Robust Distributed Sequential Detection via Robust Estimation

Wenbin Hou, Mark R. Leonard and Abdelhak M. Zoubir
Signal Processing Group, Technische Universität Darmstadt
Merckstraße 25, 64283 Darmstadt, Germany
{hou,leonard,zoubir}@spg.tu-darmstadt.de

***Abstract*—We study the problem of sequential binary hypothesis testing in a distributed multi-sensor network in non-Gaussian noise. To this end, we develop three robust extensions of the Consensus+Innovations Sequential Probability Ratio Test ($\mathcal{CI}$SPRT), namely, the Median-$\mathcal{CI}$SPRT, the M-$\mathcal{CI}$SPRT, and the Myriad-$\mathcal{CI}$SPRT, and validate their performance in a shift-in-mean as well as a change-in-variance test. Simulations show the superiority of the proposed algorithms over the alternative R-$\mathcal{CI}$SPRT.**

I. Introduction

This paper studies the problem of distributed sequential hypothesis testing in a multi-sensor network when the assumption of Gaussianity is violated. Each sensor node conducts a sequential hypothesis test based on its own observations and the information of its neighbors. The sequential test is terminated as soon as enough information has been collected to guarantee a certain level of estimation accuracy [1], so as to minimize the average run length in resource-limited and time-sensitive applications. We use a distributed architecture since it avoids the problem of a single point of failure, which is apparent in centralized networks with a fusion center [2]. In order to be able to handle non-Gaussian noise, we propose a robust version of the Consensus+Innovations Sequential Probability Ratio Test ($\mathcal{CI}$SPRT) introduced in [3],[4]. In contrast to the approach in [5], which uses the concept of least-favorable distributions to robustify the $\mathcal{CI}$SPRT, we robustify the innovations term of the update equation using a robust estimator.

The contribution of this paper is twofold. First, we develop three robust versions of the $\mathcal{CI}$SPRT algorithm, namely, the Median-$\mathcal{CI}$SPRT, the M-$\mathcal{CI}$SPRT, and the Myriad-$\mathcal{CI}$SPRT, by replacing the sample mean in the innovation term of the update equation with the respective robust estimator. Second, we evaluate the performance of the three robust algorithms and compare them to the R-$\mathcal{CI}$SPRT from [5] in terms of the average run length and the empirical probability of false alarm and misdetection.

The paper is structured as follows: In Section 2 we formulate the problem of shift-in-mean hypothesis testing in a distributed sensor network and in Section 3 we give a brief introduction of the $\mathcal{CI}$SPRT algorithm. In Section 4 we show how to robustify the $\mathcal{CI}$SPRT algorithm using a robust estimator and develop the Median-$\mathcal{CI}$SPRT, the M-$\mathcal{CI}$SPRT, and the Myriad-$\mathcal{CI}$SPRT, respectively. Section 5 is dedicated to simulations and conclusions are drawn in Section 6.

II. Problem Formulation

Consider a connected network with N sensors to decide between either of the simple hypotheses $\mathcal{H}_0$ and $\mathcal{H}_1$. The network can be modeled as an undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ and $\mathcal{E}$ denote the set of sensor nodes and the set of edges in the network, respectively. Let Ω_i, $i \in \{1, 2...N\}$, denote the set of neighbors of sensor i. For a shift-in-mean test, the null hypothesis and the alternative are formulated as

$$\begin{aligned}\mathcal{H}_0 &: y_i\left(t\right) = -\mu + g_i\left(t\right)\\ \mathcal{H}_1 &: y_i\left(t\right) = \mu + g_i\left(t\right)\end{aligned}$$

with $y_i(t)$ representing the measurement of node i at time instant t. Furthermore, μ and $-\mu$ denote the respective means of the signal and $g_i(t)$ is an independent and identically distributed, zero-mean white Gaussian noise process with variance σ_n^2. For a change-in-variance test, $\mathcal{H}_0$ and $\mathcal{H}_1$ are given by

$$\begin{aligned}\mathcal{H}_0 &: y_i\left(t\right) = g_i\left(t\right)\\ \mathcal{H}_1 &: y_i\left(t\right) = x(t) + g_i\left(t\right)\end{aligned}$$

where $x(t)$ is the zero-mean signal of interest with variance σ_x^2.

