0) = u(0) + w(0)$, where $u(0) \in E_1$ and $w(0) \in E_\alpha$. Because $u(0) \in \bigcap_{i=1}^r \mathcal{H}(W_i)$, the sequence given by $u(k+1) = W_s(k)u(k)$ is constant, and the limit $\lim u(k) = u(0)$. On the other hand, because $E_\alpha$ is invariant under all $W_s(k)$, the sequence of vectors given by $w(k+1) = W_s(k)w(k)$ all belong to $E_\alpha$ which leads to the limit $u^* \in E_1$. From the
properties of the infinite paracontracting process [34], \( w^* \) has to be in the fixed subspace of \( W \). It means \( w^* \in \text{span} \{1\} = E_1 \), which yields \( w^* = 0 \), because \( E_1 \cap E_s = \emptyset \). Therefore, the iterative process (64) has the limit \( x^* = u^* \), and \( u^* \) is given by the spectral projection of matrix \( \tilde{W} \), associated with the simple eigenvalue 1, as [31] \( x^* = u^* = \frac{1}{\sqrt{T}} \prod_{t=0}^{T-1} \delta_t \), where \( \delta = [\delta_1, \ldots, \delta_n]^T \) and \( \delta_i \) is the element of the diagonal matrix \( \Delta \). This finishes the proof of Theorem 2. 

Remark 7: Theorem 2 encompasses the average consensus as a special case when \( \delta = 1 \) and \( W_{s(k)} \) are symmetric matrices. For symmetric \( W_{s(k)} \), we have \( \|W_{s(k)}\|_2 = \rho(W_{s(k)}) \leq 1 \), based on which the convergence analysis is given in [18]. For asymmetric \( W_{s(k)} \), we adopt the \( \mathcal{L}_\infty \) norm \( \|W_{s(k)}\|_\infty \) for the convergence analysis. Meanwhile, the fixed communication topology, Theorem 1 is a special case when \( W_{s(k)} = W, \forall k \geq 0 \).

Remark 8: Theorem 2 requires weak long-term connectivity which contains both deterministic and stochastic time-varying graph sequences, and the convergence rate in general may not exist. If we further assume each link has an independent probability to fail, e.g., the link erasure model [21], we can give in the next subsection an estimation of the convergence rate of the consensus iteration.

3) Convergence Rate with Random Link Failures: For a SU network denoted as \( G = (\mathcal{E}, \mathcal{V}) \), we assume \( \mathcal{G} \) is a connected undirected graph and \( \mathcal{E} \) is the set of realizable edges. According to the erasure link model [21], we assign each pair of neighboring SUs the online and offline probabilities as \( P_{ij} \) and \( 1 - P_{ij} \), respectively. Then, at the arbitrary time index \( k \), the network of \( n \) SUs is modeled by the graph \( G(k) = (E(k), \mathcal{V}) \), where \( E(k) \) denotes the edge set at time \( k \).

Then the consensus iteration (64) becomes a random process and it is modeled as

\[
x(k+1) = W(k)x(k)
\]

where \( W(k) \) is defined as

\[
W(k) = I - \alpha \Delta^{-1} L(k)
\]

where \( \Delta = \text{diag}[\delta_1, \ldots, \delta_n] \) satisfies \( \delta_i \geq 1, \forall i \in \mathcal{I}, \) \( W(k) \) and \( L(k) \) are the Perron matrix and Laplacian matrix of the dynamic communication graph \( G(k) \) at time \( k \), respectively. We assume the link failures among the SU network happen independently, so all \( L(k) \)'s, and \( W(k) \)'s are independent and identically distributed. We have the following lemma:

