DISTRIBUTED GENERALIZED LIKELIHOOD RATIO TESTS : FUNDAMENTAL LIMITS AND TRADEOFFS

Anit Kumar Sahu and Soummya Kar

Department of Electrical and Computer Engineering,
Carnegie Mellon University, Pittsburgh PA 15213
{anits,soummyak}@andrew.cmu.edu

ABSTRACT

This paper focuses on the problem of distributed composite hypothesis testing in a network of sparsely interconnected agents, in which only a small section of the field modeling parametric alternatives is observable at each agent. A recursive generalized likelihood ratio test (GLRT) type algorithm in a distributed setup of the *consensus-plus-innovations* form is proposed, in which the agents update their parameter estimates and decision statistics by simultaneously processing the latest sensed information (*innovations*) and information obtained from neighboring agents (*consensus*). This paper characterizes the conditions and the testing algorithm design parameters which ensure that the probabilities of decision errors decay to zero asymptotically in the large sample limit. Finally, simulation studies are presented which illustrate the findings.

Index Terms— Distributed Inference, Consensus Algorithms, Generalized Likelihood Ratio Tests, Hypothesis Testing, Large Deviations Analysis.

1. INTRODUCTION

This paper revolves around testing a simple hypothesis against a composite alternative in a distributed multi-agent network. The hypotheses form a parametric family indexed by a (finite-dimensional vector) signal parameter, in which the null hypothesis corresponds to absence of signal, whereas, the collection of non-zero parameter values correspond to the (continuous) composite alternative. Broadly speaking, the objective is to simultaneously estimate the underlying parameter or state of the environment and decide which hypothesis is true based on the time-sequentially collected measurement data at the agents. This problem captures many practical applications including cooperative spectrum sensing [1] and MIMO radars [2]. The Generalized Likelihood Ratio Tests (GLRT) ([3]) algorithm is a classical approach that has been used widely in centralized setups for addressing such problems of composite testing. Apart from being inherently centralized, the GLRT is based on batch processing of observation data; further, due to the waiting time involved in obtaining a reasonably good estimate of the underlying parameter so as to ensure reasonable detection performance subsequently, its implementability in real-time applications may be limited. Moreover, the estimation and detection schemes running serially instead of in a parallel fashion consume a lot of sensing energy which may not go well with most multi-agent network scenarios that are typically energy constrained. Motivated by such constraints, we propose an algorithm $\mathcal{CIGLRT}$, a fully distributed recursive testing procedure, in which agents coordinate through local peer-to-peer information exchange and, in particular, the detection and estimation schemes run in parallel. Before elaborating further on the setup and the proposed distributed approach, we briefly review related existing work on distributed hypothesis testing in collaborative multi-agent networks. Distributed detection schemes as studied in the literature can be broadly categorized into three classes. Fusion center based architectures, where all the relevant information is transmitted to the fusion center by the agents and the subsequent inference schemes are operated by the fusion center (see, for example [4, 5]), constitutes the first class. Consensus schemes, which are distributed setups, where the data collection phase by the agents is followed by information exchange among them to reach a decision (see, for example [6, 7]) constitute the second class, whereas the third class consists of schemes which perform simultaneous assimilation of information obtained from sensing and communication in a recursive time-sequential manner (for example [8, 9]). The algorithm we present in this paper belongs to the third class, where agents make conditionally independent and temporally identically distributed (but possibly spatially heterogeneous) observations and update their parameter estimate and test statistic by simultaneous assimilation of the information obtained from the neighboring agents (*consensus*) and the latest locally sensed information (*innovation*). This justifies the name $\mathcal{CIGLRT}$ which is a distributed GLRT type algorithm of the *consensus* + *innovations* form. In this paper, so as to closely replicate typical practical sensing environments, we assume an agent's observations, say for agent n, is M_n dimensional, where $M_n << M$, M being the dimension of the underlying

This work was supported in part by NSF under grants CCF-1513936 and ECCS-1408222.

