ROBUST SEQUENTIAL TESTING OF MULTIPLE HYPOTHESES IN DISTRIBUTED SENSOR NETWORKS

Mark R. Leonard, Maximilian Stiefel†, Michael Fauß*, and Abdelhak M. Zoubir**

*Signal Processing Group
Technische Universität Darmstadt
Darmstadt, Germany

†Continental Engineering Services
Frankfurt am Main, Germany

ABSTRACT

The problem of sequential multiple hypothesis testing in a distributed sensor network is considered and two algorithms are proposed: the Consensus + Innovations Matrix Sequential Probability Ratio Test ($\mathcal{CI}$MSPRT) for multiple simple hypotheses and the robust Least-Favorable-Density-$\mathcal{CI}$MSPRT for hypotheses with uncertainties in the corresponding distributions. Simulations are performed to verify and evaluate the performance of both algorithms under different network conditions and noise contaminations.

Index Terms— sequential detection, multiple hypothesis testing, distributed detection, robustness, distributional uncertainties

1. INTRODUCTION

In sequential detection, the goal is to make a reliable decision for one out of two or more hypotheses using as few measurements as possible. This is an important problem that emerges in modern real-time applications—especially in distributed setups—such as intelligent traffic control, smart homes, or video surveillance [1].

In this work, we are concerned with the design of sequential detectors for multiple hypotheses in a distributed sensor network. In order to be suitable for real-life applications, where the assumption of Gaussianity is often violated, the detectors should be insensitive to distributional uncertainties. The latter can, for example, be caused by outliers in the observations, insufficient knowledge about the observed phenomenon, or mismatches in the mathematical model. We propose the Consensus + Innovations Matrix Sequential Probability Ratio Test ($\mathcal{CI}$MSPRT) as a fusion of the centralized Matrix Sequential Probability Ratio Test (MSPRT) [1] for multiple hypotheses and the distributed Consensus + Innovations Sequential Probability Ratio Test ($\mathcal{CI}$SPRT) [2] for binary hypotheses. In a next step, we robustify this algorithm using the concept of least-favorable densities (LFDs) [3, 4].

The paper is structured as follows. In Section 2 we formulate the problem of multiple hypothesis testing in Gaussian and non-Gaussian environments. Section 3 reviews sequential binary detection in a distributed sensor network by recapitulating Wald's Sequential Probability Ratio Test (SPRT) as well as a generalized version of the $\mathcal{CI}$SPRT [3, 5]. Subsequently, we present the MSPRT as a solution for sequential multiple hypothesis testing and propose the $\mathcal{CI}$MSPRT as a fusion of the $\mathcal{CI}$SPRT and the MSPRT in Section 4. Furthermore, we give an approximation for its expected stopping time. In Section 5, we develop the robust LFD-$\mathcal{CI}$MSPRT. The simulations in Section 6 verify and evaluate the performance of our algorithms under different network and environmental conditions. Finally, conclusions are drawn in Section 7.

2. PROBLEM FORMULATION

Let $(X^1, \dots, X^n)$ be a sequence of independent and identically distributed random variables with common distribution P. Throughout the paper we assume that P admits a continuous density p. In a network of N agents, which can be modeled as an undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with the sets of agents $\mathcal{V}$ and edges $\mathcal{E}$, the closed neighborhood of agent k is given by $\mathcal{N}_k = \{l \in \mathcal{V} \mid (k, l) \in \mathcal{E}\} \cup \{k\}$. A simple way to define the neighborhood is by considering a radius d_{max} within which nodes can communicate with each other.

In a distributed test for multiple simple hypotheses, each agent needs to decide between $M > 1$ hypotheses of the form

$$\mathcal{H}_m \colon P = P_m, \quad m = 1, \dots, M.$$

We consider the case where each agent k should decide which of the M known signals of interest $x_m(t) \sim \mathcal{N}(0, \sigma_m^2)$ is active based on its measurement $y_k(t)$ at time instant t as well as information from its neighbors. The hypotheses become

$$\mathcal{H}_m \colon y_k(t) \sim \mathcal{N}(0, \sigma_m^2 + \sigma_n^2), \quad m = 1, \dots, M,$$

where σ_n^2 is the variance of a zero-mean white Gaussian noise process, which is assumed to be independent of $x_m(t)$.