In most practical applications, the assumption of Gaussianity does not hold. Therefore, we want to design a test that is robust against deviations from this assumption. To this end, we will evaluate our algorithms in the face of measurement noise of the ϵ-contamination type [6], i.e.,

$$g_i(t) \sim G = (1-\epsilon)\, F + \epsilon\Delta,$$

where $F \sim \mathcal{N}(0, \sigma^2)$ denotes the nominal distribution, $\Delta \sim \mathcal{N}(0, \kappa\sigma^2)$ is the contaminating distribution with κ-times higher variance, and ϵ is the contamination coefficient with $0 \leq \epsilon \leq 0.5$.

III. The $\mathcal{CI}$SPRT Algorithm

The Consensus+Innovations Sequential Probability Ratio Test ($\mathcal{CI}$SPRT) introduced in [3],[4] is a distributed version

ISBN 978-0-9928626-7-1

of Wald's centralized SPRT [1]. Every node i in the network computes its test statistic $S_i(t)$ according to [3],[4]

$$S_i(t) = \underbrace{\left(w_{ii} S_i(t-1) + \sum_{j \epsilon \Omega_i} w_{ij} S_j\,(t-1) \right)}_{\text{consensus}} + \underbrace{\left(w_{ii} \eta_i(t) + \sum_{j \epsilon \Omega_i} w_{ij} \eta_j\,(t) \right)}_{\text{innovations}} \tag{1}$$

where w_{ij} denotes the entries of an appropriate weight matrix $\boldsymbol{W}$ that sum up to 1 and

$$\eta_i(t) = \log\left(\frac{f_1\,(y_i(t))}{f_0\,(y_i(t))}\right) \tag{2}$$

is the log-likelihood ratio of node i with $f_k\,(\cdot)$ denoting the probability density function under hypothesis $\mathcal{H}_k$.

The test statistic $S_i\,(t)$ is recursively updated over time until it crosses either one of the thresholds [5]

$$\gamma_u \geq \frac{4(m+1)}{7N} \frac{\sigma^2_{\eta,0}}{\mu_{\eta,0}} \left[\log\left(\frac{\alpha}{2}\right) + \log\left(1 - e^{-\frac{N}{2(m+1)} \frac{\mu^2_{\eta,0}}{\sigma^2_{\eta,0}}}\right) \right] \tag{3}$$

$$\gamma_l \leq \frac{4(m+1)}{7N} \frac{\sigma^2_{\eta,1}}{\mu_{\eta,1}} \left[\log\left(\frac{\beta}{2}\right) + \log\left(1 - e^{-\frac{N}{2(m+1)} \frac{\mu^2_{\eta,1}}{\sigma^2_{\eta,1}}}\right) \right], \tag{4}$$

that are derived based on the required false alarm and misdetection probabilities α and β, respectively. Here, $m = Nr^2$ and $r = \left\| \boldsymbol{W} - \frac{1}{N} \mathbf{1}\mathbf{1}^\top \right\|$ is the rate of information flow in the network, where $\|\cdot\|$ and $\mathbf{1}$ denote the Euclidean norm and the one-vector of length N, respectively. Furthermore, $\mu_{\eta,k}$ and $\sigma^2_{\eta,k}$ are the mean and the variance of the log-likelihood ratio under $\mathcal{H}_k$. When one of the thresholds is crossed, the test is stopped and a decision is made according to [3],[4]

$$\mathcal{H} = \begin{cases} \mathcal{H}_0, & S_i(T) \leq \gamma_l \\ \mathcal{H}_1, & S_i(T) \geq \gamma_u \end{cases}, \tag{5}$$

where T denotes the stopping time.