static vector parameter. We not only show the consistency of the parameter estimate sequence but also show the existence of feasible choice of thresholds and other algorithm design parameters which ensure that the probabilities of errors decay to zero asymptotically. (Fully) distributed detection schemes, in literature till now are concerned with either binary simple hypothesis testing (see, for example [8–10]) or multiple simple hypothesis testing (finite classification) (see, for example [11–13]) in contrast with the composite hypotheses with constant vector parameterization as studied in this paper. Furthermore, in [11–13] the observability condition assumed requires at least one agent to be able to distinguish between every possible pair of parameters, whereas, we require the weakest form of observability, i.e., the aggregate observation model is observable for the parameter of interest, referred to as global observability henceforth. The global observability requirement is necessary, even for a centralized procedure having access to all agent data at all times, for attaining consistent parameter estimates in the large sample limit. Addressing the fully composite testing setup with a continuous range of alternatives requires novel technical machinery in the form of development of analysis of efficient distributed estimation and detection procedures that interact in a closed loop fashion which we pursue in this paper. The rest of the paper is organized as follows. Spectral graph theory, preliminaries and notation are discussed next. The sensing model is described in Section 2, where we also review some preliminaries concerning the classical Generalized Likelihood Ratio Tests. Section 3 presents the proposed $\mathcal{CIGLRT}$ algorithm, while Section 4 concerns with the main results of the paper. The simulation results are stated in Section 5. Finally, Section 6 concludes the paper.

Spectral Graph Theory. The inter-agent communication network is a simple[1] undirected graph $G = (V, E)$, where V denotes the set of agents or vertices with cardinality $|V| = N$, and E the set of edges with $|E| = M$. If there exists an edge between agents i and j, then $(i, j) \in E$. A path between agents i and j of length m is a sequence ($i = p_0, p_1, \cdots, p_m = j$) of vertices, such that $(p_t, p_{t+1}) \in E$, $0 \le t \le m-1$. A graph is connected if there exists a path between all possible agent pairs. The neighborhood of an agent n is given by $\Omega_n = \{j \in V | (n, j) \in E\}$. The degree of agent n is given by $d_n = |\Omega_n|$. The structure of the graph is represented by the symmetric $N \times N$ adjacency matrix $\mathbf{A} = [A_{ij}]$, where $A_{ij} = 1$ if $(i, j) \in E$, and 0 otherwise. The degree matrix is given by the diagonal matrix $\mathbf{D} = diag(d_1 \cdots d_N)$. The graph Laplacian matrix is defined as $\mathbf{L} = \mathbf{D} - \mathbf{A}$. The Laplacian is a positive semidefinite matrix, hence its eigenvalues can be ordered and represented as $0 = \lambda_1(\mathbf{L}) \le \lambda_2(\mathbf{L}) \le \cdots \lambda_N(\mathbf{L})$. Furthermore, a graph is connected if and only if $\lambda_2(\mathbf{L}) > 0$ (see [14] for instance).

[1]A graph is said to be simple if it's devoid of self loops and multiple edges.

2. SENSING MODEL AND PRELIMINARIES

There are N agents deployed in the network. Every agent n at time index t makes a noisy observation $y_n(t)$, a noisy function of θ^*, which is a deterministic but unknown parameter and $\theta^* \in \mathcal{U} \subset \mathbb{R}^M$, where $\mathcal{U}$ is an open set in $\mathbb{R}^M$. Formally the observation model for the n-th agent is given by,

$$\mathbf{y}_n(t) = \mathbf{H}_n\theta^* + \gamma_n(t), \tag{1}$$

where $\{\mathbf{y}_n(t)\} \in \mathbb{R}^{M_n}$ is the observation sequence for the n-th agent and $\{\gamma_n(t)\}$ is a zero mean temporally i.i.d Gaussian noise sequence at the n-th agent with nonsingular covariance $\mathbf{\Sigma}_n$, where $\mathbf{\Sigma}_n \in \mathbb{R}^{M_n \times M_n}$. Furthermore, the noise processes at two different agents n, l for $n \neq l$ are independent. Motivated by most practical networked-agent applications, each agent only observes a subset of the components of θ^*, such that $M_n << M$. Under such a condition, in isolation, an agent can only estimate a part of the parameter, as the local sensing functions H_n's are not one-to-one on $\mathcal{U}$. However under appropriate network observability conditions and through inter-agent collaboration, it might be possible for each agent to get a consistent estimate of θ^*. Moreover, depending on as to which hypothesis is in force, the observation model is formalized as follows:

$$\begin{aligned} &\mathcal{H}_1 : \mathbf{y}_n(t) = \mathbf{H}_n\theta^* + \gamma_n(t) \\ &\mathcal{H}_0 : \mathbf{y}_n(t) = \gamma_n(t). \end{aligned} \tag{2}$$

We formalize the assumptions on the inter-agent communication graph and global observability.

