Supplementary Material to "Distributed Consensus-based Weight Design for Cooperative Spectrum Sensing"

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

Abstract—This material is a supplement to the paper "Distributed Consensus-based Weight Design for Cooperative Spectrum Sensing". Section 1 offers related literature review on cooperative spectrum sensing and consensus algorithms. Section 2 presents related notations and models of the consensus-based graph theory. Section 3 offers further analysis of the proposed spectrum sensing scheme including detection threshold settings and convergence properties in terms of detection performance. Section 4 presents the proofs for the convergence of the proposed consensus algorithm, and discusses the convergence of the proposed algorithm under random link failure network models. Section 5 shows additional simulation results.

Index Terms—Cooperative spectrum sensing, Weighted average consensus, Cognitive radio networks.

1 Related Literature Review

1.1 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 [1]. 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 [2]. 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 [3], soft gain combining [4], to name a few.

W. Zhang is with Western Digital Corporation, Irvine, CA 92614, USA. This work was done during W. Zhang's Ph.D. study at Stevens Institute of Technology.
Y. Guo and Y. Chen are with Department of Electrical & Computer Engineering, Stevens Institute of Technology, Hoboken, NJ 07030, USA.
H. Liu is with Department of Computer Information and Technology, Indiana University-Purdue University Indianapolis, Indianapolis, IN 46202, USA. This work was done during H. Liu's Ph.D. study at Stevens Institute of Technology.
Z. Wang is with the Department of Electrical & Computer Engineering, University of Toronto, Toronto, ON, M5S 3G4, Canada. This work was done when Z. Wang was a postdoctoral researcher at Stevens Institute of Technology.
J. Mitola is with Allied Communications, Boston, MA 02111, USA. This work was done when J. Mitola was Distinguished Professor and Vice President for The Research Enterprise at Stevens Institute of Technology.

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 [1]. Centralized cooperative sensing [5] 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 [1][6] 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 [7]. The multi-hop communication of the relay-assisted sensing may bring extra power cost than one-hop communication, since all SUs' sensing data need to be relayed from the network nodes to the fusion center or detection node. In addition, the multi-hop communication paths may degrade the sensing data quality and affect the detection performance significantly compared to one-hop communication scenarios. Other factors such as communication channel selection schemes and sensing data coding schemes also need to be considered [8] in the relay assisted cooperative sensing to overcome the disadvantage of the multi-hop paths.

Distributed cooperative sensing first appears in [3] 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. Very recently, bio-inspired consensus scheme is introduced to spectrum sensing in [9][10] for distributed measurement fusion and soft combining. Consensus-based spectrum sensing is a

This is the author's manuscript of the supplemental materials published in final edited form at: http://doi.ieeecomputersociety.org/10.1109/TPDS.2014.2307951

biologically inspired approach learned from swarming behaviors of fish schools and bird fleets. The consensus-based cooperation features self-organizable and scalable network structure and only needs one-hop communication among local neighbors. Recent research work [11] applies belief propagation to distributed spectrum sensing [11], which advances the sensing stage for heterogenous radio environment.

The fusion scheme of the sensing data from the SU network also contribute to the detection performance. There are hard bit combining such as OR rule combining and soft combining including equal gain combing and weighted gain combining. Hard bit combining adopts the decision bit from each SU to achieve global detection, which is less effective compared to soft combing schemes taking average of the statistics from all the SUs. Generally speaking, equal gain combing is to compute the average of the measured statistics of the SU network while weighted gain combining computes the weighted average considering the measurement channel conditions. Therefore, weighted combining offers better detection performance under various channel conditions such as fading and shadowing.

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 [12][10] only ensure equal gain combining of local measurements, which is incomparable with centralized *weighted* combining approaches [4]. 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.

1.2 Related Work in Average Consensus Algorithm

The consensus algorithm was studied in [13] 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 [14][15]. In [14], Jadbabaie 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 [15]. Since the average consensus problem has strong impact on distributed networked systems, it increasingly attracts research attention on decentralized estimation [16], filtering [17], and detection [18], etc.. For signal processing applications, communication constraints and the convergence rate become crucial for performance improvement. Typical problems include communication topology design and optimization [19], convergence rate analysis and optimization [20]. Interested readers are referred to the review papers [21][22] for the complete history of consensus algorithm development.

Compared to the extensively studied average consensus, much less research attention is paid to *weighted* average consensus. As stated in [21], 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.

2 Preliminaries on Graph Theory Notations and SU Network Models

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 $\mathcal{I} = \{1, 2, ..., n\}$ 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 $\mathcal{G} = (\mathcal{E}, \mathcal{V})$, where $\mathcal{V} = \{v_i | i \in \mathcal{I}\}$ is a finite nonempty set of nodes. We refer the i^{th} node as the i^{th} SU. The two names, SU and node, will be used alternatively. The edge set $\mathcal{E} = \{e_{ij} = (v_i, v_j) | i, j \in \mathcal{I}\}$. The set of neighbors of node i is denoted by $N_i = \{j : e_{ij} \in \mathcal{E}\}$. A path in $\mathcal{G}$ consists of a sequence of nodes $v_1, v_2, ..., v_l, l \geq 2$, satisfying $(e_{m,m+1}) \in \mathcal{E}, \forall 1 \leq m \leq l-1$. The graph $\mathcal{G}$ is *connected* if any two distinct nodes in $\mathcal{G}$ 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 $\mathcal{G}(k) = (E(k), \mathcal{V})$, where $E(k)$ is the set of active edges at time k. Let $N_i(k) = \{j \in \mathcal{V} | \{i, j \in E(k)\}\}$, and $d_i(k) = |N_i(k)|$ denote the degree (number of neighbors) of node i at time k.

