On Centralized Composite Detection with Distributed Sensors

Cuichun Xu, and Steven Kay

Abstract

For composite hypothesis testing, the generalized likelihood ratio test (GLRT) and the Bayesian approach are two widely used methods. This paper investigates the two methods for signal detection of a known waveform and unknown amplitude with distributed sensors. It is first proved that the performance of GLRT can be poor and hence improved for this problem and then an approximated Bayesian detector is proposed. Compared with the exact Bayesian approach, the proposed method always has a closed form and hence is easy to implement. Computer simulation results show that the proposed method has comparable performance to the exact Bayesian approach.

I. Introduction

Detection with distributed sensors has been studied for nearly three decades [1] [2] [3]. With respect to different data assumptions, it can be categorized to centralized detection and decentralized detection. In centralized detection, it is assumed that all data from all local sensors are available for processing. In decentralized detection, only compressed data or local decision from all local sensors are communicated to a central processor, where a central decision is made. Since the centralized detection can largely resort to classical detection theory, many works focus on decentralized detection [1] [2] [3]. However, there are still some problems left for centralized detection, especially when there are unknown parameters in hypotheses, i.e. composite detection.

In [4], a minimax constant false alarm rate (CFAR) centralized detector is proposed for composite detection. However, it assumes that the unknown parameters take discrete values and the detection involves large computation. In [3], many distributed detection techniques are reviewed, with focus on the locally optimum distributed detectors, which are optimal only when signal is weak. Other common composite detection methods include the generalized likelihood ratio test (GLRT) and the Bayesian approach.

Because of its ease of implementation, the GLRT is widely used and usually has satisfactory performance. However, its optimality is hard to establish, even though some attempts have been made, based

on different criteria [8] [9]. On the other hand, the Bayesian approach is optimal if the assumed prior probabilities and prior probability density functions (PDFs) of the unknown parameters are true. However, in most applications the true priors and true prior PDFs are unknown and hard to assume. Furthermore the integration involved in calculation of the marginal PDF may not be easy to evaluate [12].

In this paper, we examine the important special case of independent sensors (conditioned on hypothesis). In particular, we focus on situations that even when a signal is present only part of the sensors receive the signal. It is shown in Section III that in this case, the GLRT is not optimal. Some improvements of the GLRT are also discussed. In section IV, an approximated Bayesian detector is proposed based on vague prior PDFs, which does not need a specific form. Some computer simulation results are shown in Section V and finally Section VI offers conclusion and further discussion.

II. Statement of the Problem

To simplify the mathematical exposition, we assume the sensors are ordered with respect to their importance. That is if $i \leq j$ and sensor j receives signal then sensor i must also receive signal. This order of importance may be due to the physical locations of the sensors. When the sensors are not ordered, the discussion is given in Section VI.

We assume that there are M independent sensors or channels. The detection problem is stated as

$$\begin{aligned} H_0 &: \mathbf{x}[n] = \mathbf{w}[n] && n = 0, 1, \cdots, N-1 \\ H_1 &: \mathbf{x}[n] = \mathbf{A}s[n] + \mathbf{w}[n] && n = 0, 1, \cdots, N-1 \end{aligned} \tag{1}$$

where $\mathbf{A}$ is the $M \times 1$ unknown amplitude vector, with possibly different elements, $s[n]$ is a known waveform and $\mathbf{w}[n]$ is a Gaussian random vector. Each element of $\mathbf{w}[n]$ is white both in time domain and in space domain, i.e.

$$E[w_i[n] w_i[m]] = \begin{cases} \sigma^2, & n = m \\ 0, & n \neq m \end{cases} \quad \text{for any } 1 \leq i \leq M$$

and

$$E[w_i[n] w_j[n]] = \begin{cases} \sigma^2, & i = j \\ 0, & i \neq j \end{cases} \quad \text{for any } 0 \leq n \leq N-1$$

where $w_i[n]$ denotes the i^{th} element of $\mathbf{w}[n]$ and $E[\cdot]$ denotes the expectation operator. Since the sensors are ordered, under H_1 only the first L elements of $\mathbf{A}$ is nonzero. L is referred to as the true model order under H_1. Both L and the value of the first L elements of $\mathbf{A}$ are unknown. If we assume there are i nonzero elements in $\mathbf{A}$, the corresponding model of (1) is called model i and is denoted as $\mathcal{M}_i$. Under

model $\mathcal{M}_i$, $\mathbf{A}$ is denoted as $\mathbf{A}_i$ and $\mathbf{A}_i = [A_1, A_2, \cdots, A_i, 0, 0, \cdots, 0]_{M\times 1}$ Clearly, under H_0 the data is from model $\mathcal{M}_0$ and under H_1 the data is from model $\mathcal{M}_i$, $1 \le i \le M$. With this modeling, the detection problem can equivalently be stated as to assign the given data to $\mathcal{M}_0$ (H_0) or to $\mathcal{M}_1 \cup \mathcal{M}_2 \cup \cdots \cup \mathcal{M}_M$ (H_1).

