Sequential Bayes Factor Testing: A New Framework for Decision Fusion

1st Juan Parras
Information Processing and Telecommunications Center
Universidad Politécnica de Madrid
Madrid, Spain
j.parras@upm.es

2nd Santiago Zazo
Information Processing and Telecommunications Center
Universidad Politécnica de Madrid
Madrid, Spain
santiago.zazo@upm.es

***Abstract*—We propose using a Bayes factor sequential hypothesis test for the decision fusion problem in wireless sensor networks, in which several sensors send a report to a fusion center so that a global decision is taken. This problem is frequently modeled in current literature as a hypothesis test from Bernoulli samples. We propose using a sequential composite hypothesis test based on Bayes Factor using Beta distributions as prior distributions. We obtain closed form expressions for the distributions, which allows us to develop a very efficient algorithm to implement our approach. When we validate our approach via simulations, we observe that, when compared to the common counting rule, our algorithm provides a lower average error and requires a smaller number of samples to make a decision.**

***Index Terms*—Bayes factor, Beta prior, Decision fusion, Wireless sensor networks**

I. Introduction

The problem of testing hypothesis from Bernoulli samples is a recurrent problem that appears in several applications in the signal processing field. For instance, it appears in radar applications [1], pattern identification [2] and sensor fusion in sensor networks [3] [4], [5] [6]. A very used and simple test proposed for this problem is the Neyman Pearson test [7], which for equal confidence reduces to the Counting Rule. Other tests which are also used are the Rao and Wald tests [8].

In all these works, the number of samples used during the testing procedure is fixed. Let us focus in a sensor fusion problem, in which several wireless sensors send a report to a centralized fusion center, in which the information provided by the sensors is combined. Having a fixed number of samples means that the fusion center needs to have a fixed number of reports from the sensors. Even though this could make sense in certain setups, in others it may be better to use a sequential test, such as the Sequential Probability Ratio Test (SPRT) [9], [10]. These statistical tests are specialized on data streams, whose sample size is not fixed in advance. This means that the data is collected and evaluated and according to a predefined stopping rule, a decision is taken on-the-go. The decision might be to stop the analysis because there is enough information to make a decision or to collect more data until there is enough information.

Even though SPRT was originally developed to test simple hypotheses, several approaches have been proposed to deal with composite hypothesis, as shown in [11]. Also, a unified framework for treating composite hypotheses is found in [12], where the author also proposes a nearly-optimal Bayes sequential test for the case of one-sided composite hypothesis. In general, SPRT allows using fewer samples to make a decision, which causes a shorter delay to decide and also, means that fewer transmissions are needed from the wireless sensors to the fusion center, with a subsequent saving both in energy and communication bandwidth. These advantages of the SPRT test are very appreciated in sensor fusion related problems, and thus, it has been the base of several works as [13] or [14].

A different way to approach the problem consists in using the Bayes Factor [15], which traces back to the work of Jeffreys [16], [17]. This approach involves the use of prior probability distribution functions which must be integrated and thus, may cause this approach to be computationally very expensive if there are no closed form expressions of these integrals, which unfortunately happens often.

In this work, we propose using Bayes Factor sequential probability tests in the sensor fusion problem of Bernoulli samples. Until today, many works use fixed length tests, such as [5], [6] or [18]. However, a variable length test may bring some advantages over a fixed length one in issues of capital concern for wireless sensor network such as battery consumption or bandwidth use. In [19], a framework for change detection in sensor networks which uses a non-parametric model is proposed, however, they do not make use of any prior information that may be available. And in [20] there is a study on Bayes Factor sequential probability ratio test, which however is computationally costly and hence, is not adequate for sensor networks.

