Composite Hypothesis Testing with Intermittent Observations

Ashraf Tantawy, Xenofon Koutsoukos, and Gautam Biswas

Institute for Software Integrated Systems and the Department of Electrical Engineering and Computer Science, Vanderbilt University, Nashville, TN 37235 USA (e-mail: ashraf.tantawy; xenofon.koutsoukos;gautam.biswas@vanderbilt.edu)

Abstract: In this paper, we consider the detection problem with intermittent observations, due to the unreliable shared communication link between local sensors and the fusion center. Detection performance is analyzed using Neyman-Pearson criterion of maximizing the probability of detection, for a given probability of false alarm. The detector performance is compared, with and without intermittent observations, and a formal approach is presented to restore the original detector performance.

Keywords: NP Detection; Wireless sensor network; Passive sensor; Hypothesis testing; Z-test.

1. INTRODUCTION

In classical detection theory, statistical hypothesis testing is applied to detect noisy signals. The main problem addressed is to design the optimal detector (according to some pre-defined criteria) that distinguishes between two or more hypotheses (sometimes referred to as *phenomena* or state of nature), given noisy observations (Kay [1998]). In composite hypothesis testing, the hypotheses are not defined completely, i.e., the Probability Density Function (PDF) for the observed data sequence is not completely specified under each hypothesis. An example is the detection of the presence or absence of an event, like flammable gases in a process plant, excess vibrations in a structure, and intruders in a secured area.

Although the design of an optimal detector for a composite hypothesis testing problem is not always possible, sub-optimal detectors exist for some problem classes. A well-known example is the detection of a change in a known DC level in White Gaussian Noise (WGN), using Neyman-Pearson (NP) criterion. If the noise variance is known, the resulting detector design is a *Z-test*, while if the noise variance is unknown, the resulting design is a *t-test*. Both tests are widely used in practice. However, the main assumption in the design of these detectors is that all sensor measurements are available at the detector, which is no longer the case in Decentralized Detection (DD) applications that use a shared communication link to transmit local sensor measurements to a fusion center.

In shared communication links, multiple sensors contend to acquire channel access and submit their measurements. Therefore, collisions are unavoidable, and measurements may be lost. The detector then has to make a decision, using only a subset of the transmitted information from sensors, and a performance degradation occurs. Analysis of the detection performance is important to quantify the degradation, and to take a remedy action to restore the original performance. The shared communication link may be a wired communication cable or a wireless communication channel. In particular, information loss in Wireless Sensor Networks (WSNs) is significant due to the limited bandwidth available and the channel unreliability, caused by different sources of interference.

In decentralized detection using WSNs, multiple sensors relay information (after pre-processing) to a fusion center, and the problem is to design both the optimal local decision rules and the fusion rule to detect events as accurately as possible (Chamberland and V. Veeravalli [2007]). The detection problem in WSNs has the additional challenge of channel imperfection, which causes delays and missed observations. There is some recent work that discusses the effect of non-ideal channels on the design of decentralized detectors (Chen and Willett [2005], Chen et al. [2004]). However, the main assumption is that the sensor nodes have sufficient computational power to pre-process the observations and take a decision.

In the last few years, passive wireless sensors have emerged as a new technology for sensors that do not require any power source. The basic idea is to use a powered reader to poll the data from different passive sensors. The passive sensor uses the incident power from the reader to energize its local circuitry. A variety of ways exist for the sensor to modulate the reader incident wave and relay its information back to the reader. The most popular example for this architecture is RFID technology (Lewis [2004]).

The introduction of passive wireless sensors as a replacement for active sensors has changed the detection problem formulation. The main change comes from the fact that a passive sensor does not have power, and therefore, it cannot pre-process the measurements before transmission to the reader. In addition, because of the limited power supplied to the sensor, the reflected signal has very small power, which makes detection a much more challenging task especially if the noise power is large. Finally, in a passive WSN architecture, the reader polls the information from multiple sensors at the same time, giving rise to collisions and hence packet drops and delays.

In this paper, we study the effect of intermittent observations on the detection performance. Motivated by wireless passive sensors, we assume that observations are transmitted without preprocessing, over a shared communication link. We formulate the detection performance, assuming an Independent and Identically Distributed (IID) Bernoulli random process model for the communication link. We illustrate the degradation in the performance using a case study of the Z-test. The analytical results are verified by Monte Carlo Simulation studies, and we propose an approach for adaptive detector design, to restore the original detection performance.

