DISTRIBUTED DETECTION ALGORITHMS FOR PARALLEL AND HIERARCHICAL WIRELESS SENSOR NETWORKS

by

KAYHAN ERİTMEN

Submitted to the Graduate School of Engineering and
Natural Sciences in partial fulfillment of
the requirements for the degree of
Master of Science

Sabancı University
August 2008

DISTRIBUTED DETECTION ALGORITHMS FOR PARALLEL AND HIERARCHICAL WIRELESS SENSOR NETWORKS

APPROVED BY:

Associate Prof. Dr. Mehmet Keskinöz (Thesis Advisor): ....

Associate Prof. Dr. Mustafa Ünel: ....

Assistant Prof. Dr. Hakan Erdoğan: ....

Assistant Prof. Dr. İlker Hamzaoğlu: ....

Assistant Prof. Dr. Yücel Saygın: ....

DATE OF APPROVAL: ……………….

ABSTRACT

DISTRIBUTED DETECTION ALGORITHMS FOR PARALLEL AND HIERARCHICAL WIRELESS SENSOR NETWORKS

Wireless Sensor Networks (WSNs) have recently attracted a lot of attention in various potential applications in military, health, environment and commerce due to their detection, processing and communication capabilities. In this thesis, we consider distributed detection problem for both parallel and hierarchical topology in which sensor decisions are sent over non-ideal wireless channels. We first investigate optimal fusion rules in Neyman-Pearson sense for all considered network configuration. We then, suggest suboptimal fusion rules to decrease computational complexity of the optimal fusion rules. Thirdly, multi-bit distributed detection is investigated both analytically and numerically to increase the detection performance. Finally, we propose fusion center diversity by employing multiple antennas at the fusion center to improve the detection performance of the network and derive optimum fusion rules accordingly. Simulation results suggest that fusion center diversity increases the probability of detection for a given constant false alarm probability.

ÖZET

PARALEL VE HİYERARŞİK TELSİZ DUYARGA AĞLARI İÇİN DAĞITIK SEZİMLEME ALGORİTMALARI

Telsiz Duyarga Ağları (TDA), sezileme, bilgi işleme ve haberleşme kabiliyetleri sayesinde, askeri, sağlık, çevre ve ticari alanlardaki olası uygulamalarıyla ilgi görmektedir. Bu tezde, TDAlardaki duyargaların kararlarını ideal olmayan telsiz kanallardan ilettikleri dağıtık sezimle problemini, paralel ve hiyerarşik toplojiye sahip ağlar için inceledik. İlk olarak ilgilenilen topolojiler için en iyi tümleştirme kurallarını Neyman-Pearson sezimleme kriteri altında inceledik. Daha sonra, en iyi kuralların işlemsel karmaşıklığını azaltmak için alt-en iyi altı tümleştirme kuralları önerdik. Üçüncü olarak sezim başarımını artırmak için çoklu-bit kullanılarak yapılan dağıtık sezimleme analitik ve nümerik olarak incelendi. Son olarak, sezimleme başarımını artırmak için tümleştirme merkezinde çoklu anten kullanan tümleştirme merkezi çeşitlemesini önerdik ve bunun için en iyi tümleştirme kuralını türettik. Benzetim sonuçları tümleştirme merkezi çeşitlemesinin verilen sabit yanlış hata olasılığı için sezimleme olasılığını artırdığını göstermektedir.

ACKNOWLEDGEMENTS

I would like to offer my thanks and gratitude to my thesis advisor Mehmet Keskinöz for his technical and psychological support throughout this work.

I also would like to thank Mustafa Ünel, Hakan Erdoğan, İlker Hamzaoğlu and Yücel Saygın for reading and commenting on this thesis.

Throughout my master studies, I am financially supported by TUBITAK which helped me very much to concentrate on my thesis.

I would like to thank my colleagues, housemates and members of Sabancı University for creating an entertaining and inspirational atmosphere. Also, I would like to thank Ayşe, Burcu, Ceren and Gizem for their cheerful talks in my exhausted times. Last but not the least; I am grateful to my beloved family for their endless support during my graduate study.