Let $\mathcal{G}_i = (\mathcal{E}_i, \mathcal{V}), i = 1, \ldots, r$, denote a finite collection of graphs with common vertex set $\mathcal{V}$. Their union is a graph $\mathcal{G}$ with the same vertex set and an edge set that is the union of the $\mathcal{E}_i's$, i.e., $\mathcal{G} = \bigcup_{i=1}^r \mathcal{G}_i = (\bigcup_{i=1}^r \mathcal{E}_i, \mathcal{V})$. The set of undirected graphs $\{\mathcal{G}_1, \ldots, \mathcal{G}_r\}$ 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 $\mathcal{G}$ formed by the secondary user nodes is defined as

$$l_{ij} = \begin{cases} d_i, & \text{if } i = j, \\ -1, & \text{if } i \neq j, \quad j \in N_i, \\ 0, & \text{otherwise}, \end{cases} \tag{1}$$

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

$$d_{max} = \max_i |N_i|. \tag{2}$$

It's easy to see, the undirected graph Laplacian matrix L is symmetric and has the left and right eigenvector $\mathbf{1}^T$ and $\mathbf{1}$ associated with the eigenvalue 1, respectively. For the Laplacian matrix of strongly connected graphs, we have the following lemma:

Lemma 1: [21] Let $\mathcal{G}$ be a strongly connected digraph with n nodes and the maximum node degree Δ. Then, the associated Perron matrix W defined as $W = I - \alpha L$ with parameter $0 < \alpha < \frac{1}{d_{max}}$ satisfies the following properties. i) W is a row stochastic nonnegative matrix with a simple 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 $|\lambda_i| < 1, i = 2, \ldots, n$.

In the context of consensus-based spectrum sensing, for the n SUs modeled by the graph $\mathcal{G}$, the i^{th} SU is assigned a state variable $x_i, i \in \mathcal{I}$. The i^{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., $x_i(k) \to x^*$ as $k \to \infty, \forall i \in \mathcal{I}$, where k is the discrete time step, $k = 0, 1, 2, ...$, and $x_i(k)$ is updated based on the previous states of node i and its neighbors.

3 Weighted Consensus-based Two Stage Sensing

3.1 Convergence in terms of Detection Probability

The distributed fusion based on the Algorithm 1 (Eq. (15)) in the main paper is an iterative process that completes after the convergence is reached. In this subsection, we characterize the convergence of the proposed fusion algorithm in terms of the detection probability. If we write the algorithm in the compact form:

$$x(k+1) = W(k)x(k), \tag{3}$$

where $x = [x_1, \ldots, x_n]^T$, and $W(k)$ is the iteration transition matrix at time step k. We will prove by Theorem 1 and 2 in the main file that $\lim_{k\to\infty} \prod_{i=1}^{k} W(i) = \frac{\mathbf{1}\delta^T}{\delta^T \mathbf{1}}$ and $\lim_{k\to\infty} w_{ij} = \delta_j$, where $\delta = [\delta_1, \delta_2, \ldots, \delta_n]^T$, and δ_i is the weighting ratio set by the i^{th} SU.

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

$$x_i(k) \to x^* = \frac{\sum_{i=1}^{n} \delta_i x_i(0)}{\sum_{j=1}^{n} \delta_i} \text{ as } k \to \infty, \forall i \in \mathcal{I}. \tag{4}$$

If we assume that $x_i(0)$ follows a normal distribution as discussed in Section 2.1 of the main paper, we have $x_i(k) = [W]_i x(0)$, where $[W]_i$ denotes the ith row of the matrix $\prod_{i=1}^{k} W(k)$. Therefore, $x_i(k)$ is a weighted average of Gaussian distributed random variables, which is also Gaussian distributed, i.e.,

$$\begin{aligned} \mathcal{H}_0 : x(k)_i \sim & \quad \mathcal{N}\left\{ m\sum_{j=1}^{n} w_{ij}\sigma_i^2, \sqrt{2m\sum_{j=1}^{n} w_{ij}^2 \sigma_i^4} \right\} \\ \mathcal{H}_1 : x(k)_i \sim & \quad \mathcal{N}\left\{ \sum_{j=1}^{n} w_{ij}(m+\eta_i)\sigma_i^2, \sqrt{\sum_{j=1}^{n} w_{ij}^2 2(m+2\eta_i)\sigma_i^4} \right\} \end{aligned} \tag{5}$$

where w_{ij} is the element of matrix $\prod_{i=1}^{k} W(k)$ at the ith row and jth column. Thus, the probability of detection at the ith SU at time k is given by

$$P_f(k)_i = Q(\lambda; \mu_0, \nu_0) \tag{6}$$
$$P_d(k)_i = Q(\lambda; \mu_1, \nu_1) \tag{7}$$

where $Q(\cdot)$ is the complementary cumulative distribution function of Gaussian variable, λ is the decision threshold, and

$$\begin{aligned} \{\mu_0, \nu_0\} = & \quad \left\{ m\sum_{j=1}^{n} w_{ij}\sigma_i^2, \sqrt{2m\sum_{j=1}^{n} w_{ij}^2 \sigma_i^4} \right\} \\ \{\mu_1, \nu_1\} = & \quad \left\{ \sum_{j=1}^{n} w_{ij}(m+\eta_i)\sigma_i^2, \sqrt{\sum_{j=1}^{n} w_{ij}^2 2(m+2\eta_i)\sigma_i^4} \right\} \end{aligned} \tag{8}$$

Practically, it's unnecessary to process the algorithm for the infinite iteration. We can use Eqn. (7) 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.

Remark 1: From Eqn (6)(7)(8), we can see the choice of the threshold λ depends on the whole network connections and the channel conditions of each SU node.

3.2 Detection Threshold for Each SU

Since the converged combining (4) is independent of SU network structure, the threshold λ can also be deducted for each SU independently. Assuming the initial measurements $x_i(0)$ follows the Gaussian distribution, $x^* = \frac{\sum_{i=1}^{n} \delta_i x_i(0)}{\sum_{i=1}^{n} \delta_i}$ is also Gaussian distributed. Therefore, the false alarm (6) and detection rate (7) are still applicable when we set $\omega_i = \frac{\sum_{i=1}^{n} \delta_i x_i(0)}{\sum_{i=1}^{n} \delta_i}$ in the parameter setting (8).