III. Generalized Likelihood Ratio Test

A. Performance of the GLRT

Firstly, suppose the variance σ^2 is known. Since L and $\mathbf{A}$ are unknown, the GLRT decides H_1 if

$$L_G(\mathbf{X}) = \max_{1\le i\le M} \frac{p\left(\mathbf{X}; \hat{\mathbf{A}}_i, \mathcal{M}_i\right)}{p(\mathbf{X}; \mathcal{M}_0)} > \lambda$$

where $\mathbf{X} = [\mathbf{x}[0], \mathbf{x}[1], \cdots, \mathbf{x}[N-1]]^T$ denotes the whole data and $p\left(\mathbf{X}; \hat{\mathbf{A}}_i, \mathcal{M}_i\right)$ is the probability density function (PDF) under $\mathcal{M}_i$ parameterized by the maximum likelihood estimator (MLE) $\hat{\mathbf{A}}_i$, which is estimated assuming $\mathcal{M}_i$ is true. Since if $i \le j$, $\mathcal{M}_i$ can be thought of as a special case of $\mathcal{M}_j$, it is readily shown that the GLRT will always implement the test statistic based on the maximum model, i.e. [5]

$$L_G(\mathbf{X}) = \frac{p\left(\mathbf{X}; \hat{\mathbf{A}}_M, \mathcal{M}_M\right)}{p(\mathbf{X}; \mathcal{M}_0)} > \lambda$$

or equivalently

$$2\ln L_G(\mathbf{X}) = 2\ln \frac{p\left(\mathbf{X}; \hat{\mathbf{A}}_M, \mathcal{M}_M\right)}{p(\mathbf{X}; \mathcal{M}_0)} > \lambda' \tag{2}$$

The test statistic of (2) has the PDF [5]

$$2\ln L_G(\mathbf{X}) \sim \begin{cases} \chi^2_M, & H_0 \\ \chi'^2_M(\lambda), & H_1 \end{cases} \tag{3}$$

where χ^2_M denotes the chi-squared distribution and $\chi'^2_M(\lambda)$ denotes the noncentral chi-squared distribution with the noncentrality parameter λ. The noncentrality parameter is given by

$$\lambda = \frac{\varepsilon \sum_{i=1}^{M} A_i^2}{\sigma^2}$$

where

$$\varepsilon = \sum_{n=0}^{N-1} s^2[n]$$

is the signal waveform energy. Since only the first L elements of $\mathbf{A}$ are nonzero, we have

$$\lambda = \frac{\varepsilon \sum_{i=1}^{L} A_i^2}{\sigma^2}$$

With the PDF of the GLRT test statistic of (3), we now show the performance of the GLRT is generally not optimal and can be improved. Suppose $L < M$ and let K be an integer such that $L \leq K < M$. If there is a generalized likelihood ratio test based on $\mathcal{M}_K$, which decides H_1 if

$$2 \ln L_{G_K}(\mathbf{X}) = 2 \ln \frac{p\left(\mathbf{X}; \hat{\mathbf{A}}_K, \mathcal{M}_k\right)}{p(\mathbf{X}; \mathcal{M}_0)} > \lambda'' \tag{4}$$

the PDF of this test statistic is given by

$$2 \ln L_{G_K}(\mathbf{X}) \sim \begin{cases} \chi^2_K, & H_0 \\ \chi'^2_K(\lambda), & H_1 \end{cases} \tag{5}$$

We note that if the variance σ^2 is unknown, (2) and (4) should be changed to

$$2 \ln L_G(\mathbf{X}) = 2 \ln \frac{p\left(\mathbf{X}; \hat{\mathbf{A}}_M, \hat{\sigma}^2_M, \mathcal{M}_M\right)}{p(\mathbf{X}; \hat{\sigma}^2_0, \mathcal{M}_0)} > \lambda'$$

and

$$2 \ln L_{G_K}(\mathbf{X}) = 2 \ln \frac{p\left(\mathbf{X}; \hat{\mathbf{A}}_K, \hat{\sigma}^2_K, \mathcal{M}_K\right)}{p(\mathbf{X}; \hat{\sigma}^2_0, \mathcal{M}_0)} > \lambda''$$

where $\hat{\sigma}^2_i$ is the MLE of σ^2 under $\mathcal{M}_i$. However, asymptotically (3) and (5) still hold [5]. Since $K < M$, from Theorem 1, which is proved in Appendix I, the detector based on model $\mathcal{M}_K$ has better performance than that of the GLRT, which is based on $\mathcal{M}_M$.