IV. Robustifying the $\mathcal{CI}$SPRT using Robust Estimators

In this section we show how to robustify the $\mathcal{CI}$SPRT against outliers using a robust estimator. Looking at (1), we observe that the innovations part is a weighted average of the log-likelihood ratios of the observations of node i and its neighbors. Hence, the update equation can be reformulated as

$$S_i(t) = \left(w_{ii} S_i(t-1) + \sum_{j \epsilon \Omega_i} w_{ij} S_j\,(t-1) \right) + \hat{\eta}(t), \tag{6}$$

with $\hat{\eta}(t)$ denoting the estimate of the innovations term at time t. By weighting the information of node i and its neighbors equally – a common choice when no a priori information on the reliability of each node is available – we obtain the sample mean

$$\hat{\eta}_{\text{mean}}(t) = \frac{1}{|\Omega_i \cup \{i\}|} \sum_{j \in \Omega_i \cup \{i\}} \eta_j(t), \tag{7}$$

which is a non-robust estimator [6]. Since the update equation is recursive, replacing the sample mean in the innovations part with a robust alternative, such as the median, the M, or the Myriad estimator, will robustify the consensus part as well and, thus, yield a test statistic that can handle outliers.

An advantage of introducing robustness by changing the combination rule instead of the log-likelihood ratio as proposed in [3],[5] is the fact that the thresholds and decision rules of the original $\mathcal{CI}$SPRT remain valid. In the following we will detail our approach for three different robust estimators.

A. The Median-$\mathcal{CI}$SPRT

The most straightforward way of replacing the sample mean in Equation (6) with a robust alternative is to use the median $\hat{\eta}_{\text{median}}(t)$. The estimate of the innovations term is calculated as

$$\hat{\eta}_{\text{median}}(t) = \text{median}\left(\boldsymbol{\eta}(t)\right), \tag{8}$$

with $\boldsymbol{\eta}(t) = \text{vec}\left(\{\eta_j(t)\}_{j \in \Omega_i \cup \{i\}}\right)$ denoting the vector of the log-likelihood ratios of node i and its neighbors.

B. The M-$\mathcal{CI}$SPRT

The M-$\mathcal{CI}$SPRT is obtained by estimating the innovations part of Equation (6) with an M-estimator. M-estimates can be intuitively understood as a weighted average with weights given by [6]

$$W(x) = \begin{cases} \frac{\psi(x)}{x}, & x \neq 0 \\ \psi'(0), & x = 0 \end{cases}, \tag{9}$$

where $\psi(x)$ is a score function and $\psi'(x)$ its first derivative. In this work we consider Huber's and Tukey's score functions that are defined as [7],[8]

$$\psi_{\text{Hub}}(x) = \begin{cases} x, & |x| \leq c_{\text{Hub}} \\ c_{\text{Hub}} \text{sign}(x), & |x| > c_{\text{Hub}} \end{cases}, \tag{10}$$

and

$$\psi_{\text{Tuk}}(x) = \begin{cases} x - 2\frac{x^3}{c^2_{\text{Tuk}}} + \frac{x^5}{c^4_{\text{Tuk}}}, & |x| \leq c_{\text{Tuk}} \\ 0, & |x| > c_{\text{Tuk}} \end{cases}, \tag{11}$$

respectively.

The M-estimate of the innovations term is obtained by recursively calculating [6], [9]

$$w_{kj} = W\left(\frac{\eta_j(t) - \hat{\eta}^{(k)}_{\text{M}}(t)}{\hat{\sigma}(\boldsymbol{\eta}(t))}\right) \tag{12}$$

$$\hat{\eta}^{(k+1)}_{\text{M}}(t) = \frac{\sum_{j \in \Omega_i \cup \{i\}} w_{kj} \eta_j(t)}{\sum_{j \in \Omega_i \cup \{i\}} w_{kj}} \tag{13}$$

until $\frac{|\hat{\eta}^{(k+1)}_{\text{M}}(t) - \hat{\eta}^{(k)}_{\text{M}}(t)|}{\hat{\sigma}(\boldsymbol{\eta}(t))} < \varepsilon$ for a very small, positive constant ε. The algorithm can be initialized by setting

ISBN 978-0-9928626-7-1 © EURASIP 2017

$\hat{\eta}_{\mathrm{M}}^{(0)}(t) = \hat{\eta}_{\mathrm{median}}(t)$ and estimating the scale using the normalized median standard deviation according to [6]

$$\hat{\sigma}_{\mathrm{mad}}(\boldsymbol{\eta}(t)) = 1.483 \cdot \mathrm{median}\left(|\boldsymbol{\eta}(t) - \hat{\eta}_{\mathrm{median}}(t)|\right). \quad (14)$$