For each SU to compute the detection threshold λ, we can adopt the inverse of the false alarm (6) with parameters in (8) to obtain $\lambda = Q^{-1}(P_f, \mu_0, \nu_0)$. In Section 3.2 of the main paper, we show that under hypothesis $\mathcal{H}_0$ with the absence of PU's signal, the fusion weight of each SU $\delta_i \approx 1/\sigma_i^2$, where σ_i^2 is the variance of measurement noise mainly depends on measurement devices and general wireless environment. Although μ_0 and ν_0 depends on σ_i from all SU's, that global information can be obtained offline or during the calibration process. If the SU network size is fixed, and all the SUs have similar measurement devices, it's not a strong assumption to further assume all the SU's have the same measurement noises under $\mathcal{H}_0$. Then, the threshold can be set by each SU without centralized communication.

4 Convergence analysis of Weighted Average Consensus Algorithm

4.1 Convergence Proof under Dynamic Communication Channels

This subsection provides the theoretic proof for Theorem 2 of the main paper. For convenience, we restate the setup as follows:

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 $\{G_1, \dots, G_r\}$, and the set of corresponding Laplacian matrices and Perron matrices given by $\{L_1, \dots, L_r\}$ and $\{W_1, \dots, W_r\}$, respectively. For any $1 \le s \le r$, we have

$$W_s = I - \alpha\Delta^{-1}L_s, \tag{9}$$

where $\Delta = \text{diag}\{\delta_1, \dots, \delta_n\}$, and δ_i is the weighting ratio. The weighted average consensus algorithm is given by

$$x(k+1) = W_{s(k)}x(k), \tag{10}$$

where the indices $s(k)$ are integers and satisfy $1 \le s(k) \le r$ for all $k > 0$. Here, we use the notion $W_{s(k)}$ to denote the graph sequence in the iteration because the graph sequences could be stochastic or deterministic. We will use $W(k)$ to denote the stochastic case later.

Proof: We show that consensus iteration (10) is actually a paracontraction process under the $\mathcal{L}_\infty$ norm. Specifically, if we decompose the initial state $x(0)$ in (10) as

$$x(0) = x_c(0) + x_d(0), \tag{11}$$

where $x_c(0)$ means consensus vector that $x_c(k) \in span(1)$, and $x_d(0)$ means the difference vector that $x_c(0)x_c(0)^T = 0$, then the paracontracting means (10) will contracts the norm of the state $x_d(k)$ in each iteration and $x_c(k)$ will remain fixed. When the iteration goes to infinite, $x_d(k)$ will shrink to zero and $x(k)$ goes to $x_c(k)$ which equals to $x_c(0)$.

Before presenting the main proof, we discuss the related definitions and three lemmas as following:

A matrix $M \in \mathcal{R}^{n\times n}$ is called paracontracting [23] with respect to a vector norm $\|\cdot\|$ if

$$Mx \ne x \Leftrightarrow \|Mx\| < \|x\|. \tag{12}$$

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.

Lemma 2: [23] Suppose that a finite set of square matrices $\{W_1, \dots, W_r\}$ are paracontracting. Let $\{i(k)\}_{i=0}^{\infty}$, with $1 \le i(k) \le r$, be a sequence of integers, and denote by $\mathcal{J}$ the set of all integers that appear infinitely often in the sequence. Then for all $x(0) \in \mathbb{R}^n$ the sequence of vectors $x(k+1) = W_{i(k)}x(k), k \le 0$, has a limit $x^* \in \bigcap_{i\in\mathcal{J}} \mathcal{H}(W_i)$.

Lemma 3: [16] If a collection of graphs $\{G_1, \dots, G_p\}$ are jointly connected, then their corresponding Perron matrices satisfy

$$\bigcap_{i=1}^{p} \mathcal{H}(W_i) = \mathcal{H}\left(\frac{1}{p}\sum_{i=1}^{p} W_i\right) = span(\mathbf{1}). \tag{13}$$

The proof of Lemma 3 follows the same procedure in the proof of Lemma 2 in [16]. For the jointly connected collection of possible graphs $\{G_1, \dots, G_r\}, r \ge p$, we have

$$\bigcap_{i=1}^{r} \mathcal{H}(W_i) = \bigcap_{i=1}^{p} \mathcal{H}(W_i) = \text{span}\{\mathbf{1}\}. \tag{14}$$

Lemma 4: For any possible graph G, the associated graph Perron matrix is $W = I - \alpha\Delta^{-1}L$, we have $\|W\|_\infty \le 1$. For any graph sequence $\{G_1, \dots, G_k\}, k > 0$ containing $n-1$ collections of jointly connected graph sequence, that is $\{G_1, \dots, G_{p_j}\}, j = 1, ..., n-1, \sum_{j=1}^{n-1} p_j = k$, then the matrix

$$\tilde{W} = \prod_{j=1}^{n-1}\prod_{i=1}^{p_j} W_i \tag{15}$$

is a paracontracting matrix having $\mathbf{1}$ as the right eigenvector associated with the simple eigenvalue 1.

To prove Lemma 4, we firstly show that $\|\tilde{W}\|_\infty \le 1$, 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 $\mathcal{G}$ and the associated Perron matrix $W = I - \epsilon\Delta^{-1}L$ defined the same as the form of (9), if we assume the i^{th} SU holds the maximum value in the network, we have the algorithm (10) in distributed form as

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

because $0 < \alpha < \frac{1}{n}, n \ge N_i$ and $\delta_i > 1$, we have $0 < \frac{|N_i|\alpha}{\delta_i} < 1$, 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 $\|W\|_\infty \le 1$ which leads to $\|\tilde{W}\|_\infty \le \prod_{i=1}^{q} \|W_i\|_\infty \le 1$, where $q = \sum_{j=1}^{n-1} p_j$.

Meanwhile, from the above analysis we obtain

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

If x_{max} has a neighbor which holds non-maximum value, then we obtain

$$\begin{aligned} x_{max}(k+1) &< (1 - \alpha\frac{|N_i|}{\delta_i})x_{max}(k) + \frac{\alpha}{\delta_i}|N_i|x_{max}(k) \\ &< x_{max}(k) \end{aligned}$$

It means for each iteration of algorithm (10), if the x_{max} communicates with non-maximum nodes, x_{max} is strictly decreasing. Also, if x_{min} communicates with non-minimum nodes, x_{min} is strictly increasing.