The major problem with a Bayesian approach in sensor networks is the computational cost. In order to avoid this problem, we use Beta prior distributions, which allow us not only to obtain closed form expressions of the probability distributions involved, but also, allows us to develop an updating rule very efficient both in terms of time and

The work is supported by the Universidad Politécnica de Madrid by a Ph.D. grant given to the first author, as well as by the Spanish Ministry of Science and Innovation under the grant TEC2016-76038-C3-1-R (HERAKLES).

computational resources. Thus, our contribution consists on a very efficient algorithm that can be used as an alternative to the simple Counting Rule, which (1) may make use of prior information because it is a Bayesian approach, (2) is implemented sequentially, hence, offering all the advantages of sequential tests for wireless sensor networks and (3) has a very high computational efficiency.

The rest of the paper goes as follows: in Section II, we describe our problem and the Counting Rule commonly used to solve these problems. In Section III, we introduce the Bayes Factor approach using Beta priors, and then, Section IV introduces our algorithm, that takes advantage of several properties of the Gamma function in order to achieve efficiency. Our approach is compared to the Counting Rule in Section V, and finally, some conclusions are given in Section VI.

II. Problem description

A. Problem setup

We consider a decentralized hypothesis test, in which there are M sensors that are used to discriminate between two different hypotheses: H_0 and H_1 which represent, respectively, the absence or the presence of a certain phenomenon of interest. When required by a central fusion center (FC), the sensor $m \in \{1, 2, ..., M\}$ measures the phenomenon and takes a certain decision $x^m \in \{0, 1\}$. We define $P_d = P(x^m = 1|H_1)$ and $P_{fa} = P(x^m = 1|H_0)$ as the probability of detection and false alarm of each sensor m, where $P(a|b)$ denotes the conditioned probability of a given b.

Under a fixed sample length schema, the FC would ask K reports from the sensors, where K is the sample length size, and then, take a decision. Under a variable sample length schema, the FC asks sequentially for a report x_n^m to sensor m in each time step $n = 1, 2, ...$, where the superscript indexes the sensor and the subscript the time step. The report x_n^m is processed in the FC and either a decision is made, or a new sample $x_{n+1}^{m'}$ is collected from a sensor m'. Note that we are assuming that there are no transmission error between the sensors and the FC: this assumption is frequent in the literature [6], or in [5], where transmission errors are neglected when designing the fusion rules.

We also assume that P_{fa} and P_d are independent and identically distributed among all M sensors and that $P_d > P_{fa}$, as in [5]. Note that we do not consider the case in which each sensor has a different P_d and / or P_{fa} value [19]. Both H_0 and H_1 are modeled using a Bernoulli distribution $B(\theta)$, where:

$$\begin{cases} H_0 : x_n \sim B(\theta = P_{fa}), & n = 1, 2, ..., N \\ H_1 : x_n \sim B(\theta = P_d), & n = 1, 2, ..., N \end{cases} \tag{1}$$

where x_n is the stream of data from the sensors.

Since we assume that $P_d > P_{fa}$, the hypothesis test that the FC faces is:

$$\begin{cases} H_0 : \theta = P_{fa} \\ H_1 : \theta > P_{fa} \end{cases} \tag{2}$$

Note that, as in [5], P_d concrete value need not be explicitly known. Finally, we define $s_n = \sum_{i=1}^{n} x_i$ as the sum of the sensor data up to time n. Note that s_n follows a Binomial distribution and that it contains information coming from all the sensors called by the FC.

B. The Counting Rule

The counting rule (CR) is a decision rule used in many works due to being simple but nonetheless able to outperform more complex mechanisms [5], because it is the universally most powerful test [8] for the fixed length hypothesis test (2). It uses s_n as test statistic for a predetermined value of n, and decides H_1 if $s_n \geq \gamma$, where $P(s_n \geq \gamma|P_d) \leq \alpha$ allows fixing the decision threshold γ as a function of the significance level α fixed a priori. The power of the test depends on α and n: a larger n brings a higher power to the test, at the cost of a longer delay to the decision and a larger number of communications required. Hence, note that there is a tradeoff between precision and resources consumption.