The rest of the paper is organized as follows: In section 2, we briefly describe related work. In Section 3, we formulate the detection problem, with a review of some basic concepts about the composite hypothesis testing problem. In section 4, we formulate the detection performance with intermittent observations. Section 5 presents a case study for the Z-test. In Section 6, we propose an approach for adaptive detector design, to restore the original detector performance. Finally, in section 7, we present the conclusions and future research directions.

2. RELATED WORK

Classical detection theory is discussed in (Kay [1998]). The research on decentralized detection is largely attributed to the seminal work of Tenney and Sandell (Tenney and Sandell [1981]). The optimal decision rules for the local nodes and the fusion center are derived under various problem settings and different optimality criteria. For a more comprehensive survey in this area, the readers are referred to (Chamberland and V. Veeravalli [2007], Chen et al. [2006]) and the references therein.

Decentralized detection over unreliable communication links has been an active area of research in the last decade. The optimality of Likelihood Ratio (LR) test for local sensor decisions, with binary symmetric channel is proved in (Chen and Willett [2005]). The variations in the false alarm and detection probabilities, due to the errors caused by the communication link are studied in (Madishetty et al. [2005]). The distributed detection problem over multiple access channel is studied in (Li and Dai [2006]). For a comprehensive survey on the results in this field, the readers are referred to (Chamberland and V. Veeravalli [2007]).

In previous work, we addressed the simple hypothesis testing problem with intermittent observations, according to NP and Bayesian criteria (Tantawy et al. [2009a,b]). In this paper, we address the composite hypothesis testing problem, using Generalized Likelihood Ratio Test (GLRT) approach and NP criterion, with intermittent observations. Z-test, as an application example is thoroughly studied.

3. PROBLEM FORMULATION

Figure 1 illustrates the detection system architecture. The state of nature is sensed by a set of k local sensors, and the observations are transmitted over a shared communication link. Since we assume no local pre-processing for the

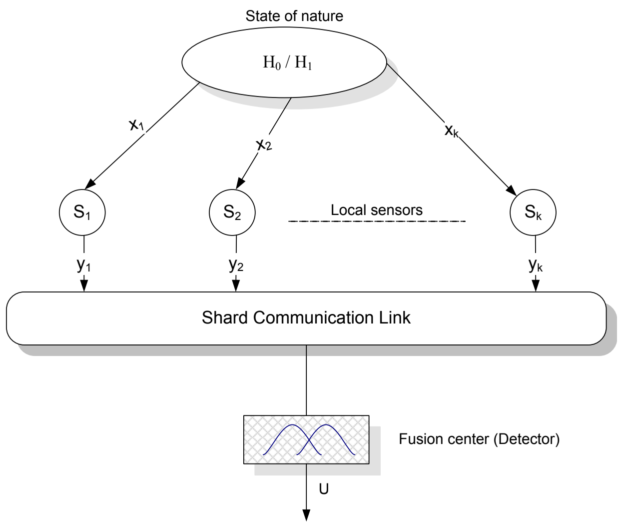

Fig. 1. Detection with intermittent observations. The shared communication link may represent a wired communication cable, or a wireless channel. The detector is to decide on one of two states of nature.

measurements at the local sensors, the transmitted observations will be equal to the sensed observations at each local sensor, i.e., $\mathbf{y}_k = \mathbf{x}_k$. The fusion center (detector) takes a final binary decision, U, based on the observation vector $\mathbf{y} = (\mathbf{y}_1, \mathbf{y}_2 \ldots \mathbf{y}_k)$. Due to missing observations, only a subset of the observation vector $\mathbf{y}$ is available at the detector. The objective of the study is to analyze the detection performance with intermittent observations, and to restore the original detection performance. The later objective is achieved by increasing the number of samples, in a proportional magnitude to the communication link drop rate.