For any state $x(k)$, assume there are l out of n nodes $(0 < l < n)$ holding maximum value and other $n-l$ nodes hold non-maximum values. Then, at least 1 out of l maximum nodes will strictly decrease when a jointly connected graph sequence happens from the k step to $k+1$ step. If no maximum nodes strictly decrease after a jointly connected graph sequence, then, it means none of the maximum nodes communicates with other non-maximum nodes and the graph sequence is not jointly connected. Therefore, for state $x(k)$, 1 jointly connected graph sequence happens from k step, at least 1 out of l maximum nodes decreases, and the network will have $l-1$ maximum nodes after. By induction, after $l-1$ jointly connected graph sequences happen, all the $l-1$ maximum nodes will strictly decreasing. For a network with n nodes, there are at most $n-1$ maximum nodes, and after $n-1$ jointly graph sequences, all the $n-1$ maximum nodes will strictly decrease. Same reason, after $n-1$ jointly connected graph sequences, all the minimum nodes will strictly increase. It means $\|\tilde{W}x\|_\infty < \|x\|_\infty$ for $x \notin span(1)$. Since all the $W_{s(k)}$ defined in (9) has $\mathbf{1}$ as right eigenvector with eignenvalue 1. Thus, we have $\tilde{W}$ is paracontracting according to the definition Eqn.(12). This finishes the proof of Lemma 4.

Under the condition that the collection of the jointly connected graphs occurs infinitely, we can write the matrix sequence as $\prod_{i=1}^{k} W_{s(k)} = \prod_{j=1}^{h} \tilde{W}_j$, that each $\tilde{W}_j$ represents n jointly connected graph sequences. Then $\tilde{W}_j$ is paracontracting and $\prod_{i=1}^{k} W_{s(k)}$ is a paracontracting sequence. Then according to Lemma 2, the iteration (10) will converge to its fixed subspace. Specifically, we can rewrite the iteration (10) as

$$x_c(k+1) = W_{s(k)} x_c(k) \tag{16}$$
$$x_d(k+1) = W_{s(k)} x_d(k) \tag{17}$$

where the initial value $x(0) = x_c(0) + x_d(0)$, $x_c(0) \in span(1)$ and $x_c(0)x_d(0)^T = 0$. Then we obtain $x_c(k+1) = x_c(k)$ and $\|x_d(k+1)\|_\infty < \|x_d(k)\|_\infty$. According to Lemma 2, when $k \to \infty$, $x_d(k) \to 0$ and $x(k) \to x_c(0)$. According to Lemma 3, the invariant subspace of $\prod W_{s(k)}$ is decided by the underlying strongly connected graph Perron matrix of each sub graph sequence. Let us denote as $W_p = \frac{1}{p}\sum_{i=1}^{p} W_i$ for each graph sequence $\{G_1, \ldots, G_p\}$ which contains $n-1$ jointly connected graph sequences. Since we have shown that in iteration (10), each W_p of a sub graph sequence, has the same simple eigenvalue 1 with same left eigenvector $\delta^T = [\delta_1, \ldots, \delta_n]$ and same right eigenvector $\mathbf{1}$. According to the Perron Frobenius Theorem [24], $x \to x_c(0) = \frac{\mathbf{1}\delta^T}{\delta^T\mathbf{1}} x(0) = \frac{\sum_{i=1}^{n}\delta_i x_i(0)}{\sum_{i=1}^{n}\delta_i}\mathbf{1}$, where $\delta = [\delta_1, \ldots, \delta_n]^T$ and δ_i is the element of the diagonal matrix Δ. This finishes the proof of Theorem 2 in the main file. □

4.2 Convergence Rate with Random Link Failures

For a SU network denoted as $\mathcal{G} = (\mathcal{E}, \mathcal{V})$, we assume $\mathcal{G}$ is a connected undirected graph and $\mathcal{E}$ is the set of realizable edges. We assign each pair of neighboring SUs the online and offline probabilities at each time step 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 $\mathcal{G}(k) = (E(k), \mathcal{V})$, where $E(k)$ denotes the edge set at time k.

Then the consensus iteration (10) becomes a random process and it is modeled as

$$x(k+1) = W(k)x(k) \tag{18}$$

where $W(k)$ is defined as

$$W(k) = I - \alpha\Delta^{-1}L(k) \tag{19}$$

where $\Delta = \mathrm{diag}\{\delta_1, \ldots, \delta_n\}$ 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 $\mathcal{G}(k)$ 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 5: If the SU network forms a connected undirected communication graph $\mathcal{G} = (\mathcal{E}, \mathcal{V})$, 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 stepsize α in Eqn. (19) satisfies the maximum node degree constraint $0 < \alpha < \frac{1}{d_{max}(\mathcal{G})}$, then the vector sequence $\{x(k)\}_{k=0}^{\infty}$ in (18) converges exponentially in the sense that

$$\lim_{k\to\infty} \|\mathrm{E}(x(k)) - x^*\mathbf{1}\|_2 = 0. \quad \forall x(0) \in \mathbb{R}^{n\times 1} \tag{20}$$

The decay factor of the convergence is given by $\rho(\overline{W} - J_1)$, where $0 < \rho(\overline{W} - J_1) < 1$ is the spectral radius of $\overline{W} - J_1$, $\overline{W} = E(W)$, and $J_1 = \frac{\mathbf{1}\delta^T}{\delta^T\mathbf{1}}$, where δ defined as

$$\delta = [\delta_1, \delta_2, \ldots, \delta_n]^T. \tag{21}$$

Proof: we have the error dynamics of the algorithm (18) as

$$\begin{aligned} x(k+1) - x^*\mathbf{1} &= W(k)x(k) - J_1 x(0) \\ &= \prod_{j=0}^{k} W(j)x(0) - J_1 x(0). \end{aligned} \tag{22}$$