Assumption B1. *We require the following global observability condition. The matrix* $\mathbf{G} = \sum_{n=1}^{N} \mathbf{H}_n^\top \mathbf{\Sigma}_n^{-1} \mathbf{H}_n$ *is full rank.*

Assumption B2. *The inter-agent communication graph, modeling the information exchange among the agents, is connected, i.e.* $\lambda_2(\mathbf{L}) > 0$*, where* $\mathbf{L}$ *denotes the associated graph Laplacian matrix.*

In order to motivate our distributed testing approach (presented in Section 3), we now review some concepts from Generalized Likelihood Ratio Testing. In a generalized target detection problem, let the absence of target be modeled by a simple hypothesis $\mathcal{H}_0$, whereas, its presence corresponds to a composite alternative $\mathcal{H}_1$ as the underlying parameter θ^* is unknown and can possibly attain a lot of values. Let $\mathbf{y}(t) = \left[\mathbf{y}_1(t)^\top \cdots \mathbf{y}_N(t)^\top\right]^\top$ represent the data from all the agents at time t, where $\mathbf{y}(t) \in \mathbb{R}^{\sum_{n=1}^{N} M_n}$. In a centralized setup, in which the fusion center has access to all the agents' observations i.e. $\mathbf{y}(t)$ at all times t, a classical testing approach is the generalized likelihood ratio test (GLRT). Formally, the GLRT decision rule is defined as follows:

$$\mathcal{H} = \begin{cases} \mathcal{H}_1, & \textit{if } \max_\theta \sum_{t=0}^{T} \log \frac{f_\theta(\mathbf{y}(t))}{f_0(\mathbf{y}(t)} > \eta, \\ \mathcal{H}_0, & \textit{otherwise}, \end{cases} \tag{3}$$

where η is a predefined threshold and

$$\begin{aligned} f_0(\mathbf{y}(t)) &= f_0^1(\mathbf{y}_1(t)) \cdots f_0^N(\mathbf{y}_N(t)) \\ f_\theta(\mathbf{y}(t)) &= f_\theta^1(\mathbf{y}_1(t)) \cdots f_\theta^N(\mathbf{y}_N(t)), \end{aligned} \tag{4}$$

which represent the likelihood of observing $\mathbf{y}$ under $\mathcal{H}_0$ and $\mathcal{H}_1$ respectively. Now, with the assumption that the observations made by the agents are conditionally independent, we have,

$$\max_{\theta}\sum_{t=0}^{T}\log\frac{f_{\theta}(\mathbf{y}(t))}{f_0(\mathbf{y}(\mathbf{t}))}=\max_{\theta}\sum_{t=0}^{T}\sum_{n=1}^{N}\log\frac{f_{\theta}^{n}(\mathbf{y}_n(t))}{f_0^{n}(\mathbf{y}_n(t))}. \quad (5)$$

The computation of the decision statistic in the maximization in (5) which uses all the data collected so far, is the key bottleneck in the implementation of the classical GLRT. In general, a maximizer of (5) is not known apriori as it depends on the raw data instance, and hence as far as communication complexity in the GLRT implementation is concerned, the maximization step incurs the major overhead- in fact, a direct implementation of the maximization (5) requires access to the entire raw data $\mathbf{y}$ at the fusion center.

To mitigate the communication complexity in realizing a fusion center having access to all the data, we present a distributed algorithm in which agents collaborate locally to obtain a maximizing θ. In order to obtain reasonable decision performance with such localized communication, we propose a distributed detector of the *consensus* + *innovations* type, which have been introduced in [15, 16]. In particular, each agent sequentially updates its parameter estimate and decision statistic in two parallelly running recursive schemes by assimilating the information obtained from its neighbors (*consensus potential*) and latest sensed local information (*innovation potential*).

3. $\mathcal{CIGLRT}$: ALGORITHM

In this section, we develop the algorithm $\mathcal{CIGLRT}$ for linear observation models. In Section 4 we state the main results concerning the characterization of the thresholds which ensure asymptotically decaying probabilities of errors. We skip the proofs due to space limitations. The proofs can be found in the longer manuscript ([17]).

Algorithm $\mathcal{CIGLRT}$

The algorithm $\mathcal{CIGLRT}$ consists of two parts, namely, the parameter estimate update and the decision statistic update.