III. Bayes factor using beta priors

Under a Bayesian scheme, we assume to know a prior probability distribution for each hypothesis, $p(H_0)$ and $p(H_1) = 1 - p(H_0)$. Using Bayes theorem, it is possible to obtain $p(H_k|x_n)$ as:

$$p(H_k|x_n) = \frac{p(x_n|H_k)p(H_k)}{\sum_k p(x_n|H_k)p(H_k)}, \quad k = \{0, 1\} \tag{3}$$

The expression (3) can be manipulated to obtain:

$$\frac{p(H_1|x_n)}{p(H_0|x_n)} = \frac{p(x_n|H_1)}{p(x_n|H_0)} \frac{p(H_1)}{p(H_0)} = B_{10} \frac{p(H_1)}{p(H_0)} \tag{4}$$

In (4), we observe that the posterior odds are the prior odds times a B_{10} term , which is the Bayes factor (BF). Intuitively, the BF carries information about how likely is the data x_n to have been generated under models H_0 or H_1. There are several thresholds proposed for the value of the BF in order to take a decision [15].

The densities $p(x_n|H_k)$ need to be computed in order to obtain the BF. By assuming that each hypothesis is modeled using a distribution $p(\theta|H_k)$ with an unknown parameter θ, the densities $p(x_n|H_k)$ can be obtained by integration as:

$$p(x_n|H_k) = \int p(x_n|\theta, H_k)p(\theta|H_k)d\theta, \quad k = \{0, 1\} \tag{5}$$

By taking into account that in our problem x_n follows a Bernoulli distribution, (5) becomes:

$$p(x_n|H_k) = \int_0^1 \theta^{s_n}(1 - \theta)^{n-s_n} p(\theta|H_k)d\theta, \quad k = \{0, 1\} \tag{6}$$

A major problem to use the BF is that (6) can be hard to obtain, see for instance [15], where several numerical methods are reviewed. In the best case, (6) can be analytically evaluated and hence, there is no need of numerical methods. This also brings a significant improvement both in computational efficiency and precision. For choosing a prior $p(\theta|H_k)$ analytically evaluable, the family of conjugate distributions is of special interest because the conjugacy property holds: the posterior distribution is in the same family of distributions

as the prior [21]. These reasons motivate us to choose Beta distributions as priors, which belong to the exponential family. As we will see, this choice will allow us to design a simple and efficient sequential update algorithm to obtain (6). The beta distribution has two parameters $\gamma_1 > 0$ and $\gamma_2 > 0$, which we fix a priori. The pdf of the Beta distribution is:

$$Beta(\theta|\gamma_1,\gamma_2) = \frac{\theta^{\gamma_1-1}(1-\theta)^{\gamma_2-1}}{B(\gamma_1,\gamma_2)} \tag{7}$$

where the normalization factor $B(\gamma_1,\gamma_2)$ is the Beta function (also known as Euler integral of first kind):

$$B(\gamma_1,\gamma_2) = \int_0^1 \theta^{\gamma_1-1}(1-\theta)^{\gamma_2-1}d\theta = \frac{\Gamma(\gamma_1)\Gamma(\gamma_2)}{\Gamma(\gamma_1+\gamma_2)} \tag{8}$$

where $\Gamma(a)$ stands for the gamma function of a.

We choose to use as prior a weighted sum of L Beta distributions, because they allow us to model complicated priors, such as multimodal distributions. The choice of the prior parameters will be chosen as to adapt to the prior information available. We define γ^k as the matrix of parameters of the prior distribution of θ under hypothesis H_k, with L rows containing the beta distribution parameters $(\gamma_1^{l,k},\gamma_2^{l,k})$. Using (7), our prior becomes:

$$\begin{aligned} p(\theta|\gamma) &= \sum_{l=1}^{L} w_l beta(\theta|\gamma_1^{l,k},\gamma_2^{l,k}) \\ &= \sum_{l=1}^{L} w_l \frac{\theta^{\gamma_1^{l,k}-1}(1-\theta)^{\gamma_2^{l,k}-1}}{B(\gamma_1^{l,k},\gamma_2^{l,k})} \end{aligned} \tag{9}$$