Since δ and $\mathbf{1}$ are respectively the left and right eigenvetor of $W(k), \forall k \geq 0$, associated with the eigenvalue $\lambda_1 = 1$, we have $W(k)J_1 = J_1$ and $J_1 W(k) = J_1, \forall k \geq 0$, which yield

$$\begin{aligned} &x(k+1) - x^*\mathbf{1} \\ &= W(k)\prod_{j=0}^{k-1} W(j)x(0) - J_1 \prod_{j=0}^{k-1} W(j)x(0), \\ &= W(k)x(k) - J_1 x(k) = (W(k) - J_1)\, x(k) \end{aligned}$$

Since $W(k)J_1 = J_1$ and $J_1 J_1 = J_1$ we have

$$x(k+1) - x^*\mathbf{1} = (W(k) - J_1)(x(k) - J_1 x(0)) = \prod_{j=0}^{k} (W(k) - J_1)(x(0) - J_1 x(0))$$

Since all $L(k)$'s, and $W(k)$'s are independent and identically distributed, we have

$$\begin{aligned} \mathrm{E}(x(k+1) - x^*\mathbf{1}) &= \mathrm{E}\left(\prod_{j=0}^{k} (W(k) - J_1)\right) \mathrm{E}(x(0) - x^*\mathbf{1}) \\ &= \prod_{j=0}^{k} \mathrm{E}(W(k) - J_1)(\mathrm{E}(x(0)) - x^*\mathbf{1}) \\ &= \left(\overline{W} - J_1\right)^k (\mathrm{E}(x(0)) - x^*\mathbf{1}), \end{aligned}$$

which yields

$$\|\mathrm{E}(x(k+1)) - x^*\mathbf{1}\|_2 = \|\left(\overline{W} - J_1\right)^k\|_2 \|(\mathrm{E}(x(0)) - x^*\mathbf{1})\|_2. \quad (23)$$

For $\overline{W} = \mathrm{E}(W)$, we have

$$\overline{W} = I - \epsilon \Delta^{-1} L_p \quad (24)$$

where L_p is from the link probability, defined as

$$l_{p_{ij}} = \begin{cases} \sum_{j=1}^{n} P_{ij}, & \text{if } i = j, \\ -P_{ij} & \text{if } i \neq j, \text{ and } (v_i, v_j) \in \mathcal{E} \\ 0 & \text{otherwise} \end{cases} \quad (25)$$

We can see that L_p is still a Laplacian matrix of a connected graph with $\underline{P_{ij}}$ as its link weights. According to Lemma 1, we have $\rho(\overline{W} - J_1) < 1$, when α satisfies the maximum node degree constraint.

According to the famous Gelfand's formula, $\|(\overline{W} - J_1)^k\|_2$ has the same growth rate as $\rho(\overline{W} - J_1)^k$ as $k \to \infty$, which leads to the exponential convergence of Eqn. (20), and the decay factor is $\rho(\overline{W} - J_1)$. □

Remark 2: The decay factor for the convergence rate is the so-called spectral gap $\rho(\overline{W} - J_1)$ which relates to the 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 [19], [20], [25].

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 ϵ close to the final convergence value in the probability sense, which is called ϵ-convergence in [22].

Theorem 1: Under the same condition of Lemma 5, $\forall \epsilon > 0$ and $k \geq T(\epsilon)$, for the iteration (18), we have

$$\Pr\{\max_{1 \leq i \leq n} |x_i(k) - x^*| \geq \epsilon | \mathcal{H}_k\} \leq \epsilon, \quad k \in \{0, 1\} \quad (26)$$

and

$$T(\epsilon) \leq \frac{3/2 \log \epsilon^{-1} + 1/2 \log(K)}{1 - \mathrm{E}(\|(W - J_1)\|_\infty)} \quad (27)$$

where

$$K = \sum_{i=1}^{n} \left(2m\sigma_i^4 + 4E_s|h_i|^2\sigma_i^2 + (m\sigma_i^2 + E_s|h_i|^2)^2\right)$$

and $J_1 = \frac{\mathbf{1}\delta^T}{\delta^T \mathbf{1}}$, where δ defined in Eqn. (21), σ_i is the measurement noise variance for the i^{th} SU, and E_s is the signal energy and h_i is the channel gain defined in Section 2.1 of the main paper.

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

$$\begin{aligned} &\Pr\{\|x(k) - x^*\mathbf{1}\|_\infty \geq \epsilon\ \mathcal{H}_k\} \\ &= \Pr\{\|x(k) - x^*\mathbf{1}\|_\infty^2 \geq \epsilon | \mathcal{H}_k^2\} \\ &\leq \frac{\mathrm{E}\{\|x(k) - x^*\mathbf{1}\|_\infty^2 | \mathcal{H}_k^2\}}{\epsilon^2} \end{aligned} \quad (28)$$

where the second equation is from the Markov inequality. Following the proof of Theorem 5, from Eqn. (23), we have

$$\|x(k) - x^*\mathbf{1}\|_\infty^2 \leq \prod_{i=1}^{k} \|(W(k) - J_1)\|_\infty^{2k} \|x(0)\|_\infty^2. \quad (29)$$

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

$$\mathrm{E}(\|x(k) - x^*\mathbf{1}\|_\infty^2) \leq \mathrm{E}(\|W - J_1\|_\infty)^{2k} \mathrm{E}(\|x(0)\|_\infty^2) \quad (30)$$

If we choose a vector $\tilde{x}$ that $\|\tilde{x}\|_\infty = 1$ and $\delta^T \tilde{x} = 0$, where δ is defined in Eqn. (21), we have $J_1 \tilde{x} = 0$ and following Lemma 4, we have

$$\|(W(k) - J_1)\tilde{x}\|_\infty = \|W(k)\tilde{x}\|_\infty \leq \|\tilde{x}\|_\infty \quad (31)$$