Dedicated to my beloved mother…

TABLE OF CONTENTS

TABLE OF FIGURES

1. INTRODUCTION

1.1 Wireless Sensor Networks

Sensors with different sizes and capabilities are being used in military, health or daily life applications. Developments in network protocols, new methods in wireless communications, micro level production and the advances in microprocessor design triggered the research in the area of wireless sensor networks (WSN) [1]. A Wireless Sensor Network consists of a large number of small, cheap and low powered units, namely: sensors. Sensors in network are positioned densely in or around the target phenomenon in a random manner and they collaboratively try to get information about the interested phenomenon. WSN has some features that distinct them from conventional wireless networks [1];

- WSN may have much more nodes compared to ordinary wireless network.
- Sensor node density is very high in WSNs (20 nodes/m^3)
- Some of the sensors may not operate after deployed
- WSNs hardware capabilities are restricted; for example, they have low battery power, memory and processor capacity.
- WSNs have no static topology based on identification no (ID)
- Sensor nodes have to use broadcast while communicating.

WSNs begin to operate after deployment to interested region and sensor nodes in network begin to collect information about the environment to sink. There are some metrics for evaluating the performance of a WSN [2]. These metrics are flexibility, robustness, security, communication, computation, time synchronization, size, cost and power. Actually almost all metrics points the importance of decreasing power consumption of the WSN.

Each sensor in WSN has the capability to sense, process and communicate. This leads to many interesting potential applications [2]. These applications can be grouped as

- Environmental Applications
    - Detection and informing of a natural disaster or forest fire
    - To collect information about air pollution
    - Monitoring natural life
- Health Applications
    - Distant-monitoring of physical data of humans
    - Tracking the medication of patients
    - Tracking patients and doctors in a facility
- Commercial applications
    - Security issues
    - Tracking kids by their families
    - Car tracking and detecting
- Military Applications
    - Battlefield surveillance
    - Expedition of an unknown area
    - Detecting position and velocity of target
    - Being informed about nuclear, biological and chemical attack.

In all scenarios, each unit in the WSN sends its observations, collected from the environment, to another unit capable of fusing this information which is called fusion center. Local sensor nodes can transmit their observations without processing or every local sensor node makes a hard decision relies on its observation, and these decisions are huddled together in the fusion center; these are called Central Detection and Distributed Detection respectively [3]. Transmission of raw data is too demanding. Due to obligation of sending too much data, too much power and bandwidth is consumed by local sensor nodes and this is a performance decreasing issue according to the performance evaluation criteria. Also many of the local sensor nodes will send same data; this will cause excessive information in the fusion center. Due to these reasons distributed detection in WSNs is pretty effective.

1.2 Distributed Detection in Wireless Sensor Networks

Distributed Detection is a cooperative detection scheme that many detectors join together to make a final decision between two or more hypotheses. Therefore; we can have distributed detection model as a fusion center have to make the final decision and other detectors send their decision to the fusion center via communication link. To avoid wasting bandwidth, these detectors should quantize their observations first and then send their decisions to the fusion center. Main problem in distributed detection is to design fusion and detection rules in global fusion center and the other detectors, namely local sensor nodes. Optimal fusion rule which joins information coming from local sensor nodes at the fusion center is derived in [4] and the optimality of the likelihood ratio test (LRT) in local sensor nodes and fusion center under conditional independence assumption is shown in [5] under Neyman-Pearson criterion. A Bayesian approach for distributed detection of a phenomenon is investigated in [6]. Obtaining optimality of LRT does not explain how to find an optimum threshold for local sensor nodes. In distributed detection systems local sensor nodes does not behave like isolated detection systems, their performance jointly affect the system performance so, person by person optimization approach, where one local sensor node optimizes its decision rule while other nodes' and fusion center's fusion rule is fixed, is used . If local sensor nodes have dependent observations, fusion rules become difficult to solve and they do not turn into LRT [7]. Up to now most of the works in literature assumes error free communication but finally Thomopoulos and Zhang analyzed the case of distributed detection with non-ideal channel. In this thesis, distributed detection problem in WSN under fading channels is investigated; optimum and suboptimum fusion rules are derived for different cases and performance of these fusion rules is given via numerical simulations.