Parameter Estimate Update. The algorithm $\mathcal{CIGLRT}$ generates the sequence $\{\theta_n(t)\}\in\mathbb{R}^M$ at the n-th agent according to the following recursive scheme

$$\theta_n(t+1)=\theta_n(t)-\underbrace{\beta_t\sum_{l\in\Omega_n}(\theta_n(t)-\theta_l(t))}_{\text{consensus}}$$
$$+\underbrace{\alpha_t\mathbf{H}_n^{\top}\boldsymbol{\Sigma}_n^{-1}\left(\mathbf{y}_n(t)-\mathbf{H}_n\theta_n(t)\right)}_{\text{innovation}}, \quad (6)$$

where Ω_n denotes the communication neighborhood of agent n, $\{\beta_t\}$ and $\{\alpha_t\}$ are consensus and innovation weight sequences respectively (to be specified shortly). The update in (6) can be written in a compact manner as follows:

$$\theta(t+1)=\theta(t)-\beta_t(\mathbf{L}\otimes\mathbf{I}_M)\theta(t)+\alpha_t\mathbf{G}_H\boldsymbol{\Sigma}^{-1}(\mathbf{y}(t)-\mathbf{G}_H^{\top}\theta(t)), \quad (7)$$

where $\theta(t)^{\top}=[\theta_1(t)^{\top}\cdots\theta_N(t)^{\top}]$, $\mathbf{G}_H=diag[\mathbf{H}_1^{\top}\cdots\mathbf{H}_N^{\top}]$, $\mathbf{y}(t)^{\top}=[y_1(t)^{\top}\cdots y_N(t)^{\top}]^{\top}$ and $\boldsymbol{\Sigma}=diag\left[\boldsymbol{\Sigma}_1,\cdots,\boldsymbol{\Sigma}_N\right]$.

We make the following assumptions on the weight sequences $\{\alpha_t\}$ and $\{\beta_t\}$.

Assumption B3. *The weight sequences $\{\alpha_t\}$ and $\{\beta_t\}$ are of the form $\alpha_t=(t+1)^{-1}, \beta_t=b(t+1)^{-\delta_2}$, where $b>0$ and $0<\delta_2<1/2$.*

Decision Statistic Update. The algorithm $\mathcal{CIGLRT}$ generates the decision statistic sequence $\{z_n(t)\}$ at the n-th agent according to the distributed recursive scheme

$$z_n(t+1)=\frac{t}{t+1}\left(z_n(t)-\underbrace{\delta\sum_{l\in\Omega_n}(z_n(t)-z_l(t))}_{\text{consensus}}\right)$$
$$+\underbrace{\frac{1}{t+1}\log\frac{f_{\theta_n(t)}(y_n(t))}{f_0(y_n(t))}}_{\text{innovation}}, \quad (8)$$

where $f_{\theta}(.)$ and $f_0(.)$ represent the likelihoods under $\mathcal{H}_1$ and $\mathcal{H}_0$ respectively and $\delta=\frac{2}{\lambda_2(\mathbf{L})+\lambda_N(\mathbf{L})}$. The decision statistic update in (8) can be written in a compact manner as follows:-

$$\mathbf{z}(t+1)=\frac{t}{t+1}(\mathbf{I}-\delta\mathbf{L})\mathbf{z}(t)$$
$$+\frac{1}{(t+1)}\mathbf{G}_{\theta}(t)\boldsymbol{\Sigma}^{-1}\left(\mathbf{y}(t)-\frac{\mathbf{G}_H^{\top}\theta(t)}{2}\right), \quad (9)$$

where $\mathbf{G}_{\theta}(t)=diag[\theta_1(t)^T\mathbf{H}_1^{\top},\cdots,\theta_N(t)^T\mathbf{H}_N^{\top}]$. It is to be noted that the entries of the weight matrix $\mathbf{W}=\mathbf{I}-\delta\mathbf{L}$ are designed in such a way that $\mathbf{W}$ is non-negative, symmetric, irreducible and stochastic, i.e., each row of $\mathbf{W}$ sums to one. Furthermore, the second largest eigenvalue in magnitude of $\mathbf{W}$, denoted by r, is strictly less than one (see [18]). Moreover, by the stochasticity of $\mathbf{W}$, the quantity r satisfies $r=||\mathbf{W}-\mathbf{J}||$, where $\mathbf{J}=\frac{1}{N}\mathbf{1}_N\mathbf{1}_N^{\top}$.

