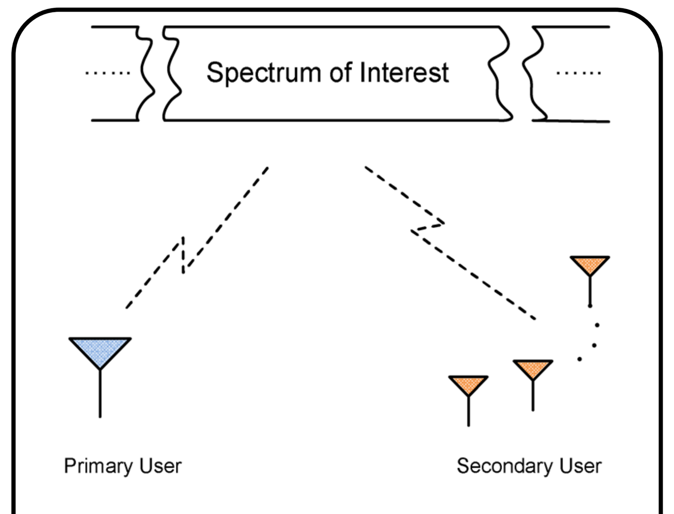
Distributed Generalized Likelihood Ratio Tests: Fundamental Limits and Tradeoffs

Anit Kumar Sahu and Soummya Kar

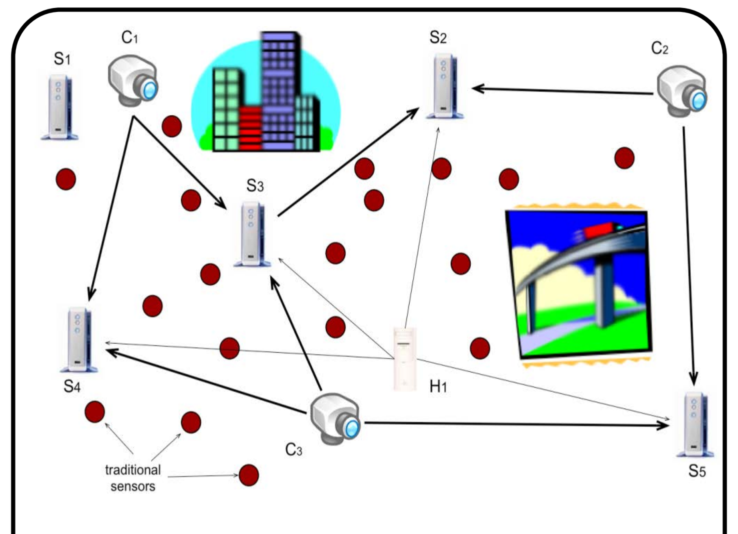
Carnegie Mellon University
Electrical & Computer ENGINEERING

Motivation and Setup

Motivation



Dynamic Spectrum Sensing



Sensor Surveillance

- Composite Hypothesis

$$\mathcal{H}_1 : \theta^* \neq 0$$
$$\mathcal{H}_0 : \theta^* = 0$$

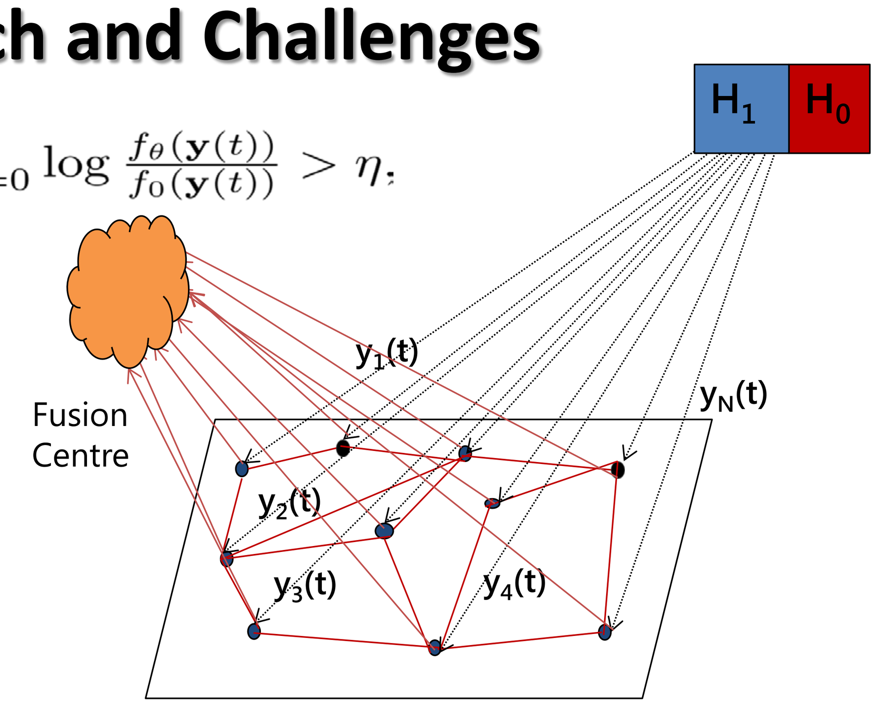
- θ^*: M-dimensional parameter, deterministic but unknown
- Network of N agents. Observation sequence at each agent