Theorem 3.1: Suppose there are two detection statistics T_1 and T_2. The detection statistic T_1 has the PDF

$$T_1 \sim \begin{cases} \chi^2_{\nu 1}, & H_0 \\ \chi'^2_{\nu 1}(\lambda), & H_1 \end{cases}$$

and the detector 1 decides a signal is present if $T_1 > \gamma_1$. The detection statistic T_2 has the PDF

$$T_2 \sim \begin{cases} \chi^2_{\nu 2}, & H_0 \\ \chi'^2_{\nu 2}(\lambda), & H_1 \end{cases}$$

and the detector 2 decides a signal is present if $T_2 > \gamma_2$. Then, if $\nu 1 < \nu 2$, the performance of detector 1 is better than that of detector 2 in a Neyman-Pearson (NP) sense, i.e. for the same false alarm rate detector 1 has a higher probability of detection than detector 2.

In other words, for this distributed detection problem, if $L < M$ the GLRT is not optimal. This is because the GLRT is based on $\mathcal{M}_M$ and therefore the channels that contain only noise samples are included in the test statistic. Including these channels will increase the degrees of freedom of the test statistic, however the noncentrality parameter remains the same.

B. Improved GLRT

From Theorem 1, any detector based on $\mathcal{M}_i$ with $i > L$ will have a performance less than the detector based on $\mathcal{M}_L$. It is clear that the unknown order L is critical to the performance of detectors. In order to improve the GLRT, L can first be estimated and a generalized likelihood ratio test based on model $\hat{L}$ can then be conducted. This can be done by using various model order selection criteria [6]. A widely used criterion is the minimum description length (MDL) criterion, which selects the model that minimizes

$$MDL(i) = -2\ln L_{G_i}(\mathbf{X}) + i\ln N, \quad 1 \le i \le M \tag{6}$$

The GLRT with estimated model is referred to as "rGLRT" in this paper.

Recently, a multifamily likelihood ratio test (MFLRT) is proposed to modify the GLRT to accommodate nested PDF families [7]. The MFLRT decides H_1 if

$$T_{MFLRT}(\mathbf{X}) = \max_{1\le i\le M}\left\{\left[L_{G_i}(\mathbf{X}) - i\left(\ln\left(\frac{L_{G_i}(\mathbf{X})}{i}\right)+1\right)\right]u\left(\frac{L_{G_i}(\mathbf{X})}{i}-1\right)\right\} > \gamma \tag{7}$$

where $u(\cdot)$ is the unit step function.

These two revised GLRTs will be compared with the GLRT in Section V.

IV. Bayesian Approach

A. Approximated Bayesian Approach

Firstly, suppose the variance σ^2 is known. If we are willing to assume Bayesian assumptions, that is model $\mathcal{M}_i$, $0 \le i \le M$ has a prior probability π_i and under each model in H_1, $\mathbf{A}_j$, $1 \le j \le M$ has a prior PDF given as $p(\mathbf{A}_j|\mathcal{M}_j)$, the optimal NP detector decides H_1 if

$$\frac{p(\mathbf{X}|H_1)}{p(\mathbf{X}|H_0)} > \lambda \tag{8}$$

From the Bayesian assumptions

$$p(\mathbf{X}|H_1) = \sum_{i=1}^{M}\pi_i p(\mathbf{X}|\mathcal{M}_i) \tag{9}$$

Plugging (9) into (8), the detector decides H_1 equivalently if

$$\frac{\sum_{i=1}^{M}\pi_i p(\mathbf{X}|\mathcal{M}_i)}{p(\mathbf{X}|\mathcal{M}_0)} > \lambda \tag{10}$$

where the marginal PDF $p(\mathbf{X}|\mathcal{M}_i)$ is given by

$$\begin{aligned} p(\mathbf{X}|\mathcal{M}_i) &= \int p(\mathbf{X},\mathbf{A}_i|\mathcal{M}_i)\,d\mathbf{A}_i \\ &= \int p(\mathbf{X}|\mathbf{A}_i,\mathcal{M}_i)\,p(\mathbf{A}_i|\mathcal{M}_i)\,d\mathbf{A}_i \end{aligned} \tag{11}$$

As mentioned previously, the true prior PDFs are usually unknown and hard to assume. Furthermore the integration in (11) may not have a closed form. If the variance σ^2 is unknown $\int p\left(\mathbf{X}|\mathbf{A}_i,\mathcal{M}_i\right)p\left(\mathbf{A}_i|\mathcal{M}_i\right)d\mathbf{A}_i$ should be replaced by $\iint p\left(\mathbf{X}|\mathbf{A}_i,\sigma^2,\mathcal{M}_i\right)p\left(\mathbf{A}_i,\sigma^2|\mathcal{M}_i\right)d\mathbf{A}_i d\sigma^2$ and it is even more difficult to assume the prior PDFs as well as calculate the integration.