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

$$\|(W(k) - J_1)\tilde{x}\|_\infty < \|\tilde{x}\|_\infty, \quad (32)$$

which means

$$\|(W(k) - J_1)\|_\infty = \max_{\|\tilde{x}\|_\infty = 1} \frac{\|(W(k) - J_1)\tilde{x}\|_\infty}{\|\tilde{x}\|_\infty} \leq 1 \quad (33)$$

and

$$\mathrm{E}(\|(W - J_1)\|_\infty) < 1, \quad (34)$$

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

$$\|x(0)\|_\infty^2 \leq \|x(0)\|_2^2. \quad (35)$$

Substitute (30) and (35) into (28), we have

$$\Pr\{\|x(k) - x^*\mathbf{1}\|_\infty \geq \epsilon | \mathcal{H}_k\} \leq \frac{\mathrm{E}(\|(W - J_1)\|_\infty)^{2k} \mathrm{E}\{\|x(0)\|_2^2\}}{\epsilon^2} \quad (36)$$

Let

$$\frac{\mathrm{E}(\|(W - J_1)\|_\infty)^{2k} \mathrm{E}\{\|x(0)\|_2\}}{\epsilon^2} = \epsilon, \quad (37)$$

we obtain

$$T(\epsilon) = \frac{3/2 \log \epsilon^{-1} + 1/2 \log(\mathrm{E}\{\|x(0)\|_2^2 | \mathcal{H}_k\})}{-\log(\mathrm{E}(\|(W - J_1)\|_\infty)} \quad (38)$$

Therefore, we have

$$T(\epsilon) = \frac{3/2 \log \epsilon^{-1} + 1/2 \log(\mathrm{E}\{\|x(0)\|_2^2 | \mathcal{H}_k\})}{-\log(\mathrm{E}(\|(W - J_1)\|_\infty)} \quad (39)$$

From the inequality $\log(1 + u) \leq u$ when u is small, let $1 + u = \mathrm{E}(\|(W - J_1)\|_\infty)$, we obtain,

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

Thus, we have

$$T(\epsilon) \leq \frac{3/2 \log \epsilon^{-1} + 1/2 \log(\mathrm{E}\{\|x(0)\|_2^2 | \mathcal{H}_k\})}{1 - \mathrm{E}\left(\|(W - J_1)\|_\infty\right)} \quad (40)$$

Meanwhile, according to Section 2.1 of the main paper, we have

$$\begin{aligned} &\mathrm{E}(\|x(0)\|_2^2 | \mathcal{H}_0) = \sum_{i=1}^{n} \mathrm{E}(x_i^2(0) | \mathcal{H}_0) \\ &< \sum_{i=1}^{n} \mathrm{E}(x_i^2(0)|\mathcal{H}_1) = \sum_{i=1}^{n} \left(\mathrm{Var}(x_i(0)|\mathcal{H}_1) + \mathrm{E}^2(x_i(0)|\mathcal{H}_1)\right) \\ &\leq \sum_{i=1}^{n} \left(2m\sigma_i^4 + 4E_s|h_i|^2\sigma_i^2 + (m\sigma_i^2 + E_s|h_i|^2)^2\right) \end{aligned}$$

where σ_i is the measurement noise variance for the i^{th} SU, and E_s is the signal energy and h_i is the channel gain. we obtain 27. □

Remark 3: ϵ-convergence of the average consensus or gossip algorithm has been extensively studied in [26] [12] [22]. Theorem 1 is a generalization to weighted average consensus convergence with random link failures. From (27), we can see clearly that the convergence rate of ϵ-convergence depends on the desired accuracy ϵ, the measurement channel noise variance σ_i, signal energy E_s, channel gain h_i and the expectation $\mathrm{E}\left(\|(W - J_1)\|_\infty\right)$.

Remark 4: Practically, $\mathrm{E}\left(\|(W - J_1)\|_\infty\right)$ is not easy to compute. Because the norm $\|\cdot\|$ is a convex function, we have

$$\mathrm{E}\left(\|(W - J_1)\|_\infty\right) \geq \|(\overline{W} - J_1)\|_\infty \geq \rho(\overline{W} - J_1), \quad (41)$$

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

5 Supplementary Simulation Results

5.1 P_d with respect to the algorithm convergence and PU transmission power

In order to characterize the convergence performance in terms of the detection probability P_d, Fig. 1(a) shows the trend of the P_d curves during the fusion process with respect to the consensus iteration step under fixed communication channels. We observe that the detection probability of 10 SU nodes converges to the same value 0.97 as the centralized Weighted Gain Combining (WGC) [27] approach within 35 steps. The false alarm is set at $P_f = 0.1$, the variance of Gaussian noise $\sigma_i = 1, \forall i$, and the channel SNR varies from 0 dB to -10 dB.

In this scenario, we evaluate the detection performance of the proposed scheme with respect to PU transmission power variation under the AWGN measurement channels. In Fig. 1(b), we compare the detection probability P_d of the proposed DWGC with existing EGC, OR and centralized WGC schemes. Under the AWGN channel condition, since the variance of Gaussian noise is fixed at $\sigma_i = 1$ for all i, the PU transmission power is directly reflected in the PU signal SNR in dB. From Fig. 1(b), we observe that DWGC always achieves the highest detection probability under both fixed and dynamic communication channels, and DWGC has comparable performance with centralized WGC. Particularly, when channel SNR is 0 dB, DWGC achieves detection probability of 0.77, which has 40% and 88% improvement over EGC and OR, respectively. When the PU transmission power becomes larger, such as close to 5 dB, the three approaches offer similar performance. The results validate that when the PU transmission power is low, the DWGC approach offers higher detection probability than EGC and OR approaches.