Lemma 4: If the SU network forms a connected undirected communication graph \( G = (\mathcal{E}, \mathcal{V}) \), each link \( e_{ij} \in \mathcal{E} \) has the online and offline probability as \( P_{ij} \) and \( 1 - P_{ij} \), where \( P_{ij} \in (0, 1) \), the stepsizes \( \alpha \) in Eqn. (73) satisfies the maximum node degree constraint \( 0 < \alpha < \frac{1}{\delta_{\text{max}}(\mathcal{E})} \), then the vector sequence \( \{x(i)\} \) in (25) converges exponentially in the sense that

\[
\lim_{k \to \infty} \|E(x(k)) - x^*\|_2 = 0. \quad \forall x(0) \in \mathbb{R}^{n \times 1}
\]

The decay factor of the convergence is given by \( \rho(\mathcal{W} - J_1) \), where \( 0 < \rho(\mathcal{W} - J_1) < 1 \) is the spectral radius of \( \mathcal{W} - J_1 \), \( \mathcal{W} = E(W) \), and \( J_1 \) is defined in Eqn. (61).

Proof: we have the error dynamics of the algorithm (72) as

\[
x(k+1) - x^* = W(k)x(k) - J_1x(0), \quad k = 0
\]

hence

\[
W(j)x(0) - J_1x(0).
\]

Since \( \delta \) and 1 are respectively the left and right eigenvector of \( W(k), \forall k \geq 0 \), associated with the eigenvalue \( \lambda_1 = 1 \), we have

\[
x(k+1) - x^* = W(k)x(k) - J_1x(0)
\]

Then the consensus iteration (64) becomes a random process and it is modeled as

\[
x(k+1) = W(k)x(k)
\]

where \( W(k) \) is defined as

\[
W(k) = I - \alpha \Delta^{-1} L(k)
\]

where \( \Delta = \text{diag}[\delta_1, \ldots, \delta_n] \) satisfies \( \delta_i \geq 1, \forall i \in \mathcal{I}, \) \( W(k) \) and \( L(k) \) are the Perron matrix and Laplacian matrix of the dynamic communication graph \( G(k) \) at time \( k \), respectively. We assume the link failures among the SU network happen independently, so all \( L(k) \)'s, and \( W(k) \)'s are independent and identically distributed. We have the following lemma:

Lemma 4: If the SU network forms a connected undirected communication graph \( G = (\mathcal{E}, \mathcal{V}) \), each link \( e_{ij} \in \mathcal{E} \) has the online and offline probability as \( P_{ij} \) and \( 1 - P_{ij} \), where \( P_{ij} \in (0, 1) \), the stepsizes \( \alpha \) in Eqn. (73) satisfies the maximum node degree constraint \( 0 < \alpha < \frac{1}{\delta_{\text{max}}(\mathcal{E})} \), then the vector sequence \( \{x(i)\} \) in (25) converges exponentially in the sense that

\[
\lim_{k \to \infty} \|E(x(k)) - x^*\|_2 = 0. \quad \forall x(0) \in \mathbb{R}^{n \times 1}
\]

The decay factor of the convergence is given by \( \rho(\mathcal{W} - J_1) \), where \( 0 < \rho(\mathcal{W} - J_1) < 1 \) is the spectral radius of \( \mathcal{W} - J_1 \), \( \mathcal{W} = E(W) \), and \( J_1 \) is defined in Eqn. (61).
network topology and the link weights, as well as the link failure probability matrix $L_p$. For optimizing the convergence rate, interested readers can refer to [21], [22], [36].

Practically, it’s unnecessary for the SU network to reach the limit in the consensus iteration. We can derive the upper bound on the iteration number at which all SUs are $\epsilon$ close to the final convergence value in the probability sense, which is called $\epsilon$-convergence in [24].