where $\sum_{l=1}^{L} w_l = 1$ and each $w_l \geq 0$, so that (9) defines a distribution. We now can compute the posterior probabilities $p(x_n|H_k)$ using (6) and (9) as follows:

$$\begin{aligned} &p(x_n|H_k) = \\ &\int_0^1 \theta^{s_n}(1-\theta)^{n-s_n} \sum_{l=1}^{L} w_l \frac{\theta^{\gamma_1^{l,k}-1}(1-\theta)^{\gamma_2^{l,k}-1}}{B(\gamma_1^{l,k},\gamma_2^{l,k})} d\theta \\ &\quad = \sum_{l=1}^{L} w_l \frac{\int_0^1 \theta^{s_n+\gamma_1^{l,k}-1}(1-\theta)^{n-s_n+\gamma_2^{l,k}-1}d\theta}{B(\gamma_1^{l,k},\gamma_2^{l,k})} \\ &\quad = \sum_{l=1}^{L} w_l \frac{B(s_n+\gamma_1^{l,k}, n-s_n+\gamma_2^{l,k})}{B(\gamma_1^{l,k},\gamma_2^{l,k})} \end{aligned} \tag{10}$$

IV. Bayes factor update algorithm

The expression obtained for $p(x_n|H_k)$ in (10) allows obtaining an efficient sequential algorithm to update the prior in a sequential test, as new samples x_n arrive. First, we express (10) in terms of the Gamma function using (8) as follows:

$$p(x_n|H_k) = \sum_{l=1}^{L} w_l S_{l,k}^n \tag{11}$$

where

$$S_{l,k}^n = \frac{\Gamma(s_n+\gamma_1^{l,k})}{\Gamma(\gamma_1^{l,k})} \frac{\Gamma(n-s_n+\gamma_2^{l,k})}{\Gamma(\gamma_2^{l,k})} \frac{\Gamma(\gamma_1^{l,k}+\gamma_2^{l,k})}{\Gamma(n+\gamma_1^{l,k}+\gamma_2^{l,k})} \tag{12}$$

The values for $S_{l,k}^n$ in (12) can be obtained recursively with the help of Lemma 1.

Lemma 1. *The following identity holds for $a \in \{0,1,2,3...\}$ and $k > 0$:*

$$\frac{\Gamma(k+a)}{\Gamma(k)} = \begin{cases} \prod_{i=k}^{k+a-1} i & \text{if} \quad a \geq 1 \\ 1 & \text{if} \quad a = 0 \end{cases}$$

Proof. For $a = 0$, the proof is straightforward:

$$\left.\frac{\Gamma(k+a)}{\Gamma(k)}\right|_{a=0} = \frac{\Gamma(k)}{\Gamma(k)} = 1$$

For $a \geq 1$, we will use the following property of the gamma function which holds for any real number $z > 0$:

$$\Gamma(z+1) = z\Gamma(z) \tag{13}$$

Proceeding by induction, for $a = 1$ and $k > 0$:

$$\frac{\Gamma(k+1)}{\Gamma(k)} = \frac{k\Gamma(k)}{\Gamma(k)} = k$$

where we used (13). For $a = 2$, we have that:

$$\frac{\Gamma(k+2)}{\Gamma(k)} = \frac{(k+1)\Gamma(k+1)}{\Gamma(k)} = (k+1)k$$

where again we used (13). Now, we assume that for $a > 1$, the following holds:

$$\frac{\Gamma(k+a)}{\Gamma(k)} = \prod_{i=k}^{k+a-1} i$$

and proceed to obtain the value for $a+1$:

$$\frac{\Gamma(k+a+1)}{\Gamma(k)} = \frac{(k+a)\Gamma(k+a)}{\Gamma(k)} = (k+a)\prod_{i=k}^{k+a-1} i = \prod_{i=k}^{k+a} i$$

which finishes the proof. □

Lemma 1 allows obtaining the values for $S_{l,k}^n$ in (12) sequentially, as new data x_n arrives. Observe that (12) can be expressed as:

$$S_{l,k}^n = S_{1,l,k}^n S_{2,l,k}^n \left(S_{3,l,k}^n\right)^{-1} \tag{14}$$

where:

$$\begin{cases} S_{1,l,k}^n = \dfrac{\Gamma(\gamma_1^{l,k}+s_n)}{\Gamma(\gamma_1^{l,k})} \\ S_{2,l,k}^n = \dfrac{\Gamma(\gamma_2^{l,k}+n-s_n)}{\Gamma(\gamma_2^{l,k})} \\ S_{3,l,k}^n = \dfrac{\Gamma(\gamma_1^{l,k}+\gamma_2^{l,k}+n)}{\Gamma(\gamma_1^{l,k}+\gamma_2^{l,k})} \end{cases} \tag{15}$$

In these expressions, s_n, $n-s_n$ and n are natural numbers and hence, we can apply Lemma 1 to recursively update $S_{l,k}^n$ as new data x_n arrives. Observe that all γ parameters are greater than zero, for the Beta distribution parameters must be positive. Thus, all conditions from Lemma 1 are satisfied.

The updating procedure depends on each new x_n. If $x_n = 1$, then s_n and n increase one unit with respect to their previous values, whereas $n - s_n$ remains the same. Hence, we must

update $S^n_{1,l,k}$ and $S^n_{3,l,k}$ only. If $x_n = 0$, then $n - s_n$ and n increase one unit with respect to their previous values, whereas s_n remains the same. Hence, we must update $S^n_{2,l,k}$ and $S^n_{3,l,k}$ only. Thus, when a sample x_n arrives, we always update $S^3_{c,l,k}$ and depending on whether $x_n = 1$ or $x_n = 0$ we update $S^n_{1,l,k}$ or $S^n_{2,l,k}$. Note that these updatings are straightforward according to Lemma 1. If $s_n = 0$, then $S^n_{1,l,k} = 1$. And if $s_n \geq 1$, then $S^n_{1,l,k} = (\gamma^{l,k}_1 + s_n - 1)S^{n-1}_{1,l,k}$. A similar reasoning applies to $S^n_{2,l,k}$ and $S^n_{3,l,k}$.

Algorithm 1 Sequential Bayes test

Input: $\gamma^k, w_l, B_{t,0}, B_{t,1}$
Initialize $stop = False$, $n = 1$, $s_0 = 0$
Initialize $S^0_{1,l,k} = S^0_{2,l,k} = S^0_{3,l,k} = 1$
while $stop$ is $False$ **do**
  Obtain a new sample x_n
  Update $s_n = \sum_{i=1}^{n} x_i = s_{n-1} + x_n$
  for $k = 0, 1$ **do**
    for $l = 1, 2, ..., L$ **do**
      $S^n_{3,l,k} = (\gamma^{l,k}_1 + \gamma^{l,k}_2 + n - 1)S^{n-1}_{3,l,k}$
      if $x_n = 1$ **then**
        $S^n_{1,l,k} = (\gamma^{l,k}_1 + s_n - 1)S^{n-1}_{1,l,k}$
        $S^n_{2,l,k} = S^{n-1}_{2,l,k}$
      if $x_n = 0$ **then**
        $S^n_{2,l,k} = (\gamma^{l,k}_2 + n - s_n - 1)S^{n-1}_{2,l,k}$
        $S^n_{1,l,k} = S^{n-1}_{1,l,k}$
  Obtain $p(x_n|H_k)$ using (11) for $k = \{0, 1\}$
  Obtain $B^n_{10} = \frac{p(x_n|H_1)}{p(x_n|H_0)}$
  if $B^n_{10} > B_{1,t}$ **then**
    Decide H_1 and set $stop = True$
  else if $B^n_{10} < B_{0,t}$ **then**
    Decide H_0 and set $stop = True$
  else
    Set $n = n + 1$
Output: Decison taken, n