It is shown in [10] that assuming vague prior PDFs, the marginal PDF $p(\mathbf{X}|\mathcal{M}_i)$ can be asymptotically approximated as

$$p\left(\mathbf{X}|\mathcal{M}_i\right)\approx const\cdot p\left(\mathbf{X}|\hat{\mathbf{A}}_i,\mathcal{M}_i\right)\left|\mathbf{I}\left(\hat{\mathbf{A}}_i\right)\right|^{-\frac{1}{2}}(2\pi e)^{\frac{i}{2}}$$

where $|\cdot|$ denotes determinant and $\mathbf{I}\left(\hat{\mathbf{A}}_i\right)$ is the observed information matrix, or

$$\mathbf{I}\left(\hat{\mathbf{A}}_i\right)=-\left.\frac{\partial^2\ln p\left(\mathbf{X}|\mathbf{A}_i,\mathcal{M}_i\right)}{\partial\mathbf{A}_i\partial\mathbf{A}_i^T}\right|_{\mathbf{A}_i=\hat{\mathbf{A}}_i}$$

For the model given by (1), it can be shown that asymptotically

$$\left|\mathbf{I}\left(\hat{\mathbf{A}}_i\right)\right|\approx N^i$$

and therefore

$$p\left(\mathbf{X}|\mathcal{M}_i\right)\approx const\cdot p\left(\mathbf{X}|\hat{\mathbf{A}}_i,\mathcal{M}_i\right)\left(\frac{N}{2\pi e}\right)^{-\frac{i}{2}} \tag{12}$$

Plugging the above into (10) we have

$$\frac{\sum_{i=1}^{M}\pi_i const\cdot p\left(\mathbf{X}|\hat{\mathbf{A}}_i,\mathcal{M}_i\right)\left(\frac{N}{2\pi e}\right)^{-\frac{i}{2}}}{p\left(\mathbf{X}|\mathcal{M}_0\right)}>\lambda'$$

or

$$\sum_{i=1}^{M}\pi_i\left(\frac{N}{2\pi e}\right)^{-\frac{i}{2}}L_{G_i}\left(\mathbf{X}\right)>\lambda'' \tag{13}$$

When σ^2 is unknown, similarly we have

$$p\left(\mathbf{X}|\mathcal{M}_i\right)\approx const\cdot p\left(\mathbf{X}|\hat{\mathbf{A}}_i,\hat{\sigma}_i^2,\mathcal{M}_i\right)\left(\frac{N}{2\pi e}\right)^{-\frac{i+1}{2}}$$

and (13) is still valid. The detector given by (13) is referred to as "aBayesian" in the simulations of Section V.

B. An Exact Bayesian Detector

In order to make a comparison to the approximated Bayesian detector in the simulations in Section V, in this subsection we provide a set of Bayesian prior PDFs that result in a detector with a closed form.

Suppose σ^2 is known, and under $\mathcal{M}_i$, the i elements of $\mathbf{A}_i$ are independent and identically distributed, or

$$p\left(\mathbf{A}_i|\mathcal{M}_i\right) = \prod_{k=1}^{i} p\left(A_k\right) \tag{14}$$

Assume under any model $\mathcal{M}_i$, $1 \leq i \leq M$

$$p\left(A_k\right) = \frac{1}{\sqrt{2\pi\sigma_{A_k}^2}} \exp\left(-\frac{1}{2\sigma_{A_k}^2}\left(A_k - \mu_k\right)^2\right), \quad 1 \leq k \leq i \tag{15}$$

Denoting $\mathbf{x}_k = [X_k[0], X_k[1], \cdots, X_k[N-1]]^T$ as the data vector from channel k, because each channel is independent to other channels we have

$$\frac{p\left(\mathbf{X}|\mathcal{M}_i\right)}{p\left(\mathbf{X}|\mathcal{M}_0\right)} = \prod_{k=1}^{i} \frac{p\left(\mathbf{x}_k|\mathcal{M}_i\right)}{p\left(\mathbf{x}_k|\mathcal{M}_0\right)} \tag{16}$$