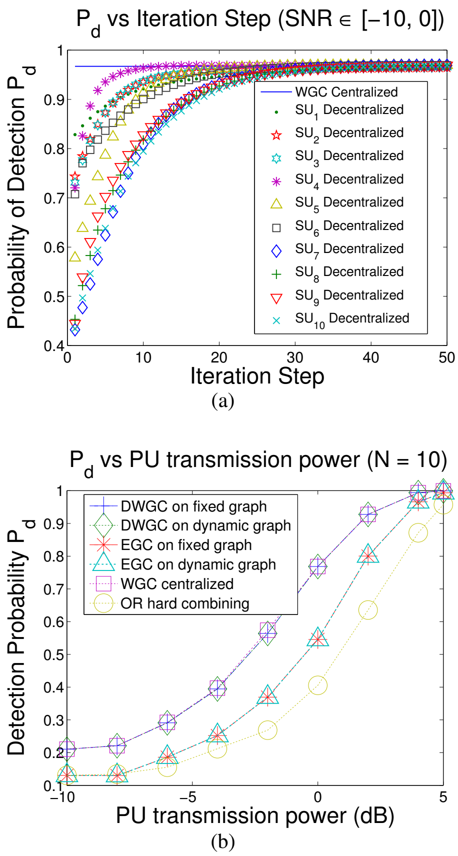
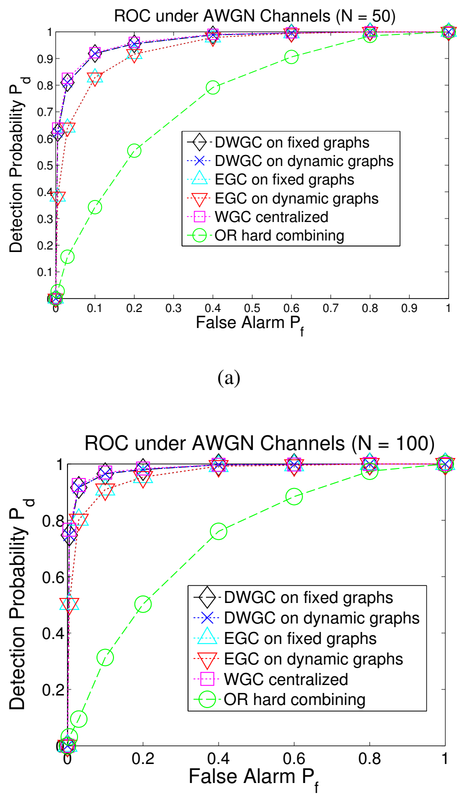


Fig. 1: (a) Detecting probability P_d with respect to iteration time step, fixed communication topology, $P_f = 0.1$. (b) Detection probability P_d with respect to PU signal SNR. $P_f = 0.1$, SUs are located 50m to 150m away from the PU.

5.2 Detection Performance with respect to Network Size

In Fig. 2, we study the ROC of the proposed DWGC, EGC, OR and centralized WGC approaches under AWGN channel with different SU network sizes. Particularly, the detection performance is evaluated under the network scenario with 50 and 100 nodes respectively, as shown in Fig. 2(a) and Fig. 2(b).

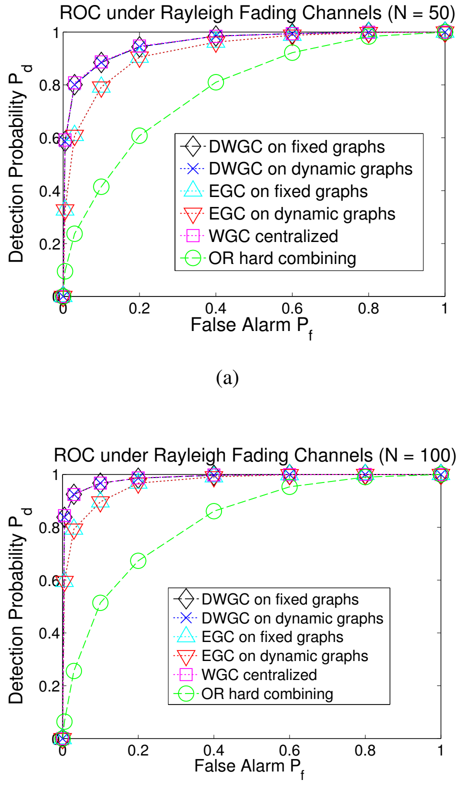
We observe that the proposed DWGC method always achieves the best performance with the centralized WGC approach under different network sizes. In particular, when the false alarm P_f is set to 0.1, the detection probability of DWGC method acheives over 0.9. Further, as the network



(a)

(b)

Fig. 2: ROC under AWGN channels with different network sizes. The channel SNR ranges from -5dB to -15dB. (a) ROC of a 50-node SU grid network. (b) ROC of a 100-node SU grid network.



(a)

(b)

Fig. 3: ROC under Rayleigh fading channels with different network sizes. The channel SNR ranges from -7dB to 3dB. (a) ROC of a 50-node SU grid network. (b) ROC of a 100-node SU grid network.

size increases, the detection probability also increases. For both 50-node and 100-node cases, the detection probability of DWGC has 10% and 50% improvement than that of EGC and OR methods, respectively. Moreover, we assume the variance of Gaussian noise is fixed at $\sigma_i = 1, \forall i$, and the channel SNR ranges from -5dB to -15dB, which is lower than the condition in Fig. 4 of the main paper with node size 10, 20 and 30. Such even harsher wireless environment further show the advantages of the proposed weighted design, which could achieve high detection probability as well as low false alarm rate, especially compared with distributed EGC and OR rule approaches.

Next, we examine the ROC curves for several different detection methods, including our proposed DWGC, EGC, OR and centralized WGC approaches under Rayleigh fading channel with different SU network sizes as shown in Fig. 3. Specifically, Fig. 3(a) and Fig. 3(b) provide the detectopm performance with the network including 50 and 100 nodes under the Rayleigh channel with identical channel conditions.

Similar as AWGN channel, we observe that DWGC method still achieves the best performance under different network sizes. In particular, when the false alarm P_f is set to 0.1, DWGC method has the detection probability above 0.9. We further find that the detection probability increases as the network size increases. For both 50-node and 100-node cases, the proposed DWGC method has 10% and 40% improvement on the detection probability than the EGC and OR methods. In addition, we also assume the variance of Gaussian noise is fixed at $\sigma_i = 1, \forall i$, and the channel SNR ranges from -7dB to 3dB. Therefore, for Rayleigh fading channel , the proposed DWGC approach has comparable performance with WGC and even outperforms EGC and OR rule in terms of detection probability under same false alarm constraints.