In practice, however, the assumption of Gaussian measurement noise often does not hold. In order to take this element of uncertainty into account, the hypothesis test is transformed into a composite one between M disjoint sets of feasible distributions $\mathcal{P}_m$ so that the hypotheses become

$$\mathcal{H}_m\colon P \in \mathcal{P}_m, \quad m = 1, \dots, M.$$

The idea of robustness is to design a test that works reliably for all feasible distributions in the uncertainty sets $\mathcal{P}_m$. For binary fixed-sample-size tests, the minimax optimal solution to this problem is a likelihood ratio test between the LFDs [6,7]. Although strict minimax optimality does not carry over, we propose to use the pairwise LFDs to robustify the MSPRT and the $\mathcal{CI}$MSPRT. This is in line with the approach of the MSPRT to perform multiple pairwise tests in parallel, as we will see in Section 4.

3. REVIEW OF DISTRIBUTED SEQUENTIAL BINARY HYPOTHESIS TESTING

3.1. The Sequential Probability Ratio Test (SPRT)

In the single-sensor binary Sequential Probability Ratio Test (SPRT) proposed by Wald [8], a test statistic $S(t)$ is calculated at each time step t as

$$S(t) = \sum_{i=1}^{t} \log\left(\frac{p_1(y(i))}{p_0(y(i))}\right),$$

and compared to the lower threshold $\gamma^l = \log\frac{\beta}{1-\alpha}$ and the upper threshold $\gamma^u = \log\frac{1-\beta}{\alpha}$, where α and β denote prespecified bounds on the probabilities of false alarm and misdetection. As soon as one of the thresholds is crossed, the test stops and a decision is made according to the following rule:

$$\begin{aligned} &\text{if } S(t) \geq \gamma^u\colon\ \text{accept } \mathcal{H}_1 \\ &\text{if } S(t) \leq \gamma^l\colon\ \text{accept } \mathcal{H}_0. \end{aligned}$$

Otherwise, we continue sampling.

3.2. The Consensus + Innovations SPRT ($\mathcal{CI}$SPRT)

The Consensus + Innovations Sequential Probability Ratio Test ($\mathcal{CI}$SPRT) introduced in [2] extends the SPRT to distributed networks. In its original form, the $\mathcal{CI}$SPRT is formulated for a specific shift-in-mean test only, which is why we refer to our general formulation from [3] in this paper.

At each time instant t every node k computes its log-likelihoood ratio (LLR) $\eta_k(t)$ according to

$$\eta_k(t) = \log\left(\frac{p_1(y_k(t))}{p_0(y_k(t))}\right),$$

and updates its test statistic using neighbor information as

$$S_k(t) = \sum_{l\in\mathcal{N}_k} w_{kl} S_l(t-1) + \sum_{l\in\mathcal{N}_k} w_{kl}\eta_l(t),$$

with appropriate combination weights w_{kl} collected in matrix $\boldsymbol{W}$. For simplicity we choose equal combination weights, i.e.,

$$w_{kl} = \begin{cases} \frac{1}{|\mathcal{N}_k|} & \text{if } l \in \mathcal{N}_k \\ 0 & \text{otherwise.} \end{cases}$$

In analogy to the SPRT, $S_k(t)$ is compared to thresholds

$$\gamma^u_{\mathcal{CI}} \geq \frac{4}{7}\frac{c\,\sigma^2_{\eta,0}}{\mu_{\eta,0}}\left(\log\left(\frac{\alpha}{2}\right) + \log\left(1 - e^{-\frac{1}{2}\frac{\mu^2_{\eta,0}}{c\sigma^2_{\eta,0}}}\right)\right) \quad (1)$$

$$\gamma^l_{\mathcal{CI}} \leq \frac{4}{7}\frac{c\,\sigma^2_{\eta,1}}{\mu_{\eta,1}}\left(\log\left(\frac{\beta}{2}\right) + \log\left(1 - e^{-\frac{1}{2}\frac{\mu^2_{\eta,1}}{c\sigma^2_{\eta,1}}}\right)\right) \quad (2)$$