C. The Myriad-$\mathcal{CI}$SPRT

The third robust estimator we consider in this work is the myriad, which estimates the innovations term according to [10],[11]

$$\hat{\eta}_{\mathrm{myriad}}(t) = \arg\min_{\eta} \prod_{j \in \Omega_i \cup \{i\}} \left[k^2 + (\eta_j(t) - \eta)^2\right] \quad (15)$$

where k is a freely tunable parameter. A common choice is to set $k = \hat{\sigma}_{\mathrm{mad}}(\boldsymbol{\eta}(t))$ [11].

V. Simulations

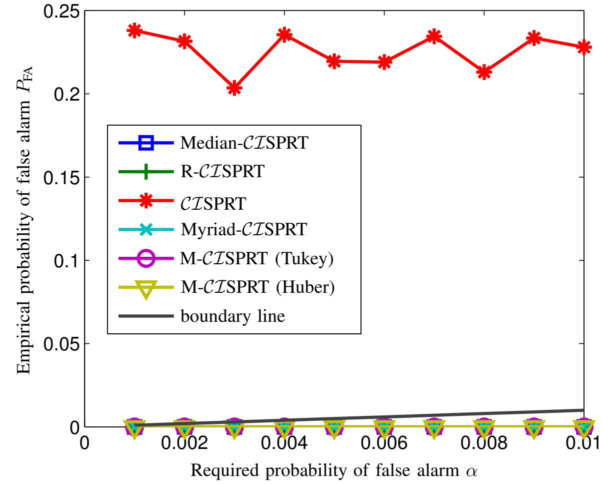
In the following we evaluate the standard $\mathcal{CI}$SPRT, our three proposed robust algorithms, namely, the Median-$\mathcal{CI}$SPRT, the M-$\mathcal{CI}$SPRT, and the Myriad-$\mathcal{CI}$SPRT, as well as the R-$\mathcal{CI}$SPRT from [5] in a shift-in-mean as well as a change-in-variance test. We compare the algorithms in terms of the empirical probabilities of false alarm P_{FA} and misdetection P_{MD}, respectively, as well as the average run length of the sequential test.

A. Simulation Setup

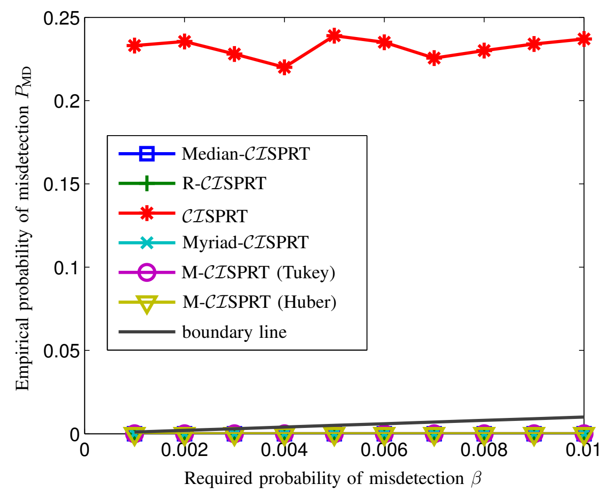
We generate a connected network with $N = 30$ sensors. The x- and y-coordinates of each node are sampled from a uniform distribution in the interval of $(0, 1)$. We set the maximal distance between two connected sensors to $g = 0.6$, meaning that only nodes within this range can communicate with each other and are considered neighbors. Furthermore, for the shift-in-mean test we set $\sigma_n^2 = 1$ and $\mu = 1$, to test between two signals with a mean of -1 and 1, respectively, and choose a noise contamination of $\epsilon = 0.3$. For the change-in-variance test, $\sigma_n^2 = 1$, $\sigma_x^2 = 8$, $\epsilon = 0.1$. The tuning constants of the M-estimator are chosen as $c_{\mathrm{Huber}} = 1.5$ and $c_{\mathrm{Tukey}} = 2.5$. We assume the required probabilities of false alarm α and misdetection β to be equal and ranging from 10^{-3} to 10^{-2}. We conduct 10,000 Monte Carlo runs under each hypothesis.