The following decision rule is adopted at all times t :

$$\mathcal{H}=\mathcal{H}_0 \textit{ if } z_n(t)\leq\eta, \mathcal{H}=\mathcal{H}_1 \textit{ otherwise}. \quad (10)$$

Under the aegis of such a decision rule, the associated probabilities of errors are as follows:

$$\mathbb{P}_{M,\theta^*}(t)=\mathbb{P}_{1,\theta^*}\left(z_n(t)\leq\eta\right), \mathbb{P}_{FA}(t)=\mathbb{P}_0\left(z_n(t)>\eta\right), \quad (11)$$

where $\mathbb{P}_{M,\theta^*}$ and $\mathbb{P}_{FA}$ refer to probability of miss and probability of false alarm respectively and $\mathbb{P}_{1,\theta^*}(.)$ and $\mathbb{P}_0(.)$ denote the probability when conditioned on hypothesis $\mathcal{H}_1$, which is in turn parameterized by θ^*, and the probability when conditioned on hypothesis $\mathcal{H}_0$ respectively.

4. $\mathcal{CIGLRT}$: MAIN RESULTS

In this section, we specifically characterize the thresholds for which the probability of miss and probability of false alarm decay to zero asymptotically. We also derive the large deviations exponent for the probability of false alarm.

Theorem 4.1. *Consider the $\mathcal{CIGLRT}$ algorithm under Assumptions B1-B3, and the sequence $\{\mathbf{z}(t)\}$ generated according to* (9). *We then have under $\mathbb{P}_{\theta^*}$, for all $\|\theta^*\|>0$,*

$$\sqrt{t+1}\left(z_n(t)-\frac{(\theta^*)^{\top}\mathbf{G}\theta^*}{2N}\right)\overset{\mathcal{D}}{\Longrightarrow}\mathcal{N}\left(0,\frac{(\theta^*)^{\top}\mathbf{G}\theta^*}{N^2}\right) \quad (12)$$

$\forall n$, where $\xRightarrow{\mathcal{D}}$ denotes convergence in distribution (weak convergence).

Theorem 4.1 asserts the asymptotic normality of the test statistic $\{z_n(t)\}$, $\forall n$ which in turn follows from the strong consistency of the parameter estimate sequence $\{\theta_n(t)\}$ which was studied in [19]. The next result concerns with the characterization of thresholds which ensures the probability of miss and probability of false alarm as defined in (11) decay to zero asymptotically.

Theorem 4.2. *Let the hypotheses of Theorem 4.1 hold. Consider the decision rule defined in* (10). *Then, we have for all θ^* that satisfy $\frac{(\theta^*)^\top \mathbf{G}\theta^*}{2N} > \frac{\left(\frac{1}{N}+\sqrt{N}r\right)\sum_{n=1}^{N} M_n}{2}$, the following choice of feasible thresholds*

$$\frac{\left(\frac{1}{N}+\sqrt{N}r\right)\sum_{n=1}^{N} M_n}{2} < \eta < \frac{(\theta^*)^\top \mathbf{G}\theta^*}{2N}, \tag{13}$$

ensures that $\mathbb{P}_{M,\theta^*}(t) \to 0$ and $\mathbb{P}_{FA}(t) \to 0$ as $t \to \infty$. Specifically, $\mathbb{P}_{FA}(t)$ decays to zero with the following large deviations exponent

$$\lim_{t\to\infty} \frac{1}{t} \log\left(\mathbb{P}_0\left(z_n(t) > \eta\right)\right) \leq -LE\left(\min\{\lambda^*, 1\}\right), \tag{14}$$

where $LE(\lambda) = \frac{\eta\lambda}{\frac{1}{N}+\sqrt{N}} + \left(\frac{\sum_{n=1}^{N} M_n}{2}\right)\log\left(1 - \frac{\lambda\left(\frac{1}{N}+\sqrt{N}r\right)}{\frac{1}{N}+\sqrt{N}}\right)$ *and* $\lambda^* = \frac{\frac{1}{N}+\sqrt{N}}{\frac{1}{N}+\sqrt{N}r} - \frac{\left(\frac{1}{N}+\sqrt{N}\right)\sum_{n=1}^{N} M_n}{2\eta}$.