With all this, we propose an Algorithm that makes use of the procedure described above to sequentially update the marginal distributions $p(x_n|H_k)$ in order to obtain a sequential test, based on the Bayes factor B_{10}. We provide as inputs to the algorithm the γ^k prior values and the w_l weights for each l value, as well as the threshold values $B_{t,0}$ and $B_{t,1}$ that we wish to establish as stopping rules. When a new sample x_n arrives, the algorithm updates the two marginal distributions $p(x_n|H_0)$ and $p(x_n|H_1)$, obtains the Bayes factor at sample n B^n_{10} and compares it to the two thresholds. If $B_{10} > B_{1,t}$, the test stops and H_1 is accepted. If $B_{10} < B_{0,t}$, the test stops and H_0 is accepted. Otherwise, a new sample is obtained. The whole procedure is summarized in Algorithm 1. Observe that this algorithm is very efficient, for:

- We do not need to evaluate any gamma function.
- Only sums, products and divisions are involved in each algorithm iteration.
- The updating is based in a constant number of operations as each new sample x_n arrives, thus, our algorithm has a linear number of operations with L and constant with n.

All these reasons make Algorithm 1 very suitable for a limited resources environment as decision fusion in a wireless sensor network.

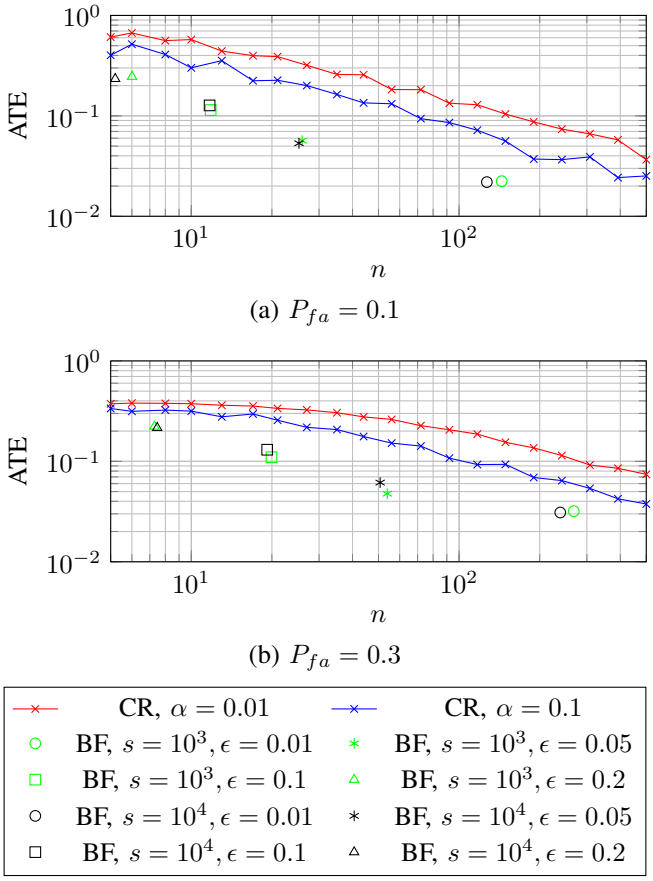

Fig. 1. Simulation result curves. Note that our proposed BF approach obtains a lower averaged total error using fewer samples n than the CR, for all the s and ϵ values tested. Note that in the BF approach, the tested values of ϵ have a greater impact than the values of s on the test ATE. This is to be expected, since ϵ controls the sensitivity of the test. For all the values tested, our BF approach significantly outperforms the CR.

V. Empirical performance evaluation

We compare the BF Algorithm we propose with the optimized Counting Rule, which, as we indicated, is commonly used and hard to beat due to being the universally most powerful test [8]. Since the CR performance is not directly comparable to our method, we use an Averaged Total Error (ATE) metric to compare them. We employ a set of 21 Bernoulli parameters θ_{test} linearly spaced in the interval $[0, 0.5]$. Then, we fix P_{fa} and we perform 100 hypothesis tests for each θ_{test}, using both CR and our BF method. After each of these hypothesis tests, we obtain the ATE by adding the number of erroneous decisions taken and dividing by the 2100 simulations performed. Note that this means that ATE gives us an averaged measure of the decision error over the θ_{test} values. We simulate for $P_{fa} \in \{0.1, 0.3\}$. For the CR, we use $\alpha \in \{0.01, 0.1\}$ and employ 20 logarithmically spaced values for n in the interval $[5, 500]$.