**Theorem 3:** Under the same condition of Lemma 4, $\forall \epsilon > 0$ and $k \geq T(\epsilon)$, for the iteration (25), we have

$$\Pr\{\max_{1 \leq i \leq n} \|x_i(k) - x^*\|_\infty \geq \epsilon \|H_k\| \} \leq \epsilon, \quad k \in \{0, 1, \ldots\}$$  \hspace{1cm} (90)

and

$$T(\epsilon) \leq \frac{3/2 \log \epsilon^{-1} + 1/2 \log(K)}{1 - E(\|(W - J_1)\|_\infty)}$$  \hspace{1cm} (91)

where

$$K = \sum_{i=1}^{n} (2m\sigma_i^4 + 4E_\epsilon|h_i|^2\sigma_i^2 + (m\sigma_i^2 + E_\epsilon|h_i|^2)^2)$$  \hspace{1cm} (92)

and $J_i$ is defined in (61), $\sigma_i$ is the measurement noise variance for the $i^{th}$ SU, and $E_\epsilon$ is the signal energy defined in (9) and $h_i$ is the channel gain.

**Proof:** Since $\max_{1 \leq i \leq n} \|x_i(k) - x^*\|_\infty = \|x(k) - x^*\|_\infty$,

we have

$$\Pr\{\|x(k) - x^*1\|_\infty \geq \epsilon \|H_k\| \} = \Pr\{\|x(k) - x^*1\|_\infty \geq \epsilon \|H_k^2\| \} \leq E(\|x(k) - x^*1\|_\infty^2 \|H_k^2\|)$$  \hspace{1cm} (94)

where the second equation is from the Markov inequality. Following the proof of Theorem 4, from Eqn. (80), we have

$$\|x(k) - x^*1\|_\infty^2 \leq \frac{1}{\epsilon^2} (\|W(k) - J_1\|_\infty^{2k}) \|x(0)\|_\infty^2$$  \hspace{1cm} (96)

Since $W(k)$’s are identically and independently distributed, we have

$$E(\|x(k) - x^*1\|_\infty^2) \leq E(\|W - J_1\|_\infty^{2k}) E(\|x(0)\|_\infty^2)$$  \hspace{1cm} (97)

If we choose a vector $\tilde{x}$ such that $\|	ilde{x}\|_\infty = 1$ and $\delta_T\tilde{x} = 0$, where $\delta$ is defined in Eqn. (55), we have $J_1\tilde{x} = 0$ and following Lemma 3 in the proof of Theorem 2, we have

$$\|W(k) - J_1\|_\infty = \|W(k)\|_\infty \leq \|\tilde{x}\|_\infty$$  \hspace{1cm} (98)

when $W(k)$ has 1 as a simple eigenvalue, we have

$$\|W(k) - J_1\|_\infty < \|\tilde{x}\|_\infty$$  \hspace{1cm} (99)

which means

$$\|W(k) - J_1\|_\infty = \max_{\|\tilde{x}\|_\infty = 1} \|W(k) - J_1\|_\infty \|\tilde{x}\|_\infty \leq 1$$  \hspace{1cm} (100)

and

$$E(\|W - J_1\|_\infty) < 1$$  \hspace{1cm} (101)

we drop the index of $W$ because $W(k)$ are identically distributed. We also have

$$\|x(0)\|^2_\infty \leq \|x(0)\|^2_2$$  \hspace{1cm} (102)

Substitute (97) and (102) into (95), we have

$$\Pr\{\|x(k) - x^*1\|_\infty \geq \epsilon \|H_k\| \} \leq E\left(\left(\frac{\|W - J_1\|_\infty^{2k}}{\epsilon^2}\right) \|x(0)\|^2_2\right)$$  \hspace{1cm} (103)

Let

$$E(\|W - J_1\|_\infty^{2k}) \|x(0)\|^2_2 = \epsilon,$$  \hspace{1cm} (105)

we obtain

$$T(\epsilon) = \frac{3/2 \log \epsilon^{-1} + 1/2 \log(E(\|x(0)\|^2_2))}{1 - \log(\|W - J_1\|_\infty)}$$  \hspace{1cm} (106)

Therefore, we have

$$T(\epsilon) \leq \frac{3/2 \log \epsilon^{-1} + 1/2 \log(E(\|x(0)\|^2_2))}{1 - \log(\|W - J_1\|_\infty)}$$  \hspace{1cm} (107)