Denoting $\frac{p(\mathbf{x}_k|\mathcal{M}_i)}{p(\mathbf{x}_k|\mathcal{M}_0)}$ as $R\left(\mathbf{x}_k\right)$, with the assumed prior PDFs, after some algebra we have

$$R\left(\mathbf{x}_k\right) = \sqrt{\frac{\sigma^2}{N\sigma_{A_k}^2 + \sigma^2}} \exp\left(\frac{N^2\sigma_{A_k}^2\left(\bar{x}_k\right)^2 + 2N\mu_k\bar{x}_k\sigma^2 - N\mu_k^2\sigma^2}{2\sigma^2\left(N\sigma_{A_k}^2 + \sigma^2\right)}\right) \tag{17}$$

where $\overline{x}_k$ denotes the mean of the data from sensor k, given as

$$\bar{x}_k = \frac{1}{N}\sum_{n=0}^{N-1} \mathbf{X}_k\left[n\right]$$

From (10) and (16) the Bayesian detector decides H_1 if

$$\sum_{i=1}^{M} \pi_i \prod_{k=1}^{i} R\left(\mathbf{x}_k\right) > \lambda \tag{18}$$

V. Computer Simulations

In this section the previously discussed detectors are compared via computer simulations. We note, these detectors are derived from different origins and have quiet different properties. The performance of each detector requires further investigation. This section only serves to give some examples of these detectors and to show the GLRT can be improved.

A. Known Variance

Define the SNR of channel k as

$$\text{SNR}_k = 10\log 10\left(\frac{\varepsilon A_k^2}{\sigma^2}\right)$$

With this definition, when $\text{SNR}_k \geq 20\text{dB}$ all methods yield pretty good results. Therefore the main concern is the situation when $\text{SNR}_k < 20\text{dB}$, i.e. when $|A_k| < \sqrt{\frac{100\sigma^2}{\varepsilon}}$. Hence we assume $\sigma_{A_k}^2 =$

$\frac{100\sigma^2}{9\varepsilon}, 1 \leq k \leq i$ for $p(A_k)$ given by (15), so that $3\sigma_{A_k} = \sqrt{\frac{100\sigma^2}{\varepsilon}}$ and for most values that A_k will take the SNR$_k$ is less than 20dB. Since the SNR$_k$ for each sensor is the same, we simply denote it as SNR. We also assume $\mu_k = 0, 1 \leq k \leq i$ for $p(A_k)$ and the priors π_i, $1 \leq i \leq M$ are all equal.

Now, all the assumptions required for the exact Bayesian detector given by (17) and (18) are made. If all these assumptions are true, the exact Bayesian detector is optimal. However, since in practice the assumptions are rarely completely true, the Bayesian approach is only to find a realizable detector and we hope the detector will be robust. That is when the true priors and prior PDFs are not the same as the assumptions, the detector still has satisfactory performance. In the simulation the true priors under H_1 are set to equal, which is the same as the assumed priors. However, the true A_k takes fixed value, in other words A_k is deterministic.

Since the detection performance only depends on the energy of the signal waveform, we simply let $s[n] = 1$, $0 \leq n \leq N-1$ and $\sigma^2 = N$, so that $\frac{\varepsilon}{\sigma^2} = 1$. The number of sensors M is set to be 5. With the parameters, it can be shown that

$$L_{G_i}(\mathbf{X}) = \exp\left(\frac{1}{2}\left(\sum_{k=1}^{i} \bar{x}_i^2\right)\right)$$

We compare the receiver operating characteristics (ROC) curves of the GLRT given by (2), the rGLRT with model selected by the MDL, the MFLRT given by (7), the approximated Bayesian detector given by (13) and the exact Bayesian detector given by (17) and (18). When SNR$= 6$ or $A_k = 2$ the performances of detectors of 5000 realizations are shown in Fig. 1, and when SNR$= 12$ or $A_k = 4$ the performances of detectors of 5000 realizations are shown in Fig. 2. For relatively low SNR case, it can be seen in the Fig. 1 that the Bayesian detector, aBayesian detector and the MFLRT are better than the GLRT for all region of the ROC curve. The rGLRT are better than GLRT for the region of the ROC curve where $P_{FA} > 0.04$. For relatively high SNR case, it can be seen in the Fig. 2 that the GLRT is the worst for all region of the ROC curve. Comparing Fig. 1 and Fig. 2, it is seen that the aBayesian detector has very similar performance to the exact Bayesian detector. It is also seen from Fig. 1 and Fig. 2, when SNR is low, the rGLRT is close to the GLRT, which is the worst detector, and when SNR is high rGLRT is the best detector. Compared with the rGLRT, the MFLRT is more robust. Actually, it can be shown the MFLRT has some minimax properties, which will be addressed in a future work.