For the BF approach, we use as decision thresholds $B_{0,t} = 3^{-1}$ and $B_{1,t} = 3$, following [15]. We also need to choose the prior parameters. For the simulations, we use $L = 1$ Beta prior distribution for simplicity, and we define the Beta parameters as a function of two values: the strength of the prior, defined as $s = \gamma_1^k + \gamma_2^k$, which denote the confidence we have in the prior, and $\epsilon = P_d - P_{fa}$, which controls the sensitivity of the BF test. Since the mean of the beta distribution is:

$$\mu = \frac{\gamma_1}{\gamma_1 + \gamma_2}$$

we set $\mu = P_{fa}$ and obtain the Beta distribution parameters for the set $\langle s, P_{fa}, \epsilon \rangle$ as:

$$\begin{cases} \gamma_1^{1,0} = s \cdot P_{fa} \\ \gamma_2^{1,0} = s \cdot (1 - P_{fa}) \\ \gamma_1^{1,1} = s \cdot (P_{fa} + \epsilon) \\ \gamma_2^{1,1} = s \cdot (1 - P_{fa} - \epsilon) \end{cases} \tag{16}$$

In our simulations, we use (16) to define the prior distributions with $s = \{10^3, 10^4\}$ and $\epsilon = \{0.01, 0.05, 0.1, 0.2\}$, that is, we test for two different confidence value in the prior and for different sensitivities for the BF test. The results are in Figure 1, where we can observe that our proposed BF approach performs significantly better than CR in both ATE and the number of samples required to take a decision. This means that using a sequential BF hypothesis test provides a lower average error of decision taking a smaller number of samples, and both are crucial in a sensor network.

In our simulations, we also observed that the total error using BF is distributed around $\theta_{test} = P_{fa}$, while the CR strongly concentrates its error on $\theta_{test} > P_{fa}$ in order to satisfy the restriction $P(H_1|H_0) \leq \alpha$. That is, the CR provides a bound in the type I error (α), but the type II error depends on n and increases as n decreases: lower sample sizes yield a higher type II error. Finally, we note that the BF test performance could be improved by having a more detailed knowledge about the prior distributions. Observe that we used a simple prior for these simulations, and in real life environments, in which a certain knowledge of the prior may be present [22] [23], our proposed BF algorithm may perform even better.

VI. Conclusions

In this work, we propose an algorithm to perform a Bayes Factor test on a decision fusion problem for Bernoulli distributions. We obtain closed form expressions of the probability distributions, which allows us to obtain a sequential test implementation of special interest in wireless sensor networks, as such a test reduces the number of transmissions required to the FC. Our algorithm is very efficient, and in terms of performance, it provides a lower average error and requires fewer samples to decide than the commonly used Counting Rule. Thus, the proposed algorithm is specially suitable for decision fusion in wireless sensor networks. Future work may include extending our algorithm to the case in which each sensor has a different false alarm and detection probabilities.

References

[1] David A Shnidman, "Binary integration for swerling target fluctuations," *IEEE Transactions on Aerospace and Electronic systems*, vol. 34, no. 3, pp. 1043–1053, 1998.

[2] Manuel F Fernandez and Tom Aridgides, "Measures for evaluating sea mine identification processing performance and the enhancements provided by fusing multisensor/multiprocess data via an m-out-of-n voting scheme," in *Detection and Remediation Technologies for Mines and Minelike Targets VIII*. International Society for Optics and Photonics, 2003, vol. 5089, pp. 425–437.

[3] Ruixin Niu and Pramod K Varshney, "Decision fusion in a wireless sensor network with a random number of sensors," in *Acoustics, Speech, and Signal Processing, 2005. Proceedings.(ICASSP'05). IEEE International Conference on*. IEEE, 2005, vol. 4, pp. iv–861.