Overall speaking, when network size varies from 10 to 100 nodes, our proposed DWGC method achieves the best detection performance with the DWGC, and outperforms all the other existing spectrum sensing methods. Meanwhile, we also observe the detection performance of the weighted combining and equal gain combining become closer when the network size increases under the same measurement condition. The underlying reason is that the detection performance of cooperative spectrum sensing depends on the variety brought by the different weights from the SU network. As the network size increases, more SU nodes are involved in the cooperative spectrum sensing, which provides more reliability and robustness on the detection performance, especially for EGC method. Therefore, the detection performance for EGC method becomes better. In conclusion, DWGC outperforms EGC and OR rule and performs equivalently with WGC over the network sizes from 10 to 100 nodes. DWGC shows more benefits for relatively smaller SU networks.

References

[1] I. F. Akyildiz, B. F. Lo, and R. Balakrishnan, "Cooperative spectrum sensing in cognitive radio networks: A survey," *Physical Communication*, vol. 4, no. 1, pp. 40 – 62, 2011.
[2] D. Cabric, S. Mishra, and R. Brodersen, "Implementation issues in spectrum sensing for cognitive radios," *Conference Record of the Thirty-Eighth Asilomar Conference on Signals, Systems and Computers*, vol. 1, pp. 772 – 776, 2004.
[3] A. Ghasemi and E. Sousa, "Collaborative spectrum sensing for opportunistic access in fading environments," *First IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks*, pp. 131 –136, 2005.
[4] J. Ma and Y. Li, "Soft combination and detection for cooperative spectrum sensing in cognitive radio networks," *IEEE Global Telecommunications Conference*, pp. 3139 –3143, 2007.
[5] K. Ben Letaief and W. Zhang, "Cooperative communications for cognitive radio networks," *Proceedings of the IEEE*, vol. 97, pp. 878 –893, May 2009.
[6] W. Zhang and K. Letaief, "Cooperative spectrum sensing with transmit and relay diversity in cognitive radio networks," *IEEE Transactions on Wireless Communications*, vol. 7, no. 12, pp. 4761 –4766, 2008.
[7] S. Mishra, A. Sahai, and R. Brodersen, "Cooperative sensing among cognitive radios," in *IEEE International Conference on Communications*, vol. 4, pp. 1658 –1663, Istanbul, 2006.
[8] Y. Zou, Y.-D. Yao, and B. Zheng, "A cooperative sensing based cognitive relay transmission scheme without a dedicated sensing relay channel in cognitive radio networks," *Signal Processing, IEEE Transactions on*, vol. 59, no. 2, pp. 854–858, 2011.
[9] F. Yu, M. Huang, and H. Tang, "Biologically inspired consensus-based spectrum sensing in mobile ad hoc networks with cognitive radios," *IEEE Network*, vol. 24, no. 3, pp. 26 –30, 2010.
[10] Z. Li, F. R. Yu, and M. Huang, "A distributed consensus-based cooperative spectrum-sensing scheme in cognitive radio," *IEEE Transaction on Vehicular Technology*, vol. 59, no. 1, pp. 383–393, 2010.
[11] Z. Han, H. Li, D. Yang, and C. Pei, "Belief propagation based cooperative compressed spectrum sensing in wideband cognitive radio networks," *IEEE Transactions on Wireless Communications*, vol. 10, no. 9, pp. 3020 – 3031, 2011.
[12] M. E. Yildiz, T. C. Aysal, and K. E. Barner, "In-network cooperative spectrum sensing," in *EURASIP European Signal Processing Conference, Glasgow, Scotland*, 2009.
[13] J. Tsitsiklis, *Problems in Decentralized Decision Making and Computation*. PhD thesis, Dept. of Electrical Engineering and Computer Science, M.I.T., Boston, MA, 1984.
[14] A. Jadbabaie, J. Lin, and A. Morse, "Coordination of groups of mobile autonomous agents using nearest neighbor rules," *IEEE Transactions on Automatic Control*, 2003.
[15] R. Olfati-Saber and R. Murray, "Consensus problems in networks of agents with switching topology and time-delays," *IEEE Transactions on Automatic Control*, vol. 49, no. 9, pp. 1520 – 1533, 2004.
[16] L. Xiao, S. Boyd, and S. Lall, "A scheme for robust distributed sensor fusion based on average consensus," *Proc. 4th Int. Symp. Inf. Process. Sens. Netw.*, pp. 63–70, 2005.
[17] R. Olfati-Saber and J. Shamma, "Consensus filters for sensor networks and distributed sensor fusion," in *44th IEEE Conference on Decision and Control and 2005 European Control Conference*, pp. 6698 – 6703, 2005.
[18] S. Kar, S. Aldosari, and J. Moura, "Topology for distributed inference on graphs," *IEEE Transactions on Signal Processing*, 2008.
[19] S. Kar and J. Moura, "Sensor networks with random links: Topology design for distributed consensus," *IEEE Transactions on Signal Processing*, 2008.
[20] L. Xiao and S. Boyd, "Fast linear iterations for distributed averaging," *Systems & Control Letters*, vol. 53, pp. 65–78, 2004.
[21] R. Olfati-Saber, J. Fax, and R. Murray, "Consensus and cooperation in networked multi-agent systems," *Proceedings of the IEEE*, vol. 95, pp. 215 –233, jan. 2007.
[22] A. Dimakis, S. Kar, J. Moura, M. Rabbat, and A. Scaglione, "Gossip algorithms for distributed signal processing," *Proceedings of the IEEE*, 2010.
[23] L. Elsner, I. Koltracht, and M. Neumann, "On the convergence of asynchronous paracontractions with applications to tomographic reconstruction from incomplete data," *Linear Algebra Appl.*, vol. 130, pp. 65–82, 1990.
[24] M. Carl, *Matrix Analysis and Applied Linear Algebra*. SIAM, 2000.
[25] D. Jakovetic, J. Xavier, and J. M. F. Moura, "Consensus in correlated random topologies: Weights for finite time horizon," in *IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP)*, 2010.
[26] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah, "Randomized gossip algorithms," *IEEE Transactions on Information Theory*, 2006.
[27] Z. Quan, S. Cui, and A. Sayed, "Optimal linear cooperation for spectrum sensing in cognitive radio networks," *IEEE Journal of Selected Topics in Signal Processing*, 2008.