1.3 Scope and Contribution of Thesis

This thesis scope is as follows:

Chapter 2 will give brief information about distributed detection in wireless sensor networks with paralell topology considering Rayleigh faded communication channels. Optimum and sub-optimum fusion rules are given and and their detection performance evaluation for a given constant average channel SNR and constant false alarm probability are shown via numerical simulations.

In chapter 3 we will analyze binary distributed detection strategies in wireless sensor networks with hierarchical topology. We derived the optimum and suboptimum fusion rules for cases channel state information (CSI) or only channel statistics (CS) is available at the fusion center. We derive probability distribution function of fusion statistics assuming CS is available at the fusion center. We give the detection performance of derived fusion rules for constant signal to noise ratio (SNR) and for constant false alarm rate.

In chapter 4, multi-bit decision case is investigated for both parallel and hierarchical topology. 2-bit decision case is described as an example; fusion rules are derived for this 2-bit case and their performance evaluations are given via numerical simulations.

In chapter 5 we explore the effects of receive diversity at fusion center on the detection performance for parallel topology. We model the communication between local sensor nodes and fusion center as multiple input multiple output (MIMO) communication and derive the optimum fusion rule. For different number of antennas and local sensor nodes detection performance is given for constant average SNR and false alarm rate at fusion center.

We conclude and give possible future work in chapter 6.

2. PARALLEL DISTRIBUTED DETECTON for WIRELESS SENSOR NETWORKS UNDER FADING CHANNELS

Distributed detection problem, as mentioned before, in wireless sensor networks was investigated with Bayes and Neyman-Pearson (N-P) methods in [3], [6] and [7] comprehensively. However, in all of these works, the fading effect of the channel during local decision transmission to the fusion center was not considered in the fusion during detection. Chen et al [9] proposed a parallel distributed detection method based on Neyman-Pearson considering fading effect of the channel. They derived the likelihood ratio test (LRT) based optimum fusion rule and suboptimum fusion rules by simplifying the optimum one. Bahceci et al [10] proposed a method to find optimum threshold at the fusion center for a desired false alarm probability for parallel network topology. In this chapter, N-P based distributed detection under fading channels is recalled for parallel WNSs. The rest of the chapter is organized as follows: fusion rules based on channel state information (CSI) is investigated in section 2.1 and fusion rules when only channel statistics (CS) is available are derived in section 2.2.

2.1 Fusion Rules When Channel State Information is Available

In this section, parallel distributed detection with fading channel based on N-P decision criterion is summarized which is analyzed in mere details in [10]. Figure 2.1 has the conventional schema of parallel distributed detection in wireless sensor networks. Two hypotheses are assumed for detection, which are null hypothesis (no phenomenon- H_0) and the other is alternative hypothesis (phenomenon-H_1). S_j , stands for local sensor nodes, z_j shows the observation of j^{th} sensor and there is a wireless channel between local sensor nodes and fusion center. h_j is the Rayleigh distributed channel gain between the j^{th} local sensor node, x_j is the binary phase shift keying (BPSK) modulated signal which can take the values of -1 or 1 if local sensor decision u_j is 0 or 1 respectively. n_j is the additive white Gaussian noise with zero mean and variance of σ^2, and r_j shows the signal coming from the local sensor.