B. Unknown Variance

When the variance σ^2 is unknown, it can be shown

$$L_{G_i}(\mathbf{X}) = \left(\frac{\hat{\sigma}_0^2}{\hat{\sigma}_i^2}\right)^{\frac{MN}{2}}$$

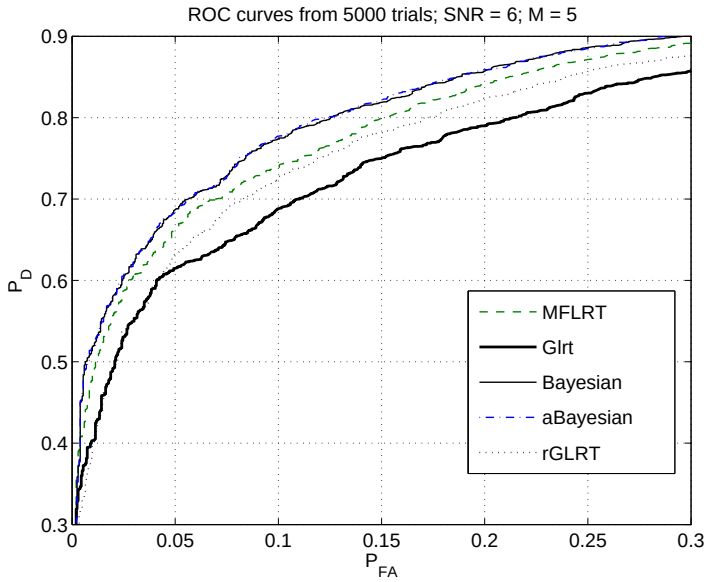

Fig. 1. ROC curves of the detectors in relatively low SNR (known variance).

.

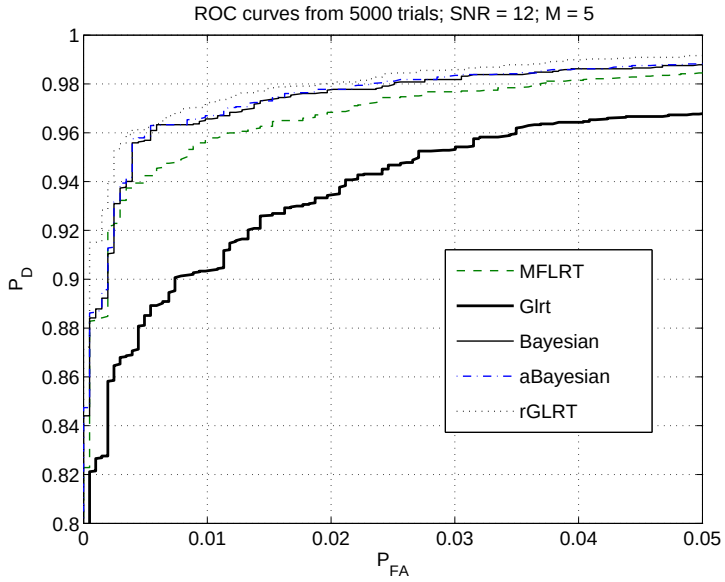

Fig. 2. ROC curves of the detectors in relatively high SNR (known variance).

.

where

$$\hat{\sigma}_k^2 = \frac{1}{NM} \sum_{n=0}^{N-1} \left(\sum_{i=1}^{k} \left(\mathbf{X}_i [n] - \bar{x}_i \right)^2 + \sum_{j=k+1}^{M} \left(\mathbf{X}_j [n] \right)^2 \right)$$

Because in this case, it is even more difficult to assume valid prior PDFs, we only compare the other 4 detectors. When SNR$=6$ the performances of detectors of 5000 realizations are shown in Fig. 3, and when SNR$=12$ the performances of detectors of 5000 realizations are shown in Fig. 4. For relatively