where $\mu_{\eta,0}, \sigma^2_{\eta,0}$ and $\mu_{\eta,1}, \sigma^2_{\eta,1}$ denote the mean and the variance of the LLR under $\mathcal{H}_0$ and $\mathcal{H}_1$, respectively. The constant c depends only on the network. More precisely, $c = r^2 + \frac{1}{N}$ where N is the network size and $r = \|\boldsymbol{W} - \frac{1}{N}\mathbf{1}\mathbf{1}^\top\|$ is the information flow. Here $\|\cdot\|$ denotes the Euclidean norm and $\mathbf{1}$ denotes the one-vector of length N. Each node makes a decision according to the rule

$$\begin{aligned} &\text{if } S_k(t) \geq \gamma^u_{\mathcal{CI}}\colon\ \text{accept } \mathcal{H}_1 \\ &\text{if } S_k(t) \leq \gamma^l_{\mathcal{CI}}\colon\ \text{accept } \mathcal{H}_0 \\ &\text{else}\colon\ \text{continue sampling.} \end{aligned}$$

4. SEQUENTIAL MULTIPLE HYPOTHESIS TESTING IN A DISTRIBUTED SENSOR NETWORK

4.1. The Matrix SPRT (MSPRT)

The Matrix SPRT (MSPRT) [1, Chapter 4] extends the single-sensor SPRT to multiple hypotheses. This is done by considering all possible hypothesis pairs $\mathcal{H}_m, \mathcal{H}_n$ with $m, n = 1, \dots, M$ and computing the pairwise test statistics

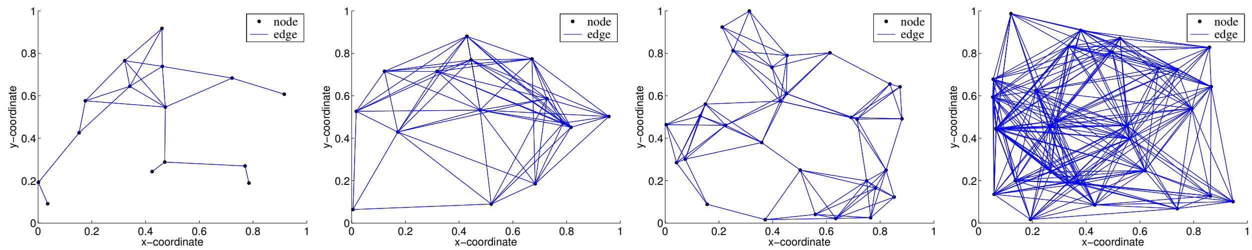
$$S_{mn}(t) = \sum_{i=1}^{t} \log\left(\frac{p_m(y(i))}{p_n(y(i))}\right).$$

Next, the test statistics are collected in a matrix $\boldsymbol{S}$, all entries of which are compared to a threshold matrix $\boldsymbol{\gamma}^u$ with entries

$$\gamma^u_{mn} = \log\left(\frac{1-\beta_{mn}}{\alpha_{mn}}\right) \approx \log\left(\frac{1}{\alpha_{mn}}\right),$$

where α_{mn} and β_{mn} denote bounds on the probabilities of false alarm and misdetection of the pairwise hypothesis test. We perform an acceptance test, i.e., the test is stopped and a decision is made in favor of $\mathcal{H}_m$ once all entries in the mth row of matrix $\boldsymbol{S}$—excluding the (m,m)th entry —exceed the corresponding thresholds. Formally, this can be written as

$$\begin{aligned} &\text{if } \exists\, m \in \{1,\dots,M\} \text{ such that} \\ &S_{mn}(t) \geq \gamma^u_{mn}\ \ \forall\, n \in \{1,\dots,M\}\setminus\{m\}\colon\ \text{accept } \mathcal{H}_m \\ &\text{else}\colon\ \text{continue sampling.} \end{aligned}$$



(a) $N = 15$, $d_{\max} = 0.3$. (b) $N = 15$, $d_{\max} = 0.6$. (c) $N = 30$, $d_{\max} = 0.3$. (d) $N = 30$, $d_{\max} = 0.6$.

Fig. 1. Sample networks obtained by randomly generating simple, connected and undirected graphs.