B. Simulation results

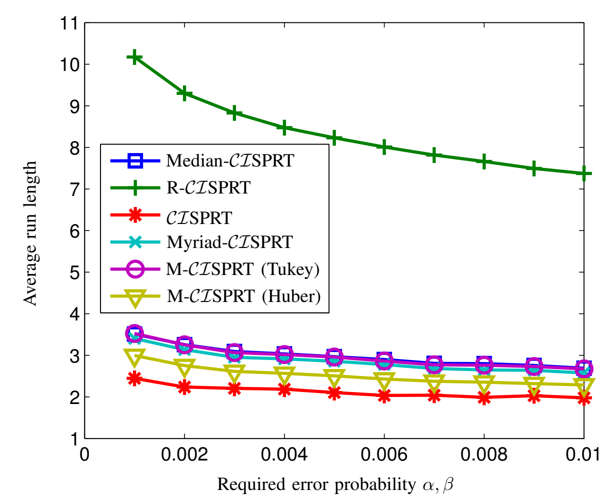
1) Shift-in-mean test: The simulation results for the shift-in-mean test are depicted in Figure 1. Figures 1(a) and 1(b) show the empirical probability of false alarm P_{FA} and misdetection P_{MD}, respectively, over the required probabilities α and β. We observe that all robust algorithms meet the required error probabilities with $P_{\mathrm{FA}} = P_{\mathrm{MD}} = 0$, while the standard $\mathcal{CI}$SPRT, which uses a sample mean to estimate the innovations term, is non-robust against outliers and therefore unable to meet the requirements. In contrast, all robust algorithms meet and even exceed the required error probabilities with $P_{\mathrm{FA}} = P_{\mathrm{MD}} = 0$. Note that the results for both error probabilities are almost identical. This is due to the fact that the shift-in-mean test is a symmetric problem and, hence, outliers affect both hypotheses equally.



(a) Empirical probability of false alarm P_{FA}



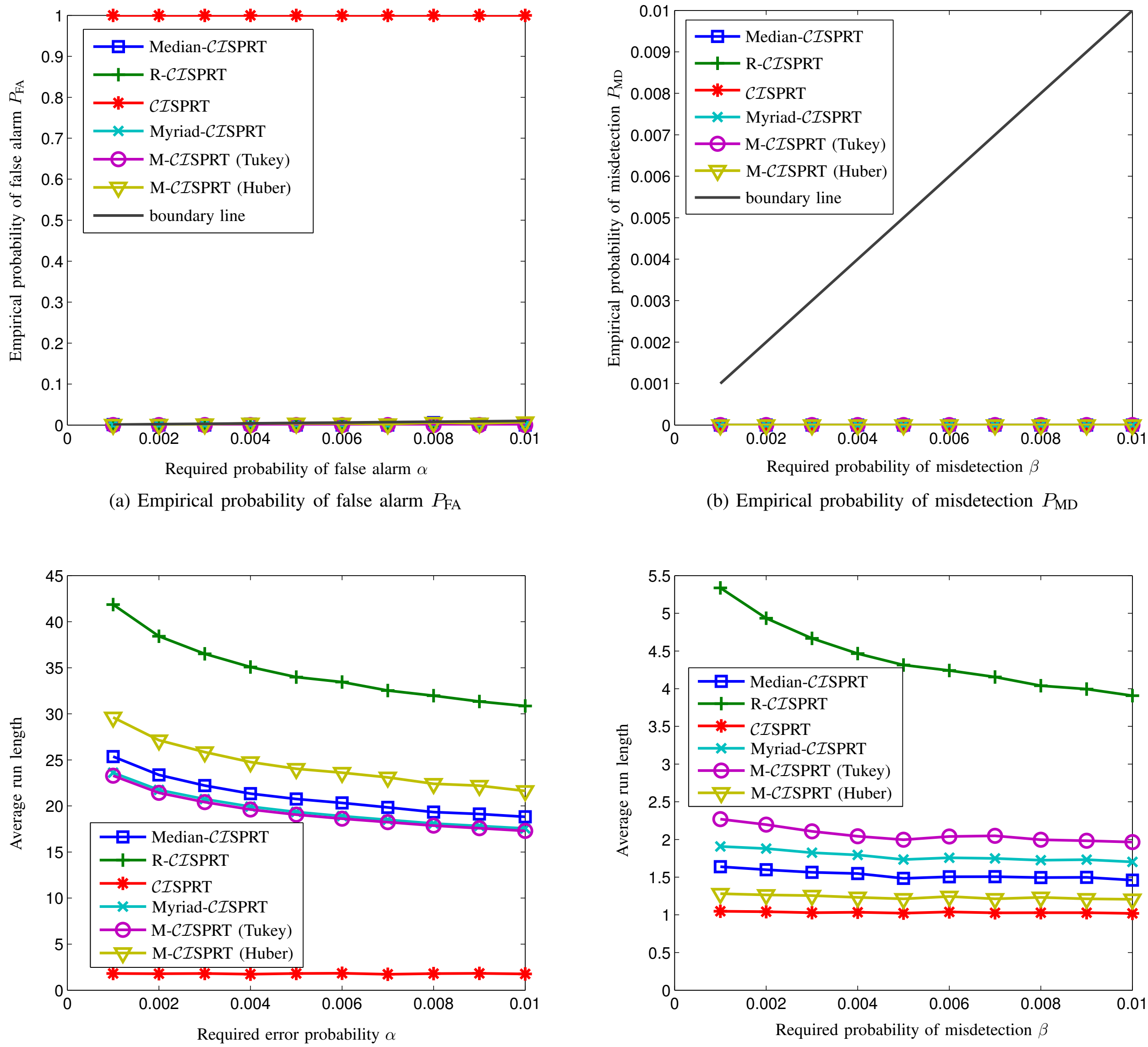
(b) Empirical probability of misdetection P_{MD}