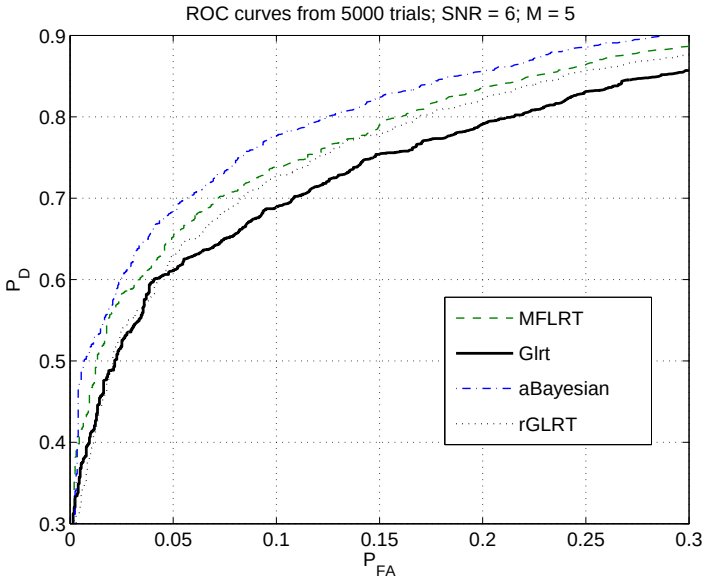

Fig. 3. ROC curves of the detectors in relatively low SNR (unknown variance).

.

low SNR case, it can be seen in the Fig. 3 that the aBayesian detector and the MFLRT are better than the GLRT for all region of the ROC curve. The rGLRT are better than the GLRT for the region of the ROC curve where $P_{FA} > 0.04$. For relatively high SNR case, it can be seen in the Fig. 4 that the GLRT is the worst for all region of the ROC curve. The minimax property of the MFLRT can also be observed comparing Fig. 3 and Fig. 4.

VI. Conclusions and Discussion

We have investigated the GLRT and the Bayesian approach for centralized composite distributed detection. In particular, In particular, we have proved that the performance of GLRT is poor and can be improved for this problem. An approximated Bayesian detector has also been proposed based on vague prior PDFs, and therefore the marginal PDF can be obtained without integration. As a result, the detector always has a closed form.

In our exposition, we assume the sensors have already been ordered with respect to their importance. Therefore when there are M sensors, only $M+1$ models need to be considered. When the original data

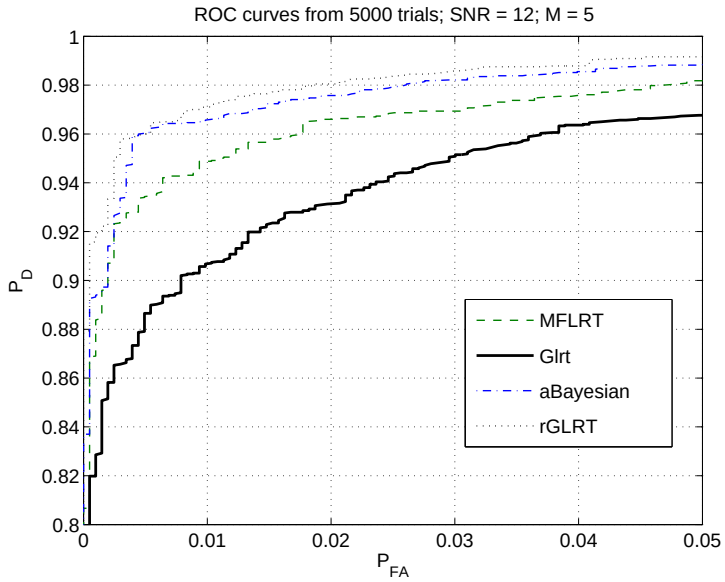

Fig. 4. ROC curves of the detectors in relatively high SNR (unknown variance).

.

from sensors are not ordered, for M sensors we have to consider 2^M models as in [4]. This will greatly increase the computational load. Alternatively, some preliminary processing can be conducted to order the sensors and similar techniques as in [11] can be used. However, the performance degradation as a result of ordering needs further investigation.

REFERENCES

[1] R. R. Tenney and N. R. Sandell Jr., "Detection with distributed snesors," *IEEE Trans. Aerospace Elect. Syst.*,, vol. AES-17, pp. 501-510, July 1981.

[2] R. Viswanathan and P. K. Varshney, "Distributed detection with multiple sensors-Part I: Fundamentals," *Proceedings of the IEEE*,, vol. 85, pp. 54-63, Jan. 1997.

[3] R. S. Blum, S. A. Kassam and H. V. Poor, "Distributed detection with multiple sensors-Part II: Advanced Topics," *Proceedings of the IEEE*,, vol. 85, pp. 64-79, Jan. 1997.

[4] B. Baygun and A. O. Hero III, "Optimal Simultaneous Detection and Estimation Under a False Alarm Constraint," *IEEE Trans. Inform. Theory*,, vol. 41, pp. 688-703, may 1995.

[5] S. M. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory*, Upper Saddle River, NJ: Prentice-Hall, 1998.