From the inequality $\log(1 + u) \leq u$ when $u$ is small, let

$$1 + u = E(\|W - J_1\|_\infty),$$

we obtain,

$$- \log(\|W - J_1\|_\infty) \geq -u = 1 - E(\|W - J_1\|_\infty).$$

Thus, we have

$$T(\epsilon) \leq \frac{3/2 \log \epsilon^{-1} + 1/2 \log(E(\|x(0)\|^2_2))}{1 - E(\|W - J_1\|_\infty)}$$  \hspace{1cm} (108)

Meanwhile, according to (11) and (12), we have

$$E(\|x(0)\|^2_2) = \sum_{i=1}^{n} E(x_i^2(0))$$  \hspace{1cm} (109)

$$< \sum_{i=1}^{n} E(x_i^2(0))$$  \hspace{1cm} (110)

$$= \sum_{i=1}^{n} \left(\text{Var}(x_i(0)) + E^2(x_i(0))\right)$$  \hspace{1cm} (111)

$$\leq \sum_{i=1}^{n} (2m\sigma_i^4 + 4E_\epsilon|h_i|^2\sigma_i^2 + (m\sigma_i^2 + E_\epsilon|h_i|^2)^2)$$  \hspace{1cm} (112)

where $\sigma_i$ is the measurement noise variance for the $i^{th}$ SU, and $E_\epsilon$ is the signal energy defined in (9) and $h_i$ is the channel gain.

**Remark 10:** $\epsilon$-convergence of the average consensus or gossip algorithm has been extensively studied in [37] [13] [24]. Theorem 3 is a generalization to weighted average consensus convergence with random link failures. From (91), we can see clearly that the convergence rate of $\epsilon$-convergence depends on the desired accuracy $\epsilon$, the measurement channel noise variance $\sigma_i$, signal energy $E_\epsilon$, channel gain $h_i$ and the expectation $E(\|W - J_1\|_\infty)$.

**Remark 11:** Practically, $E(\|W - J_1\|_\infty)$ is not easy to compute. Because the norm $\| \cdot \|$ is a convex function, we have

$$E(\|W - J_1\|_\infty) \geq \|W - J_1\|_\infty$$  \hspace{1cm} (113)

$$\geq \rho(\|W - J_1\|_\infty)$$  \hspace{1cm} (114)

the second inequality is from the property of the matrix spectral radius. Therefore, we can use $\rho(W - J_1)$ as an estimation of the minima of $E(\|W - J_1\|_\infty)$ so that we have an approximation of $T(\epsilon)$. 


V. SIMULATIONS AND DISCUSSIONS

In this section, we conduct simulations to evaluate the effectiveness of our proposed distributed weighted combining scheme. We show the convergence of the weighted consensus algorithm under fixed and dynamic communication channel conditions. We adopt the Monte Carlo simulation to validate the effectiveness of the proposed spectrum sensing scheme by introducing metrics of the spectrum sensing performance evaluation.

A. Network Setup and Consensus Convergence

To compare with the existing consensus-based scheme, we set up a similar scenario as in [14]. In particular, 10 SUs cooperate with each other and form a communication graph as shown in Fig. 2. All SUs are assumed to be static and have uncorrelated measuring channel. They are running the synchronized clock during the sensing process and the consensus iteration.

In the sensing stage, we directly generate the output $Y_i$ of every SU’s energy detector individually under the hypothesis $H_1$ in (7), with $m = 12$ at the selected center frequency and bandwidth of interest. Each SU sets $x_i(0) = Y_i$ and starts the measurement fusion using the proposed scheme with the step size $\alpha = 0.19$ in Eqn. (25). The final decision is made after the consensus result $x^*$ is reached.

Fig. 3(a) shows the convergence of the proposed algorithm iteration under fixed communication graph. We found that within 30 steps the difference of $x_i$ among the SU network are less than 1dB, which indicates the consensus has been reached on the global decision statistics. Fig. 3(b) shows the convergence of the algorithm under random graph process with independent link failures. We set the failure probability of each link to be 0.4. We observed that the consensus is achieved within 35 iteration steps, which is slightly slower than the fixed graph case.