(c) Average run length of the sequential test under both $\mathcal{H}_0$ and $\mathcal{H}_1$

Fig. 1: Simulation results of the shift-in-mean test using the $\mathcal{CI}$SPRT, Median-$\mathcal{CI}$SPRT, M-$\mathcal{CI}$SPRT, Myriad-$\mathcal{CI}$SPRT, and R-$\mathcal{CI}$SPRT.

ISBN 978-0-9928626-7-1



(a) Empirical probability of false alarm P_{FA}

(b) Empirical probability of misdetection P_{MD}

(c) Average run length under $\mathcal{H}_0$

(d) Average run length under $\mathcal{H}_1$

Fig. 2: Simulation results of the change-in-variance test using the $\mathcal{CI}$SPRT, the Median-$\mathcal{CI}$SPRT, M-$\mathcal{CI}$SPRT, Myriad-$\mathcal{CI}$SPRT, and the R-$\mathcal{CI}$SPRT.

Figure 1(c) shows the average run length of the sequential test over the required error probabilities α and β. The result is identical for both hypotheses. We observe that the standard $\mathcal{CI}$SPRT is the fastest algorithm. However, since it is not able to meet the required error probabilities, it is of no use in a non-Gaussian environment. The three proposed robust algorithms, however, are robust against outliers and only slightly slower than the $\mathcal{CI}$SPRT. Clearly, the best performing algorithm in the shift-in-mean test is the M-$\mathcal{CI}$SPRT with Huber's score function, which only needs half a time step more on average than the non-robust $\mathcal{CI}$SPRT.

While the R-$\mathcal{CI}$SPRT introduced in [5] is also robust against outliers, it is considerably slower than either of the proposed three algorithms. This is most likely due to the fact that this approach is derived using the least-favorable distributions, i.e., considering the worst-case for the scenario under test. Since we sample from a non-Gaussian distribution that is not necessarily the least-favorable one, it makes sense that the proposed algorithms, which try to remove or downweight occuring outliers, outperform the approach that is optimized for the worst case.