[6] P. Stoica and Y. Selen, "Model-Order Selection-A review of information criterion rules," *IEEE Signal Processing Magzine*,, vol. 21, pp. 36-47, July 2004.

[7] S. M. Kay, "The Multifamily Likelihood Ratio Test for Multiple Signal Model Detection," *IEEE Signal Process. Letters*,, vol. 12, no. 5, pp. 369-371, may 2005.

[8] L. L. Scharf and B. Friedlander, "Matched subspace detectors," *IEEE Trans. Signal Process.,*, vol. 46, no. 8, pp. 2146-2157, Aug. 1994.

[9] O. Zeitouni, J. Ziv and N. Merhav, "When is the generalized likelihood ratio test optimal," *IEEE Trans. Inform. Theory,*, vol. 38, pp. 1597-1602, Sept. 1992.

[10] P. M. Djuric, "Asymptotic MAP criteria for model selection," *IEEE Trans. Signal Process.,*, vol. 46, no. 10, pp. 2726-2735, Oct. 1998.

[11] R. R. Hocking, "The analysis and selection of variables in linear regression," *Biometrics,*, 32, pp. 1-49, Mar. 1976.

[12] A. Gelman, J. Carlin, H. Stern and D. Rubin *Bayesian Data Analysis*, London, U.K.: Chapman & Hall, 1995.

APPENDIX I
PROOF OF THEOREM 1

Since $\nu 1 < \nu 2$, consider a hypothesis testing problem, in which,

$$\begin{aligned} H_0 &: x\left[1\right] \sim \chi^2_{\nu 1}, \qquad x\left[2\right] \sim \chi^2_{\nu 2-\nu 1} \\ H_1 &: x\left[1\right] \sim \chi'^2_{\nu 1}\left(\lambda\right), \quad x\left[2\right] \sim \chi^2_{\nu 2-\nu 1} \end{aligned}$$

and $x[1]$ is independent of $x[2]$ in each hypothesis. The optimal NP detector decides H_1 if [5]

$$\frac{p_{\chi'^2_{\nu 1}(\lambda)}\left(x[1]\right) p_{\chi^2_{\nu 2-\nu 1}}\left(x[2]\right)}{p_{\chi^2_{\nu 1}}\left(x[1]\right) p_{\chi^2_{\nu 2-\nu 1}}\left(x[2]\right)} = \frac{p_{\chi'^2_{\nu 1}(\lambda)}\left(x[1]\right)}{p_{\chi^2_{\nu 1}}\left(x[1]\right)} > \gamma$$

Plugging in the expression of the PDF of $\chi^2_{\nu 1}$ and $\chi'^2_{\nu 1}\left(\lambda\right)$ yields [5]

$$\frac{\frac{1}{2}\left(\frac{x[1]}{\lambda}\right)^{\frac{\nu 1-2}{4}} \exp\left[-\frac{1}{2}\left(x[1]+\lambda\right)\right] \sum_{k=0}^{\infty} \frac{\left(\frac{1}{2}\sqrt{\lambda x[1]}\right)^{2k+\frac{\nu 1}{2}-1}}{k!\Gamma\left(\frac{\nu 1}{2}+k\right)}}{\frac{1}{2^{\frac{\nu 1}{2}}\Gamma\left(\frac{\nu 1}{2}\right)} x[1]^{\frac{\nu 1}{2}-1} \exp\left(-\frac{1}{2}x[1]\right)} > \gamma$$

The above expression can be simplified to

$$\sum_{k=0}^{\infty} c_k x[1]^k > \gamma$$

with all positive $c_k{}'s$. Since $x[1] \geq 0$, $\sum_{k=0}^{\infty} c_k x\left[1\right]^k$ is a monotonically non-decreasing function of $x[1]$, the detector decides H_1 equivalently if

$$T_{NP}\left(\left[x[1], x[2]\right]\right) = x[1] > \gamma' \tag{19}$$

This is the optimal detector in the NP sense. Now consider another detector which decides H_1 if

$$T_G\left(\left[x[1], x[2]\right]\right) = x[1] + x[2] > \gamma''$$

This detector is different from the optimal detector (19) and therefore it has a poorer performance. Since

$$T_{NP} \sim \begin{cases} \chi^2_{\nu 1}, & H_0 \\ \chi'^2_{\nu 1}\left(\lambda\right), & H_1 \end{cases}$$

and

$$T_G \sim \begin{cases} \chi^2_{\nu 2}, & H_0 \\ \chi'^2_{\nu 2}(\lambda), & H_1 \end{cases}$$

The theorem is proved by letting $T_1 = T_{NP}$ and $T_2 = T_G$.