[4] Ruixin Niu and Pramod K Varshney, "Performance analysis of distributed detection in a random sensor field," *IEEE Transactions on Signal Processing*, vol. 56, no. 1, pp. 339–349, 2008.

[5] Domenico Ciuonzo and P Salvo Rossi, "Decision fusion with unknown sensor detection probability," *IEEE Signal Processing Letters*, vol. 21, no. 2, pp. 208–212, 2014.

[6] Mohammad Fayazur Rahaman and Mohammed Zafar Ali Khan, "Low-complexity optimal hard decision fusion under the neyman–pearson criterion," *IEEE Signal Processing Letters*, vol. 25, no. 3, pp. 353–357, 2018.

[7] Steven M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1993.

[8] Domenico Ciuonzo, Antonio De Maio, and P Salvo Rossi, "A systematic framework for composite hypothesis testing of independent bernoulli trials," *IEEE Signal Processing Letters*, vol. 22, no. 9, pp. 1249–1253, 2015.

[9] Abraham Wald, "Statistical decision functions which minimize the maximum risk," *Annals of Mathematics*, pp. 265–280, 1945.

[10] Abraham Wald, *Sequential analysis*, Courier Corporation, 1973.

[11] Tze Leung Lai, "Sequential analysis: Some classical problems and new challenges," *Statistica Sinica*, vol. 11, no. 2, pp. 303–351, 2001.

[12] Tze Leung Lai, "Nearly optimal sequential tests of composite hypotheses," *The Annals of Statistics*, vol. 16, no. 2, pp. 856–886, 1988.

[13] Ruiliang Chen, J-M Park, and Kaigui Bian, "Robust distributed spectrum sensing in cognitive radio networks," in *INFOCOM 2008. The 27th Conference on Computer Communications. IEEE*. IEEE, 2008, pp. 1876–1884.

[14] Yeelin Shei and Yu T Su, "A sequential test based cooperative spectrum sensing scheme for cognitive radios," in *Personal, Indoor and Mobile Radio Communications, 2008. PIMRC 2008. IEEE 19th International Symposium on*. IEEE, 2008, pp. 1–5.

[15] Robert E Kass and Adrian E Raftery, "Bayes factors," *Journal of the american statistical association*, vol. 90, no. 430, pp. 773–795, 1995.

[16] Harold Jeffreys, "Some tests of significance, treated by the theory of probability," vol. 31, no. 2, pp. 203–222, 1935.

[17] H Jeffreys, *Theory of Probability*, Oxford University Press, 1961.

[18] Fayazur Rahaman Mohammad, Domenico Ciuonzo, and Zafar Ali Khan Mohammed, "Mean-based blind hard decision fusion rules," *IEEE Signal Processing Letters*, vol. 25, no. 5, pp. 630–634, 2018.

[19] Domenico Ciuonzo and P Salvo Rossi, "Dechade: Detecting slight changes with hard decisions in wireless sensor networks," *International Journal of General Systems*, vol. 47, no. 5, pp. 535–548, 2018.

[20] Felix D Schönbrodt, Eric-Jan Wagenmakers, Michael Zehetleitner, and Marco Perugini, "Sequential hypothesis testing with bayes factors: Efficiently testing mean differences," *Psychological Methods*, vol. 22, no. 2, pp. 322, 2017.

[21] Persi Diaconis, Donald Ylvisaker, et al., "Conjugate priors for exponential families," *The Annals of statistics*, vol. 7, no. 2, pp. 269–281, 1979.

[22] Ruixin Niu, Biao Chen, and Pramod K Varshney, "Fusion of decisions transmitted over rayleigh fading channels in wireless sensor networks," *IEEE Transactions on signal processing*, vol. 54, no. 3, pp. 1018–1027, 2006.

[23] Ying Lin, Biao Chen, and Pramod K Varshney, "Decision fusion rules in multi-hop wireless sensor networks," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 41, no. 2, pp. 475–488, 2005.