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

where $\mathbf{h}_n(.)$ is a non-linear function, $\mathbf{h}_n(0) = 0$, $\mathbf{h}_n(.) : \mathbf{R}^M \rightarrow \mathbf{R}^{M_n}$, $M_n << M$, $\gamma_n(t) \sim N(0, \mathbf{\Sigma}_n)$

Centralized Approach and Challenges

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



$$f_0(\mathbf{y}(t)) = f_0^1(\mathbf{y}_1(t)) \cdots f_0^N(\mathbf{y}_N(t)) : \text{likelihood of } \mathbf{y}(t) \text{ under } \mathcal{H}_0$$
$$f_\theta(\mathbf{y}(t)) = f_\theta^1(\mathbf{y}_1(t)) \cdots f_\theta^N(\mathbf{y}_N(t)) : \text{likelihood of } \mathbf{y}(t) \text{ under } \mathcal{H}_{1,\theta}$$

- Key bottleneck

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

- Depends on the entire sensed data.
- High-dimensional data exchange at all times.
- Involves batch processing: Detection cannot start until the maximizer is obtained

Distributed Approach

Assumptions

(A.1)-Global Observability The sensing model is globally observable, i.e., any two distinct values of θ and θ^* in the parameter space satisfy

$$\sum_{n=1}^{N} \|\mathbf{h}_n(\theta) - \mathbf{h}_n(\theta^*)\|^2 = 0$$

if and only if $\theta = \theta^*$.

(A.2)-Network Connectivity The inter-agent communication graph is connected, i.e., $\lambda_2(\mathbf{L}) > 0$, where $\mathbf{L}$ denotes the associated graph Laplacian matrix.

(A.3)-Smoothness For each agent n, $\forall \theta \neq \theta_1$, the sensing functions $\mathbf{h}_n$ are continuously differentiable and Lipschitz continuous with constants $k_n > 0$, i.e.,

$$\|\mathbf{h}_n(\theta) - \mathbf{h}_n(\theta_1)\| \leq k_n \|\theta - \theta_1\|.$$

(A.4)-Monotonicity For each pair of θ and $\acute{\theta}$ with $\theta \neq \acute{\theta}$, there exists a constant $c^* > 0$ such that the following aggregate strict monotonicity condition holds

$$\sum_{n=1}^{N} \left(\theta - \acute{\theta}\right)^\top (\nabla \mathbf{h}_n(\theta)) \mathbf{\Sigma}_n^{-1} \left(\mathbf{h}_n(\theta) - \mathbf{h}_n\left(\acute{\theta}\right)\right) \geq c^* \left\|\theta - \acute{\theta}\right\|^2.$$

CIGLRT: Consensus+Innovations GLRT Algorithm

Parameter Estimation Update

$$\theta_n(t+1) = \theta_n(t) - \underbrace{\beta_t \sum_{l \in \Omega_n} (\theta_n(t) - \theta_l(t))}_{\text{neighborhood consensus}} + \underbrace{\alpha_t \nabla \mathbf{h}_n(\theta_n(t)) \mathbf{\Sigma}_n^{-1} (\mathbf{y}_n(t) - \mathbf{h}_n(\theta_n(t)))}_{\text{local innovation}}$$

The weight sequences $\{\alpha_t\}_{t \geq 0}$ and $\{\beta_t\}_{t \geq 0}$ are given by

$$\alpha_t = \frac{1}{(t+1)} \;\; \beta_t = \frac{b}{(t+1)^{\tau_2}},$$

Decision Statistic Update

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

$$\delta \in \left(0, \frac{2}{\lambda_N(\mathbf{L})}\right),$$

Decision Rule

$$\mathcal{H}_n(t) = \begin{cases} \mathcal{H}_0 & z_n(t) \leq \eta \\ \mathcal{H}_1 & z_n(t) > \eta, \end{cases}$$

Probability of Errors

$$P_{M,\theta^*}(t) = P_{1,\theta^*}(z_n(t) \leq \eta)$$
$$P_{FA}(t) = P_0(z_n(t) > \eta),$$