Theorem 4.2 characterizes the range of θ^*'s for which a range of feasible thresholds exist that guarantee $\mathbb{P}_{M,\theta^*}(t)$, $\mathbb{P}_{FA}(t) \to 0$ as $t \to \infty$. The incorporation of inaccurate initial parameter estimates into the decision statistic, though sub-optimal, makes the detection scheme of $\mathcal{CIGLRT}$ a recursive *online* procedure, while the classical GLRT is an *offline* batch procedure as the corresponding parameter estimate used at any time instant depends on the entire raw data obtained at all agents so far and needs to be estimated first before computing the decision statistic. In spite of the sub-optimality in the update of the corresponding decision statistic, the algorithm $\mathcal{CIGLRT}$ ensures that the probabilities of errors decay to zero in the large sample (time) limit.

5. SIMULATIONS

We generate a planar random geometric network of 10 agents. We consider the underlying parameter to be scalar to give better intuition of our algorithm. Out of the 10 agents, 5 agents are defective i.e. observe only noise. The other 5 agents, observe noisy scaled versions of the underlying parameter with the scaling factors being $1, 1.5, 0.8, 2, 0.9$ for agents 1 to 5 respectively. We emphasize that the above design ensures global observability (in the sense of Assumption B1), however, the defective agents are locally unobservable for θ^*. The network is poorly connected which in turn is reflected by the quantity $r = \|\mathbf{W} - \mathbf{J}\|$ and is given by 0.9161[2]. The underlying parameter is considered to be $\theta^* = 7.8$, with the noise power being 3. In particular, for the parameter estimation algorithm, $b = 0.2$ and $\delta_2 = 0.1$, where b, δ_2 are as defined in Assumption B3. Figure 1 shows the convergence of the parameter estimates of the agents to the underlying parameter, which in turn demonstrates the consistency of the parameter estimation scheme. For the analysis of the probability of miss, we run the algorithm for 2000 sample paths. Figure 2 verifies the assertion in Theorem 4.2, with the probability of miss across the ideal agents and the defective agents going to zero. It is to be noted that, from Figure 2 the probability of miss starts decaying even before the parameter estimates get reasonably close to the true underlying parameter, which further indicates the *online* nature of the proposed algorithm $\mathcal{CIGLRT}$.

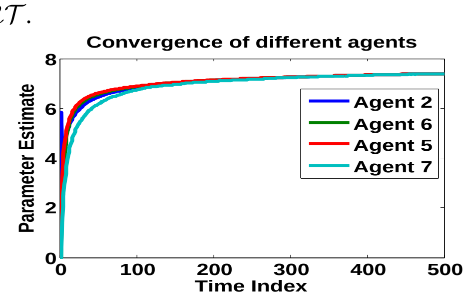

Fig. 1: Convergence analysis of the agents

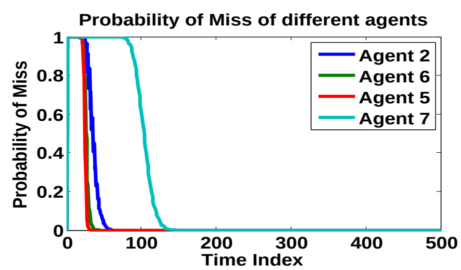

Fig. 2: Probability of miss of the agents

6. CONCLUSION

In this paper, we proposed a *consensus* + *innovations* type algorithm, $\mathcal{CIGLRT}$, in which every agent updates its parameter estimate and decision statistic by simultaneous processing of neighborhood information and local newly sensed information and where the inter-agent collaboration is restricted to a possibly sparse communication graph. Under rather generic assumptions, and a global observability criterion we establish the consistency of the parameter estimate sequence and characterize the feasible choice of thresholds which ensure that the probabilities of errors pertaining to the detection scheme decay to zero asymptotically. A natural direction for future research consists of considering models with non-linear observation functions and non-Gaussian noise.

7. REFERENCES

[1] S. Zarrin and T. J. Lim, "Composite hypothesis testing for cooperative spectrum sensing in cognitive radio,"

[2] Intuitively, smaller the value of r, better is the information flow in the network. For example, for $\mathbf{W} = \mathbf{J}$ i.e. in the all-to-all connected graph $r = 0$, while in case of no collaboration i.e. $\mathbf{W} = \mathbf{I}$, $r = 1$.

in *IEEE International Conference on Communications, 2009. ICC'09.* IEEE, 2009, pp. 1–5.