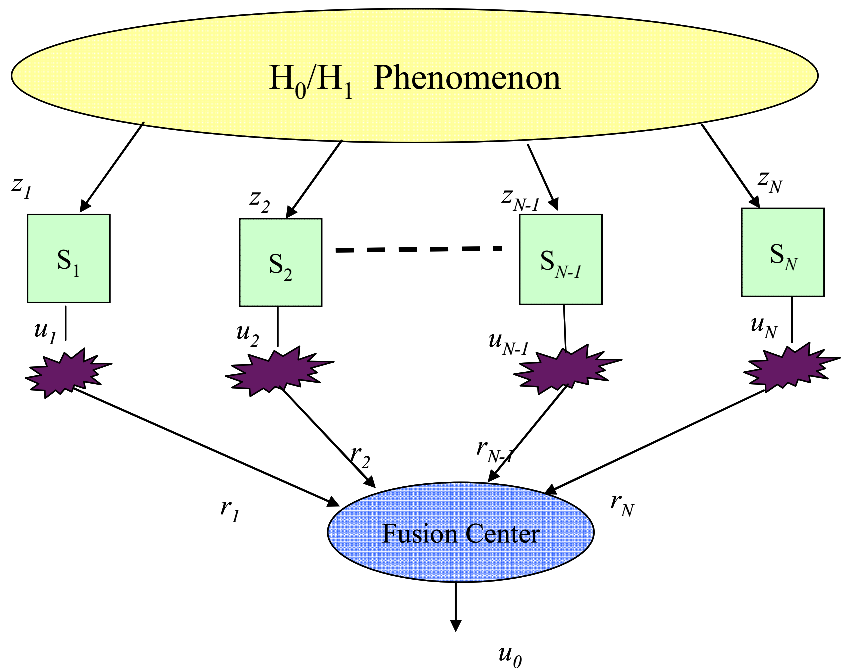

Figure 2.1 Parallel Distributed Detection Schema

2.1.1 Optimum Fusion Rule for Wireless Sensor Networks with Parallel Topology

According to Neyman Pearson under the conditional independence assumption, the optimal detection rule in fusion centre is as follows,

$$\Lambda(\mathbf{r})=\prod_{j=1}^{N}\frac{f(r_j \mid H_1,h_j)}{f(r_j \mid H_0,h_j)}=\prod_{j=1}^{N}\frac{P_{D_j}\,\mathrm{e}^{-\frac{(r_j-h_j)^2}{2\sigma^2}}+\left(1-P_{D_j}\right)\mathrm{e}^{-\frac{(r_j+h_j)^2}{2\sigma^2}}}{P_{F_j}\,\mathrm{e}^{-\frac{(r_j-h_j)^2}{2\sigma^2}}+\left(1-P_{F_j}\right)\mathrm{e}^{-\frac{(r_j+h_j)^2}{2\sigma^2}}} \tag{2.1}$$

defined as a type of likelihood ratio test. In (2.1) P_{D_j} and P_{F_j} are detection and false alarm probability of j^{th} local sensor node and can be expressed as

$$\begin{aligned} P_{D_j} &= \Pr\left(u_j = 1 \middle| H_1\right) \\ P_{F_j} &= \Pr\left(u_j = 1 \middle| H_0\right) \end{aligned} \tag{2.2}$$

and r_j shows the signal coming from the local sensor and represented as,

$$r_j = h_j x_j + n_j \tag{2.3}$$

2.1.2 Sub-Optimum Fusion Rules for Wireless Sensor Networks with Parallel Topology

Optimum fusion rule has the best detection performance for parallel topology as stated in [9] however it needs performance indices (P_{D_j} and P_{F_j}) and channel state information.

Therefore, some sub-optimum fusion rules are driven by analyzing low and high SNR behaviors of optimum fusion rule. For high SNRs we can re-express the optimal fusion rule as

$$\prod_{r_j<0}\frac{P_{D_j}+\left(1-P_{D_j}\right)\mathrm{e}^{-\frac{2r_jh_j}{\sigma^2}}}{P_{F_j}+\left(1-P_{F_j}\right)\mathrm{e}^{-\frac{2r_jh_j}{\sigma^2}}}\times\prod_{r_j>0}\frac{P_{D_j}\,\mathrm{e}^{\frac{2r_jh_j}{\sigma^2}}+\left(1-P_{D_j}\right)}{P_{F_j}\,\mathrm{e}^{\frac{2r_jh_j}{\sigma^2}}+\left(1-P_{F_j}\right)} \tag{2.4}$$

Since we are dealing with high-SNR channels we can conclude $\sigma^2 \to 0$ therefore $e^{\pm\frac{2r_jh_j}{\sigma^2}} \gg 1$ when the simplifications done we can obtain

$$\Lambda_{highSNR}(\mathbf{r}) = \sum_{sign(r_j)=-1} \log\left(\frac{1-P_{d_j}}{1-P_{f_j}}\right) + \sum_{sign(r_j)=1} \log\left(\frac{P_{d_j}}{P_{f_j}}\right) \tag{2.5}$$