We consider the composite hypothesis testing problem, where it is required to discriminate between two hypotheses, $\mathcal{H}_0$ and $\mathcal{H}_1$, given an observation vector, $\mathbf{x}$. The PDF under one (or both) hypothesis is not completely known, and depends on a set of unknown parameters, $\boldsymbol{\theta}_i$, for each hypothesis $i, i \in \{0, 1\}$. The PDF is expressed as:

$$\mathcal{H}_0 : p(\mathbf{x}; \boldsymbol{\theta}_0, \mathcal{H}_0)$$
$$\mathcal{H}_1 : p(\mathbf{x}; \boldsymbol{\theta}_1, \mathcal{H}_1)$$

The detector design problem is to find the region(s) in the multidimensional space of the observation vector $\mathbf{x}$, where we reject the null hypothesis, $\mathcal{H}_0$, and decide on the alternative hypothesis, $\mathcal{H}_1$. This region is called the critical region and denoted by R. There are different optimization criteria used to define the critical region. In NP criterion, it is required to maximize the probability of detection, P_D, given the probability of false alarm, P_{FA}, where:

$$P_{FA} = P(\text{decide } \mathcal{H}_1; \mathcal{H}_0) = P(\mathbf{x} \in R; \mathcal{H}_0)$$
$$P_D = P(\text{decide } \mathcal{H}_1; \mathcal{H}_1) = P(\mathbf{x} \in R; \mathcal{H}_1)$$

The relationship between P_D and P_{FA} is called the Receiver Operating Characteristics (ROC) curve.

There are two approaches to composite hypothesis testing problems; the Bayesian approach and the Generalized

Likelihood Ratio Test (GLRT) (Kay [1998]). Since GLRT is more widely used due to its ease of implementation, and its asymptotic optimality properties, we use it in this paper.

In GLRT, the unknown parameter is considered a deterministic quantity, and replaced by its maximum likelihood estimate (MLE), and the likelihood ratio is expressed as:

$$L(\mathbf{x}) = \frac{p(\mathbf{x}; \hat{\theta}_1, \mathcal{H}_1)}{p(\mathbf{x}; \hat{\theta}_0, \mathcal{H}_0)}$$

The likelihood ratio is calculated and compared to a threshold value, which is determined using the NP criterion for maximizing P_D, given a fixed value for P_{FA}.

Z-test is an example of a composite hypothesis testing problem, according to NP criterion, where it is assumed that the signal to be detected is a known DC level embedded in WGN. The decision to be taken is whether the DC level is equal to the known value or not. Therefore, we have a simple main hypothesis and a composite alternative hypothesis. The Z-test detection problem is explained in section 5.

In the centralized detection framework, the existence of the unreliable communication link is ignored, and, therefore, perfect knowledge about the observations is assumed. For the work presented in this paper, we assume the communication link is modeled as an IID Bernoulli random process, with a communication link drop rate, λ, which is assumed known from statistical measurements. Therefore, the communication link is characterized by a random variable C with a Probability Mass Function (PMF):

$$P_C(c = 0) = \lambda$$
$$P_C(c = 1) = 1 - \lambda$$

4. NP DETECTION WITH BERNOULLI CHANNEL

In classical centralized detection, the detector is designed based on a fixed number of observations N, received in a time period T, to achieve a required performance (P_{FA}, P_D). With the shared communication link, not all observations are received at the detector. Therefore, in a time period T, the number of observations received k is a random variable, with PMF:

$$P(K = k) = \sum_{k=0}^{N} \binom{N}{k} (1-\lambda)^k \lambda^{(N-k)}$$

If we designate the probability of false alarm and the probability of detection, when receiving k observations, by P_{FA}^k and P_D^k, respectively, then the detection performance is defined by the pair of expected values $(E_K[P_{FA}^k], E_K[P_D^k])$:

$$E_K[P_{FA}^k] = \sum_{k=0}^{N} \binom{N}{k} (1-\lambda)^k \lambda^{(N-k)} P_{FA}^k \qquad (1)$$

$$E_K[P_D^k] = \sum_{k=0}^{N} \binom{N}{k} (1-\lambda)^k \lambda^{(N-k)} P_D^k \qquad (2)$$

To overcome the performance degradation, one way is to increase the number of samples N, which can be increased by increasing the decision time period T. Section 6 illustrates the approach, where it is shown that the performance is restored, with the penalty of increasing the delay for detection.