Hence, the multiple hypothesis test corresponds to performing $M(M-1)$ one-sided pairwise tests in parallel. Here the term one-sided is used to refer to tests with only an upper but no lower threshold. Note that it is also possible to perform a rejection test by inverting the LLRs and comparing to lower thresholds γ_{mn}^{l} instead.

4.2. The Consensus + Innovations MSPRT ($\mathcal{CI}$MSPRT)

In order to perform multiple hypothesis testing in a distributed sensor network, we fuse the concepts of the $\mathcal{CI}$SPRT and the MSPRT. In the proposed Consensus + Innovations Matrix Sequential Probability Ratio Test ($\mathcal{CI}$MSPRT) each node k first computes the LLRs for all hypothesis pairs $\mathcal{H}_m, \mathcal{H}_n$ as

$$\eta_{mn}^{k}(t) = \log\left(\frac{p_m(y_k(t))}{p_n(y_k(t))}\right). \tag{3}$$

Next, the LLRs are distributed over the neighborhood and the pairwise test statistics $S_{mn}^{k}(t)$ are calculated for all hypothesis pairs according to

$$S_{mn}^{k}(t) = \sum_{l\in\mathcal{N}_k} w_{kl} S_{mn}^{l}(t-1) + \sum_{l\in\mathcal{N}_k} w_{kl}\eta_{mn}^{l}(t). \tag{4}$$

Each node k performs an acceptance test according to the rule

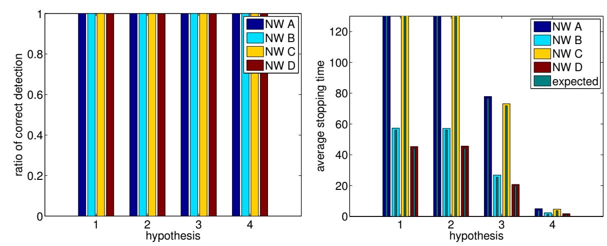
$$\begin{aligned}&\text{if } \exists\, m \in \{1,\dots,M\} \text{ such that}\\ &S_{mn}^{k}(t) \geq \gamma_{mn}^{u} \;\; \forall\, n \in \{1,\dots,M\}\setminus\{m\}: \text{ accept } \mathcal{H}_m\\ &\text{else}: \text{ continue sampling,}\end{aligned}$$

where γ_{mn}^{u} denotes the upper threshold for hypothesis pair $\mathcal{H}_m, \mathcal{H}_n$, which is calculated according to (2).

4.3. Expected Stopping Time of the $\mathcal{CI}$MSPRT

In this section we give an approximation to the expected stopping time of the $\mathcal{CI}$MSPRT. The stopping time T of the $\mathcal{CI}$MSPRT is defined as the first time instant t where a hypothesis is accepted, i.e.,

$$T = \inf\left\{t \mid \exists\, m \mid S_{mn}^{k}(t) \geq \gamma_{mn}^{u}\; \forall\, n \neq m\right\}.$$



(a) Ratio of correct detection (b) Average stopping time

Fig. 2. Simulation results for the $\mathcal{CI}$MSPRT.

Using Wald's identity [8] and following his derivations, the expected stopping time T_{mn} of a one-sided pairwise test between $\mathcal{H}_m$ and $\mathcal{H}_n$ under $\mathcal{H}_m$ can be approximated as

$$\mathbb{E}_m\left[T_{mn}\right] \approx \frac{\log\left(\gamma_{mn}^{u}\right)}{D(p_m|p_n)},$$

where $D(p_m|p_n)$ denotes the Kullback–Leibler divergence. Since the $\mathcal{CI}$MSPRT decides for $\mathcal{H}_m$ once the corresponding $M-1$ one-sided pairwise tests have stopped, its expected stopping time under $\mathcal{H}_m$ can be approximated by the expected stopping time of the slowest one-sided pairwise test, i.e.,

$$\mathbb{E}_m\left[T_m\right] \approx \max_{\substack{n=1,\dots,M\\ n\neq m}} \frac{\log\left(\gamma_{mn}^{u}\right)}{D(p_m|p_n)}. \tag{5}$$