Main Results

Asymptotic Normality and Error Analysis

Theorem 1: Asymptotic Normality Consider the $\mathcal{CIGLRT}$ algorithm under Assumptions A.1-A.4, and the sequence $\{\mathbf{z}(t)\}$. We then have under P_{θ^*}, for all $\|\theta^*\| > 0$,

$$\sqrt{t+1}\left(z_n(t) - \frac{\mathbf{h}^\top(\theta_N^*) \mathbf{\Sigma}^{-1} \mathbf{h}(\theta_N^*)}{2N}\right) \overset{\mathcal{D}}{\Longrightarrow} \mathcal{N}\left(0, \frac{\mathbf{h}^\top(\theta_N^*) \mathbf{\Sigma}^{-1} \mathbf{h}(\theta_N^*)}{N^2}\right), \forall n$$

where $\theta_N^* = \mathbf{1}_N \otimes \theta^*$, $\mathbf{h}(\theta_N^*) = \left[\mathbf{h}_1^\top(\theta^*) \cdots \mathbf{h}_N^\top(\theta^*)\right]^\top$ and $\overset{\mathcal{D}}{\Longrightarrow}$ refers to convergence in distribution (weak convergence).

Theorem 2: Error Analysis Consider the decision rule of $\mathcal{CIGLRT}$. For all θ^* which satisfy $\frac{\mathbf{h}^\top(\theta_N^*)\mathbf{\Sigma}^{-1}\mathbf{h}(\theta_N^*)}{2N} > \frac{\left(\frac{1}{N} + \sqrt{N}r\right)\sum_{n=1}^{N} M_n}{2}$, we have the following choice of the thresholds $\frac{\left(\frac{1}{N} + \sqrt{N}r\right)\sum_{n=1}^{N} M_n}{2} < \eta < \frac{\mathbf{h}^\top(\theta_N^*)\mathbf{\Sigma}^{-1}\mathbf{h}(\theta_N^*)}{2N}$ for which we have that $P_{M,\theta^*}(t) \rightarrow 0$ and $P_{FA}(t) \rightarrow 0$ as $t \rightarrow \infty$. Specifically, $P_{FA}(t)$ decays to zero exponentially with the following large deviations exponent

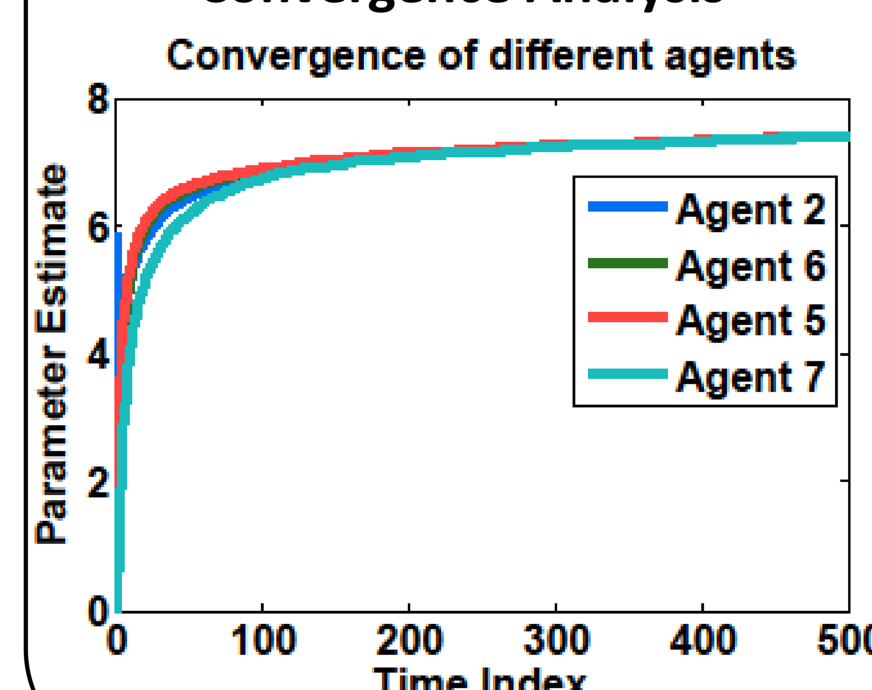
$$\limsup_{t \rightarrow \infty} \frac{1}{t} \log \left(P_0(z_n(t) > \eta)\right) \leq -LE\left(\min\{\lambda^*, 1\}\right),$$

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

Simulation Results

- Linear Scalar Model: $\mathbf{h}_n(\theta) = h_n\theta$.
- Network of 10 agents: 5 are defective, only observe noise.
- Scaling factors of the other 5 agents: 1, 1.5, 0.8, 2, 0.9
- $\theta^* = 7.8$, noise power $= 3$.
- Random geometric network: $\|\mathbf{I} - \delta\mathbf{L} - \frac{1}{N}\mathbf{1}\mathbf{1}^\top\| = 0.9161$.

Convergence Analysis



Probability of Miss