5. Z-TEST WITH BERNOULLI CHANNEL

5.1 Detector Design

The problem of detecting a change in a known DC level, assuming WGN, can be formulated by the following composite hypothesis testing problem:

$$\mathcal{H}_0 : x[n] = A_0 + w[n] \qquad n = 0, 1, \ldots, N-1$$
$$\mathcal{H}_1 : x[n] = A_1 + w[n] \qquad n = 0, 1, \ldots, N-1$$

A_0 represents the normal DC level, while $A_1 \neq A_0$ is not known a priori. Therefore, hypothesis $\mathcal{H}_1$ is not completely defined, and the PDF under $\mathcal{H}_1$ is not completely known. In fact, this hypothesis testing problem could be expressed as the parameter testing problem:

$$\mathcal{H}_0 : A = A_0$$
$$\mathcal{H}_1 : A \neq A_0$$

In this section we assume that the noise variance is known, and denoted by σ^2. Using the GLRT approach, we replace the unknown parameter, A, by its ML estimate. The MLE of A_1 is $\hat{A} = \bar{x}$, the sample mean (Kay [1998]). Substituting for the MLE for A, we can express the likelihood ratio as:

$$L(\mathbf{x}) = \frac{\frac{1}{(2\pi\sigma^2)^{\frac{N}{2}}} e^{-\frac{1}{2\sigma^2}\sum_{n=0}^{N-1}(x[n]-\bar{x})^2}}{\frac{1}{(2\pi\sigma^2)^{\frac{N}{2}}} e^{-\frac{1}{2\sigma^2}\sum_{n=0}^{N-1}(x[n]-A_0)^2}}$$

Taking the logarithm we get the detector design:

$$|\bar{x} - A_0| \geq \gamma \qquad 0 \leq \gamma < \infty \qquad (3)$$

5.2 Detection Performance

To calculate the performance of the detector in Equation (3), we calculate P_{FA} and P_D, for a given sample size k, as follows:

$$P_{FA}^k = P(\bar{X} - A_0 \geq \gamma; \mathcal{H}_0) + P(\bar{X} - A_0 \leq -\gamma; \mathcal{H}_0)$$
$$P_D^k = P(\bar{X} - A_0 \geq \gamma; \mathcal{H}_1) + P(\bar{X} - A_0 \leq -\gamma; \mathcal{H}_1)$$

We note that the sample mean, $\bar{X}$, is a Gaussian random variable with mean $A_i, i = 0, 1$, and variance σ^2/k. It is straightforward to show that:

$$P_{FA}^k = 2Q\left(\frac{\gamma}{\sqrt{\sigma^2/k}}\right) \qquad (4)$$

$$P_D^k = Q\left(\frac{\gamma - D}{\sqrt{\sigma^2/k}}\right) + Q\left(\frac{\gamma + D}{\sqrt{\sigma^2/k}}\right) \qquad (5)$$

where $Q(.)$ is the error function defined by:

$$Q(x) = \int_x^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2} dt,$$

and $D = A_1 - A_0$, represents the distance between the two mean values.

From Equations (4) and (5) we get:

$$P_D^k = Q\left(Q^{-1}\left(\frac{P_{FA}^k}{2}\right) - \sqrt{\frac{kD^2}{\sigma^2}}\right) + Q\left(Q^{-1}\left(\frac{P_{FA}^k}{2}\right) + \sqrt{\frac{kD^2}{\sigma^2}}\right) \quad (6)$$

From Equation (6), the detection performance is symmetric with respect to D and independent on the specific values A_0 and A_1. Also, the performance increases with (kD^2/σ^2). This quantity is referred to as the *deflection coefficient*, d^2. Therefore, to improve the detection performance, either the signal to noise ratio has to be high (here the signal represents the difference between the two DC levels), or k should be increased, which implies a larger delay.

Now we assume a shared communication link, modeled as an IID Bernoulli random process. From Equations (1,2,4,5) we get:

$$E[P_{FA}] = 2\sum_{k=0}^{N}\binom{N}{k}(1-\lambda)^k\lambda^{(N-k)}Q\left(\frac{\gamma}{\sqrt{\sigma^2/k}}\right) \quad (7)$$

$$E[P_D] = \sum_{k=0}^{N}\binom{N}{k}(1-\lambda)^k\lambda^{(N-k)}\left[Q\left(\frac{\gamma-D}{\sqrt{\sigma^2/k}}\right) + Q\left(\frac{\gamma+D}{\sqrt{\sigma^2/k}}\right)\right] \quad (8)$$

We note that Equations (7) and (8) reduce to Equations (4) and (5), respectively, when $\lambda = 0$ $(k = N)$.