5. ROBUSTIFYING THE $\mathcal{CI}$MSPRT

In order to robustify the $\mathcal{CI}$MSPRT against distributional uncertainties such as outliers, we resort to the concept of LFDs as proposed in [4], which we also used in our earlier work [3] to robustify the binary $\mathcal{CI}$SPRT. Assuming the measurements are ε-contaminated [9], i.e.,

$$p_m = (1-\varepsilon)p_m^0 + \varepsilon h_m$$

where $0 \leq \varepsilon < 0.5$ is the contamination factor, and p_m^0 and h_m denote the density of the nominal and the contaminating distribution under $\mathcal{H}_m$, respectively, we can find the fixed-sample-size LFDs for each hypothesis pair $\mathcal{H}_m, \mathcal{H}_n$ as

$$\begin{aligned} q_m &= \max\left\{c_m q_n\,,\ (1-\varepsilon)p_m^0\right\}, \\ q_n &= \max\left\{c_n q_m\,,\ (1-\varepsilon)p_n^0\right\}, \end{aligned} \tag{6}$$

for some $c_m, c_n > 0$. An efficient way for solving (6) can be found in [4, Table 1]. Replacing the nominal densities in (3) with the LFDs yields

$$\eta_{mn}^{k,\text{clipped}}(t) = \log\left(\frac{q_m(y_k(t))}{q_n(y_k(t))}\right).$$

This expression corresponds to Huber's clipped LLR [6, 7], which clips the LLR at certain levels to bound the influence of outliers. Hence, we can replace the pairwise LLR in (4) by its clipped counterpart and perform the matrix test at each node as in Section 4.2. We dub this algorithm LFD-$\mathcal{CI}$MSPRT.

6. SIMULATIONS

In this section, we analyze the performance of the proposed algorithms under different network sizes, $N \in \{15, 30\}$, and connectivity, $d_{\text{max}} \in \{0.3, 0.6\}$. We consider the four different networks depicted in Fig. 1 and perform a shift-in-variance test to decide which of the four zero-mean signals with variances $\sigma_m^2 \in \{1, 2, 4, 16\}$ is active. Furthermore, we fix the required probability of false alarm to $\alpha_{mn} = \alpha = 0.01$, and, in case of the LFD-$\mathcal{CI}$MSPRT, consider one type of contamination under all hypotheses, namely, $h_m = h = \mathcal{N}(0, 81)$. The contamination ratio ε is swept over the interval $[0, 0.3]$. For each hypothesis 1 000 Monte Carlo runs are performed. We consider the ratio of correct detection as well the average stopping time of the algorithm as performance metrics. In case of the $\mathcal{CI}$MSPRT we compare the average stopping time to the approximation (5).

The simulation results for the $\mathcal{CI}$MSPRT are shown in Fig. 2. Our algorithm clearly meets the required false-alarm probability of $\alpha = 0.01$. In all cases, the average stopping time deviates from the analytic approximation by at most two time instants, which means that the performance of the algorithm can be accurately predicted. Furthermore, the average stopping time is equal under $\mathcal{H}_1$ and $\mathcal{H}_2$, and drops as the signal variance increases under $\mathcal{H}_3$ and $\mathcal{H}_4$. This is due to the fact that, in a variance test, it is easier to match large values to their corresponding hypotheses. Finally, we see that a higher network connectivity can drastically reduce the average stopping time, while increasing the network size only has a marginal effect. This indicates a favorable scaling probability of the proposed algorithm.

The simulation results for the LFD-$\mathcal{CI}$MSPRT are depicted in Fig. 3. The LFD-$\mathcal{CI}$MSPRT delivers accurate detection results up to a contamination of 10 % irrespective of network size, connectivity, or underlying hypothesis. In the case where $\mathcal{H}_4$ is true, i.e., the signal with the largest variance