[2] A. Tajer, G. H. Jajamovich, X. Wang, and G. V. Moustakides, "Optimal joint target detection and parameter estimation by mimo radar," *IEEE Journal of Selected Topics in Signal Processing*, vol. 4, no. 1, pp. 127–145, 2010.

[3] O. Zeitouni, J. Ziv, and N. Merhav, "When is the generalized likelihood ratio test optimal?" *IEEE Transactions on Information Theory*, vol. 38, no. 5, pp. 1597–1602, 1992.

[4] R. S. Blum, S. A. Kassam, and H. V. Poor, "Distributed detection with multiple sensors i. advanced topics," *Proceedings of the IEEE*, vol. 85, no. 1, pp. 64–79, 1997.

[5] J. N. Tsitsiklis, "Decentralized detection," *Advances in Statistical Signal Processing*, vol. 2, no. 2, pp. 297–344, 1993.

[6] S. Kar and J. M. F. Moura, "Consensus based detection in sensor networks: Topology optimization under practical constraints," *Proc. 1st Intl. Wrkshp. Inform. Theory Sensor Networks*, 2007.

[7] R. Olfati-Saber, E. Franco, E. Frazzoli, and J. S. Shamma, "Belief consensus and distributed hypothesis testing in sensor networks," in *Networked Embedded Sensing and Control*. Springer, 2006, pp. 169–182.

[8] D. Bajovic, D. Jakovetic, J. Xavier, B. Sinopoli, and J. M. F. Moura, "Distributed detection via Gaussian running consensus: Large deviations asymptotic analysis," *IEEE Transactions on Signal Processing*, vol. 59, no. 9, pp. 4381–4396, 2011.

[9] S. Kar, R. Tandon, H. V. Poor, and S. Cui, "Distributed detection in noisy sensor networks," in *IEEE International Symposium on Information Theory Proceedings (ISIT), 2011*. IEEE, 2011, pp. 2856–2860.

[10] D. Jakovetic, J. M. F. Moura, and J. Xavier, "Distributed detection over noisy networks: Large deviations analysis," *IEEE Transactions on Signal Processing*, vol. 60, no. 8, pp. 4306–4320, 2012.

[11] A. Lalitha, A. Sarwate, and T. Javidi, "Social learning and distributed hypothesis testing," in *IEEE International Symposium on Information Theory (ISIT), 2014*. IEEE, 2014, pp. 551–555.

[12] A. Lalitha, T. Javidi, and A. Sarwate, "Social learning and distributed hypothesis testing," *arXiv preprint arXiv:1410.4307*, 2015.

[13] A. Jadbabaie, P. Molavi, A. Sandroni, and A. Tahbaz-Salehi, "Non-Bayesian social learning," *Games and Economic Behavior*, vol. 76, no. 1, pp. 210–225, 2012.

[14] F. R. Chung, *Spectral graph theory*. American Mathematical Soc., 1997, vol. 92.

[15] S. Kar and J. M. Moura, "Distributed linear parameter estimation in sensor networks: Convergence properties," in *Signals, Systems and Computers, 2008 42nd Asilomar Conference on*. IEEE, 2008, pp. 1347–1351.

[16] S. Kar, J. M. F. Moura, and K. Ramanan, "Distributed parameter estimation in sensor networks: nonlinear observation models and imperfect communication," *IEEE Transactions on Information Theory*, vol. 58, no. 6, pp. 3575 – 3605, June 2012.

[17] A. K. Sahu and S. Kar, "Recursive distributed detection for composite hypothesis testing:Algorithms and Asymptotics," *arXiv preprint arXiv:1601.04779*, 2016.

[18] A. G. Dimakis, S. Kar, J. M. F. Moura, M. G. Rabbat, and A. Scaglione, "Gossip algorithms for distributed signal processing," *Proceedings of the IEEE*, vol. 98, no. 11, pp. 1847–1864, 2010.

[19] S. Kar and J. M. F. Moura, "Convergence rate analysis of distributed gossip (linear parameter) estimation: Fundamental limits and tradeoffs," *IEEE Journal of Selected Topics in Signal Processing: Signal Processing in Gossiping Algorithms Design and Applications*, vol. 5, no. 4, pp. 674–690, August 2011.