Example 1. We assume it is required to design a Z-test with parameters $A_0 = 0, A_1 = 0.4, \sigma = 1, N = 100$, with the constraint $P_{FA} = 0.05$. From Equations (4) and (5) we get $\gamma = 0.196, P_D = 0.9793$, which defines the detector operating point $(0.05, 0.9793)$. Now the detector is designed with $\gamma = 0.196$. Assuming a shared communication link modeled as an IID Bernoulli random process, with $\lambda = 0.2$, and using Equations (7) and (8), we get the new detector performance as $(0.08, 0.9658)$. Therefore, with the shared communication link, the detector works with higher probability of false alarm and lower probability of detection.

5.3 ROC Curve

Equations (7) and (8) represent the ROC curve for the Z-test detector with intermittent observations. Using detector parameters in Table 1, the ROC curve is plotted in Figure 2 for different values of λ, including the ideal case $(\lambda = 0)$.

Table 1. Original Detector Parameters

Parameter	A_0	A_1	σ	N
Value	0	0.4	1	50

To verify the analytical results, a Monte Carlo simulation experiment is conducted. Table 1 lists the parameters used in the simulation experiment. Samples are generated from two different Gaussian distributions (corresponding to the two hypotheses), and a Bernoulli channel is introduced in the signal path to the detector. The detector calculates a running average, while ignoring dropped observations. The detector compares the running average to a threshold value, γ, that varies from 0 to ∞, to generate the complete ROC curve. For each value of the detector threshold, 5000

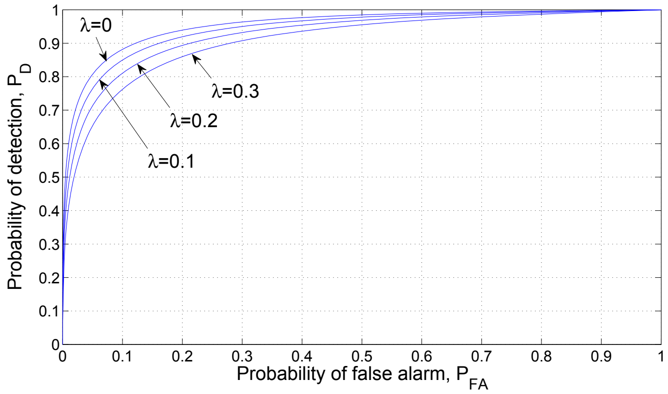

Fig. 2. ROC curve-Z-test with intermittent observations. The performance degrades with the increasing percentage of channel drop rate, λ.

Monte Carlo trials were performed to get accurate values for P_{FA} and P_D.

Figure 3 shows the theoretical ROC curve (Equations (7) and (8)) versus the ROC curve obtained from Monte Carlo simulation, for $\lambda = 0$ and $\lambda = 0.3$. As illustrated in the figure, the two ROC curves for each value of λ are very similar, due to the high number of Monte Carlo Simulation runs.

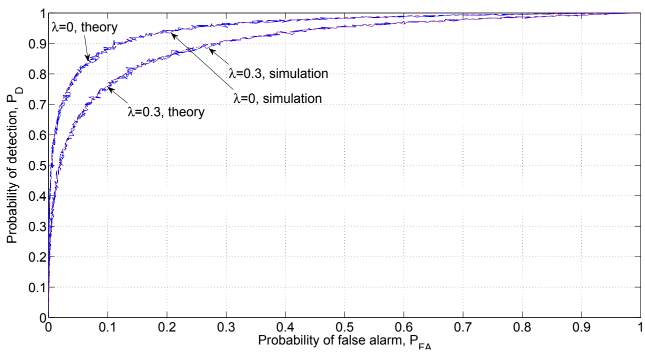

Fig. 3. ROC curve-Z-test with intermittent observations-Monte Carlo simulations and analytical results.

The degradation in the performance is clear from the new ROC curve. For a fixed P_{FA}, the original P_D cannot be obtained using the same number of observations. This is intuitive, since the detector is working with a subset of the information available in the case of an ideal communication link. In order to restore the original performance, additional information has to be acquired from local sensors. This approach is explained in Section 6.