In order to fully characterize the transient performance of the weighted consensus algorithm, we plot the detecting probability $P_d$ (probability of detection) curve with respect to the consensus iteration time step under fixed communication topology shown in Fig. 2(a). We set the false alarm $P_f = 0.2$, the variance of Gaussian noise $\sigma_i = 1, \forall i$, the channel SNR varies from -5 dB to -15 dB. We compare the performance for $\epsilon = 0.19$. Link failure probability = 0.4. Convergence probability = 1.

(c) Detecting probability $P_d$ with respect to iteration time step, $\alpha = 0.19$, fixed communication topology, SU number $N = 10$, $P_f = 0.2$, channel SNR at the output of the energy detection ranges from -5 dB to -15 dB.

Fig. 3. Convergence of the proposed consensus algorithm. (a) Fixed graph, (b) Random subgraph
of decentralized (WGC) and centralized (WGC) approaches. The detection probability $P_d$ is calculated according to Eqn. (34).

As shown in Fig. 3(c), we clearly see the trend of the $P_d$ with respect to the iteration step. Within 40 steps, the detection probability of the 10 SU nodes converges to the one of centralized (WGC).

B. Metrics

For comparison, we mainly consider $P_d$, $P_f$ (probability of false alarm), and average SNR $\gamma$ as metrics. A high $P_d$ will result in high $P_f$, which increases the interference to primary users. On the other hand, a low $P_f$ will result in low $P_d$ and lead to low spectrum utilization. The SNR in the measuring channel reflects the channel condition. It is challenging to achieve high detecting performances under the condition of low SNR.

The threshold $\lambda$ is computed from the false alarm constraints under the hypothesis $H_0$. Soft combining schemes share the same threshold, since the measurement output under $H_0$ is independent of the average SNR when the primary user signal is absent.

C. Performance Evaluation

In this subsection, we compare detection performance of the proposed distributed weighted combing (WGC) with the centralized weighted combining, the decentralized equal gain combining (EGC) [14], and hard bit combining OR-rule scheme (OR) [7], with respect to the measurement channel conditions, PU signal strength and network sizes. During the performance evaluation, we mainly consider the AWGN measurement channel which is an approximation of general fading channels. Meanwhile, we also evaluate the effect of the Rayleigh fading on the detecting performance and we compare the proposed optimal weighted design DWGC with the special heuristic weights (AWGC) in [25] based on estimated average channel SNR.

1) Receiver operating curves under AWGN channel: In Fig. 4, we plot the receiver operating curves (ROC) under AWGN channel for the 10-node SU network shown in Fig. 2. The measurement channel SNR of the SU network ranges from -5dB to -15dB. In this case, the centralized WGC and the proposed DWGC achieve the best performance. The EGC approach has satisfactory performance but worse than the WGC approaches. OR-rule gives the worst performance. As expected, the communication channel failures do not affect the detection performance of DWGC and EGC approaches. Particularly, when the false alarm probability is 0.2, the detection probability of DWGC and centralized WGC is 0.63. The EGC and OR-rule approaches have detection probability of 0.53 and 0.40, respectively. Fig. 4 shows clearly the proposed DWGC scheme achieves comparable performance with centralized weighted combining design and outweighs EGC and OR-rule combining.

2) Receiver operating curves under Rayleigh fading channels: Fig. 5 shows the ROC curves of the DWGC, EGC, centralized WGC, OR-rule and AWGC approaches under Rayleigh fading channel. To compare with the performance under AWGN channel, the average SNR of the Rayleigh fading is set as -10 dB which indicates the channel condition is the same as the AWGN channel in Fig. 4 in average sense. From Fig. 5, we can clearly observe that weighted combining schemes are robust with respect to Rayleigh fading effects. DWGC and centralized WGC schemes still achieve the best performance. The AWGC based heuristic weights design in [25] offers comparable performance with the optimal weighted design, EGC approach and OR-rule approach suffers from the fading and the performance degrade severely. Comparing Fig. 4 and Fig. 5, we see that the performance degradation of the DWGC and centralized WGC due to Rayleigh fading channel effects is within 5%, while the performance of EGC and OR-rule degrades about 50% after the Rayleigh fading is considered.