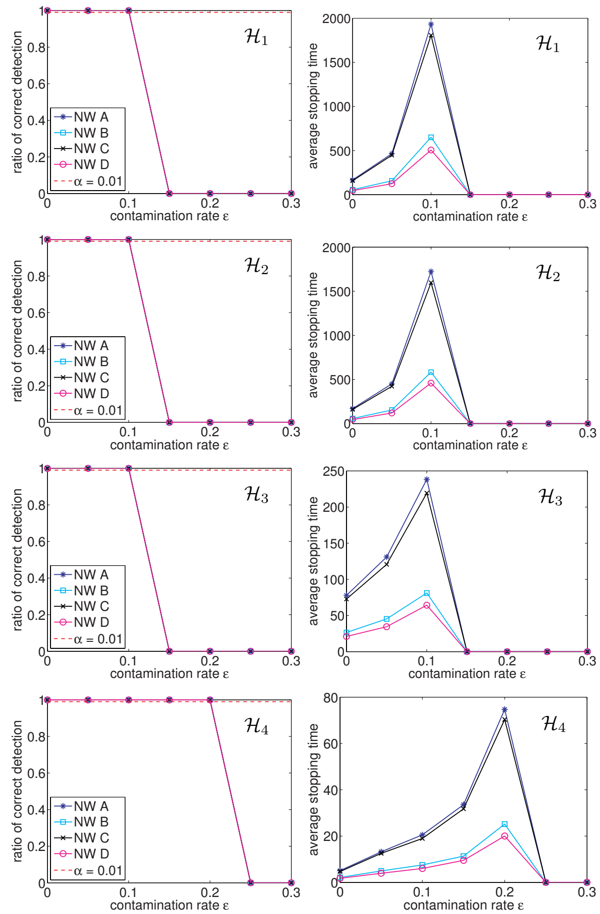


Fig. 3. Ratio of correct detection (left column) and average stopping time (right column) for the LFD-$\mathcal{CI}$MSPRT under $\mathcal{H}_1, \ldots, \mathcal{H}_4$ (from top to bottom).

is active, even 20 % outliers can be tolerated. Again, the impact of connectivity on the stopping time is visible, while the network size only has a marginal effect.

7. CONCLUSION

We proposed the $\mathcal{CI}$MSPRT and its robust extension—the LFD-$\mathcal{CI}$MSPRT—to perform sequential multiple hypothesis tests in distributed sensor networks. The algorithms are based on a fusion of the MSPRT and a generalized version of the $\mathcal{CI}$SPRT. Robustness is induced using LFDs. Our simulations show that in all considered scenarios the $\mathcal{CI}$MSPRT delivers accurate detection results—the LFD-$\mathcal{CI}$MSPRT even under 10-20 % contamination. Furthermore, we showed that the average stopping time can be predicted and is mainly controlled by the network connectivity so that the performance of our algorithms scales well with the network size.

8. REFERENCES

[1] A. Tartakovsky, I. Nikiforov, and M. Basseville, *Sequential Analysis: Hypothesis Testing and Changepoint Detection*, Chapman & Hall/CRC Monographs on Statistics & Applied Probability. Taylor & Francis, 2014.

[2] A. K. Sahu and S. Kar, "Distributed sequential detection for Gaussian shift-in-mean hypothesis testing," *IEEE Transactions on Signal Processing*, vol. 64, no. 1, pp. 89–103, Jan 2016.

[3] M. R. Leonard and A. M. Zoubir, "Robust distributed sequential hypothesis testing for detecting a random signal in non-Gaussian noise," in *Proceedings of the 25th European Signal Processing Conference (EUSIPCO)*, Aug. 2017.

[4] M. Fauß and A. M. Zoubir, "Old bands, new tracks – revisiting the band model for robust hypothesis testing," *IEEE Transactions on Signal Processing*, vol. 64, no. 22, pp. 5875–5886, Nov 2016.

[5] W. Hou, M. R. Leonard, and A. M. Zoubir, "Robust distributed sequential detection via robust estimation," in *Proceedings of the 25th European Signal Processing Conference (EUSIPCO)*, August 2017.

[6] P. J. Huber, "A robust version of the probability ratio test," *The Annals of Mathematical Statistics*, vol. 36, no. 6, pp. 1753–1758, 1965.

[7] P. J. Huber, *Robust Statistics*, Wiley, Hoboken, New Jersey, USA, 1981.

[8] A. Wald, "Sequential tests of statistical hypotheses," *The Annals of Mathematical Statistics*, vol. 16, no. 2, pp. 117–186, 1945.

[9] A. M. Zoubir, V. Koivunen, Y. Chakhchoukh, and M. Muma, "Robust estimation in signal processing: A tutorial-style treatment of fundamental concepts," *IEEE Signal Processing Magazine*, vol. 29, no. 4, pp. 61–80, July 2012.