6. ADAPTIVE DETECTOR DESIGN

To compensate for the degradation in the detection performance, one approach is to increase the number of total observations used for decision-making, by increasing the decision period T. However, since the number of observations received, k, is a random variable, we need different number of additional observations, N_k, for each realization k. The main idea behind using different values of N_k for different realizations k is to allow for more observations (and hence, more delay) for improbable values of k, while allowing less delay for highly probable values of k. This enables us to minimize the expected value of the added delay needed to restore the performance. Noting that we have two constraints on the detector performance, in terms of $E[P_{FA}], E[P_D]$, the optimization problem can be formulated as follows:

$$\begin{cases} \min\limits_{P_{FA}^k} E[N_k] \\ \text{subject to} \\ E[P_{FA}] = \alpha, \\ E[P_D] = \beta \end{cases}$$

It should be highlighted that waiting for N_k observations, for every realization k, is not equivalent to receiving N_k observations. This is due to the unreliability of the communication link. For example, if the detector waits for a time equivalent to receiving $N_k = 3$ observations, then it may get 0, 1, 2 or 3 observations. Accordingly, when deriving an expression for the relationship between $E[N_k]$ and P_{FA}^k, the unreliable channel effect should also be taken into account for the added observations.

7. CONCLUSIONS AND FUTURE WORK

In this paper we have addressed the problem of composite hypothesis testing, using GLRT approach and NP criterion, assuming intermittent observations between local sensors and the detector. Shared communication link is modeled by an IID Bernoulli random process. It is shown that the detector performance is degraded by an amount proportional to the channel drop rate. An adaptive detector is shown to be useful to restore the original performance, with the penalty of increasing the delay for detection.

We are currently working on the solution of the optimization problem, presented in Section 6, to find the minimum delay required for performance restoration. We are also investigating the use of higher order Markov models, as a modeling approach for the shared communication link.

ACKNOWLEDGMENT

This work was supported in part by the National Science Foundation under Grant CNS-06152140, and NASA NRA NNX07AD12A.

REFERENCES

Jean-Francois Chamberland and Venugopal V. Veeravalli. Wireless sensors in distributed detection applications. *IEEE Signal Processing Magazine*, 2007.

Biao Chen and Peter K. Willett. On the optimality of the likelihood-ratio test for local sensor decision rules in the presence of nonideal channels. *IEEE Transactions on Information Theory*, 51(2):693–699, Feb. 2005. ISSN 0018-9448. doi: 10.1109/TIT.2004.840879.

Biao Chen, Ruixiang Jiang, Teerasit Kasetkasem, and Pramod K. Varshney. Channel aware decision fusion in wireless sensor networks. *IEEE Transactions on Signal Processing*, 52(12):3454–3458, Dec. 2004. ISSN 1053-587X. doi: 10.1109/TSP.2004.837404.

Biao Chen, Lang Tong, and Pramod K. Varshney. Channel-aware distributed detection in wireless sensor networks. *IEEE Signal Processing Magazine*, July 2006.

Steven M. Kay. *Fundamentals of Statistical Signal Processing, Volume 2: Detection Theory.* Prentice Hall Signal Processing Series. Prentice Hall PTR, 1998.

Steve Lewis. A basic introduction to RFID technology and its use in the supply chain. Technical report, LARAN RFID, 2004.

Wenjun Li and Huaiyu Dai. Distributed detection of a deterministic signal in correlated gaussian noise over mac. *IEEE International Symposium on Information Theory*, pages 2134–2138, July 2006. doi: 10.1109/ISIT.2006.261914.

M. Madishetty, V. Kanchumarthy, R. Viswanathan, and C.H. Gowda. Distributed detection with channel errors. *Proceedings of the Thirty-Seventh Southeastern Symposium on System Theory, SSST '05.*, pages 302–306, March 2005. ISSN 0094-2898. doi: 10.1109/SSST.2005.1460926.

Ashraf Tantawy, Xenofon Koutsoukos, and Gautam Biswas. Detection using intermittent observations for passive wireless sensors. *American Control Conference, ACC2009*, 2009a.

Ashraf Tantawy, Xenofon Koutsoukos, and Gautam Biswas. Maximum likelihood detection with intermittent observations. *IEEE Conference on Information Sciences and Systems*, 2009b. to appear.

Robert R. Tenney and Nils R. Sandell. Detection with distributed sensors. *IEEE Transactions on Aerospace and Electronic Systems*, AES-17(4):501–510, July 1981. ISSN 0018-9251. doi: 10.1109/TAES.1981.309178.