2) Change-in-variance test: The simulation results for the change-in-variance test are depicted in Figure 2. Figures 2(a) and 2(b) show the empirical probability of false alarm P_{FA}

ISBN 978-0-9928626-7-1

and P_{MD}, respectively. In contrast to the shift-in-mean test, the results are quite different. While all robust algorithms again meet and even exceed the required error probabilities, the $\mathcal{CI}$SPRT fails completely under $\mathcal{H}_0$ with $P_{\text{FA}} = 1$. However, it works under the alternative with $P_{\text{MD}} = 0$. This is due to the fact that the change-in-variance test is an asymmetric problem, meaning that outliers have a different effect on each hypothesis. When testing between a small and a large variance of an incoming signal, outliers, i.e., extremely large values, will only have a bad effect under the null hypothesis. Under $\mathcal{H}_1$ they will even be beneficial in correctly accepting the alternative. Hence, no detection will be missed, but false alarms are unavoidable when relying on a non-robust detector.

For the same reason, the average run length shown in Figures 2(c) and 2(d) is different under $\mathcal{H}_0$ and $\mathcal{H}_1$. While the non-robust $\mathcal{CI}$SPRT always finishes the test after approximately 1 sample, the robust algorithms need a considerably larger amount of samples to make a decision under the null hypothesis than under the alternative since the appearance of outliers makes the test easier in the latter case. Under $\mathcal{H}_1$, the M-$\mathcal{CI}$SPRT with Huber's score function is again the best performing algorithm, which makes sense, since this side of the asymmetric problem is similar to the shift-in-mean test in terms of the effect of outliers. Under $\mathcal{H}_0$, however, both the M-$\mathcal{CI}$SPRT with Tukey's score function and the Myriad-$\mathcal{CI}$SPRT prevail, and the Median-$\mathcal{CI}$SPRT comes in second.

Similar to the shift-in-mean test, the R-$\mathcal{CI}$SPRT – while robust – is considerably slower than our proposed algorithms since it is optimized for the worst-case scenario.

VI. Conclusion

In this paper we proposed an approach for robustifying the $\mathcal{CI}$SPRT via robust estimators. To this end we developed three robust distributed sequential detectors, namely, the Median-$\mathcal{CI}$SPRT, the M-$\mathcal{CI}$SPRT, and the Myriad-$\mathcal{CI}$SPRT. We evaluated their performance in a shift-in-mean as well as a change-in-variance test in terms of the average run length and the empirical error probabilities. Furthermore, we showed that they outperform the R-$\mathcal{CI}$SPRT, which is based on the least-favorable distributions.

References

[1] A. Wald, *Sequential Analysis*. New York City, New York, USA: Wiley, 1947.

[2] J. F. Chamberland and V. V. Veeravalli, "Decentralized detection in sensor networks," *IEEE Transactions on Signal Processing*, vol. 51, no. 2, pp. 407–416, Feb 2003.

[3] A. K. Sahu and S. Kar, "Distributed sequential detection for Gaussian binary hypothesis testing: Heterogeneous networks," in *48th Asilomar Conference on Signals, Systems and Computers*, Nov 2014, pp. 723–727.

[4] ——, "Distributed sequential detection for Gaussian shift-in-mean hypothesis testing," *IEEE Transactions on Signal Processing*, vol. 64, no. 1, pp. 89–103, Jan 2016.

[5] 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)*, Sep 2017, submitted.

[6] 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.

[7] L. Denby and W. A. Larsen, "Robust regression estimators compared via Monte Carlo," *Communications in Statistics - Theory and Methods*, vol. 6, no. 4, pp. 335–362, 1977.

[8] G. Shevlyakov, S. Morgenthaler, and A. Shurygin, "Redescending M-estimators," *Journal of Statistical Planning and Inference*, vol. 138, no. 10, pp. 2906 – 2917, 2008.

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

[10] J. G. Gonzalez and G. R. Arce, "Weighted myriad filters: A robust filtering framework derived from alpha-stable distributions," in *IEEE International Conference on Acoustics, Speech, and Signal Processing Conference Proceedings (ICASSP)*, vol. 5, May 1996, pp. 2833–2836.

[11] ——, "Statistically-efficient filtering in impulsive environments: Weighted myriad filters," in *EURASIP Journal on Advances in Signal Processing*, 2002, pp. 4–20.

ISBN 978-0-9928626-7-1 © EURASIP 2017