3) $P_d$ with respect to the PU signal strength: In Fig. 6, we compare the detection performance $P_d$ of proposed DWGC with existing schemes in the literature with respect to the
signal strength of PU. Under AWGN measuring channel, we assume the variance of Gaussian noise $\sigma_i = 1, \forall i$, so the measuring channel SNR $\eta$ directly reflects the signal received signal strength of the PU.

We can clearly see that WGC approaches offer higher detection probability than EGC and OR-rule approaches, when the PU channel has low SNR. Particularly, when channel SNR is 0 dB, the WGC approaches achieve detection probability as 0.77, while EGC approach achieves 0.55 and OR rule has 0.41. When the PU channel SNR is close to 5 dB, the three approaches offer relative the same performance. Fig. 6 illustrates that the DWGC outweighs EGC and OR-rule approaches under low PU signal strength condition.

4) Receiver operating curves with respect to SU network size: In Fig. 7, we show the ROC curves of the DWGC, EGC, OR-rule and centralized WGC, under AWGN measuring channel with different SU network sizes. The results show the proposed DWGC approach achieves comparable performance with centralized WGC. The performance of weighted combining design is sensitive to the network size increase that the detection probability increases with the increase of the network size. We assume the variance of Gaussian noise $\sigma_i = 1, \forall i$, and the measuring channel SNR $\eta$ of the SU network ranges from $-5$dB to $-20$dB.

In Fig. 7(a) and Fig. 7(b), we choose the connected SU network with 20 and 30 nodes, respectively. The simulation runs under identical channel condition. Comparing the two figures, we found that as the network size increases, the performance of WGC approaches and EGC approach are improved significantly and DWGC approach work much better than EGC approach. The performance of OR approach almost remains the same because the decision depends only on the SU node with the best channel available, which is less relevant to the network size. Particularly, when the false alarm is set to 0.2, and the network size increases from 20 nodes to 30 nodes, the detection probability of WGC and DWGC increase from 0.74 to 0.85, the detection probability of EGC increases from 0.52 to 0.69, and the detection probability of OR rule increases from 0.38 to 0.46. We can clearly see that DWGC achieves the best performance under different network sizes.

In Fig. 8, we plot the ROC curves of DWGC, AWGC, EGC, OR-rule and centralized WGC under Rayleigh fading channel with network size 20 and 30 nodes respectively. We can see the proposed DWGC achieves comparable performance with centralized WGC and the performance is much less affected by the Rayleigh fading than EGC and OR-rule and DWGC improves the performance with the network size increases. While the EGC rule is much less sensitive to the network size increase compared to the AWGN measurement channel case shown in Fig. 7. Particularly, when the false alarm is set to 0.2, and the network size increases from 20 nodes to 30 nodes, the detection probability of centralized WGC, DWGC and AWGC increase 10%, while the perfor-
average consensus algorithm for both fixed and time-varying graphs. Through weighted local fusion iteration, each SU holds the global decision statistic from the weighted soft measurement combining throughout the network. We prove the convergence of the iteration and characterize the convergence rate of the consensus algorithm under independent link failure condition. We discuss the optimal weights design for the distributed soft combining and compare the results with the centralized optimal weighted combining. Simulation results show the proposed decentralized combining approach offers a comparable performance with the centralized weighted combining and outperforms existing average consensus-based equal gain combining schemes and the OR-rule hard bit combining scheme.

ACKNOWLEDGMENT

Wenlin Zhang, Zheng Wang, and Yi Guo would like to acknowledge the support by the National Science Foundation under Grant EFRI-1024660. Hongbo Liu and Yingying Chen would like to acknowledge the support by the National Science Foundation under Grant CNS-0954020.

REFERENCES