This fusion rule is so called Chair-Varshney rule in [4] and does not need channel state information, it only use performance indices of local sensor nodes for detection but it has poor performance for low channel SNR's. When low-SNR case analyzed we should rewrite the optimum fusion as

$$\Lambda(\mathbf{r}) = \prod_{j=1}^{N} \frac{P_{D_j} + \left(1-P_{D_j}\right) e^{-\frac{2r_jh_j}{\sigma^2}}}{P_{F_j} + \left(1-P_{F_j}\right) e^{-\frac{2r_jh_j}{\sigma^2}}} \tag{2.6}$$

To examine low SNR region we can assume $\sigma^2 \to \infty$ and $e^{-\frac{2r_jh_j}{\sigma^2}} \to 1$. Therefore we can use first order Taylor series expansion $e^{-\frac{2r_jh_j}{\sigma^2}} \approx \left(1 - 2r_jh_j/\sigma^2\right)$ in (2.6) we get

$$\Lambda(\mathbf{r}) = \prod_{j=1}^{N} \frac{P_{D_j} + \left(1-P_{D_j}\right)\left(1-2r_jh_j/\sigma^2\right)}{P_{F_j} + \left(1-P_{F_j}\right)\left(1-2r_jh_j/\sigma^2\right)} \tag{2.7}$$

Taking logarithm of both sides and letting $\sigma^2 \to \infty$ as in [9] we have

$$\Lambda_{lowSNR}(\mathbf{r}) = \sum_{j=1}^{N} \left(P_{d_j} - P_{f_j}\right) h_j r_j \tag{2.8}$$

When all performance indices assumed to be same (2.8) becomes

$$\Lambda_{MRC}\left(r\right)=\sum_{j=1}^{N}h_j r_j \tag{2.9}$$

This fusion rule is in the form of maximum ratio combining (MRC) and it needs only channel state information as prior information. Although, Chair-Varshney rule has significantly bad performance for low SNR case, it requires only performance indices which are fixed, but MRC fusion rule needs instantaneous prior information about channel's gain; it should be noted as a power consuming issue.

2.2 Fusion Rules When Only Channel Statistics is Available

A WSN has very limited energy and bandwidth therefore it is not preferable trying to estimate channel gain at each transmission. Therefore, in this section, we try to derive fusion rules that use only channel statistics at the expense of small performance degradation.

2.2.1 Optimum Fusion Rule with Channel Statistics for Wireless Sensor Networks with Parallel Topology

First of all we have to write likelihood as

$$\Lambda(\mathbf{r})=\frac{f(\mathbf{r}\,|\,H_1)}{f(\mathbf{r}\,|\,H_0)}=\prod_{j=1}^{N}\frac{f(r_j\,|\,H_1)}{f(r_j\,|\,H_0)} \tag{2.10}$$

Assuming unit power Rayleigh namely

$$f(h_j)=2h_j e^{-h_j^2} \qquad h_j\geq 0 \tag{2.11}$$

and additive white Gaussian noise in the receiver under hypothesis H_1 we can obtain

$$
\begin{aligned}
f\left(r_j \middle| H_1\right) &= \sum_{x_j} p\left(x_j \middle| H_1\right) f\left(r_j \middle| x_j\right) \\
&= P_{D_j} f\left(r_j \middle| x_j = 1\right) + \left(1 - P_{D_j}\right) f\left(r_j \middle| x_j = -1\right) \\
&= \frac{2\sigma}{\sqrt{2\pi}\left(1 + 2\sigma^2\right)} e^{-\frac{r_j^2}{2\sigma^2}} \left\{ 1 + \left[P_{D_j} - Q\left(ar_j\right) \right] \sqrt{2\pi} a r_j e^{\frac{\left(ar_j\right)^2}{2}} \right\}
\end{aligned} \tag{2.12}
$$

$f\left(r_j \middle| H_1\right)$ can be obtained using P_{F_j} instead of P_{D_j} and LRT based fusion rule become

$$
\Lambda_{CS}\left(\mathbf{r}\right) = \prod_{j=1}^{N} \frac{1 + \left[P_{D_j} - Q\left(ar_j\right) \right] \sqrt{2\pi} a r_j e^{\frac{\left(ar_j\right)^2}{2}}}{1 + \left[P_{F_j} - Q\left(ar_j\right) \right] \sqrt{2\pi} a r_j e^{\frac{\left(ar_j\right)^2}{2}}} \tag{2.13}
$$

where $a = 1 \Big/ \left(\sigma \sqrt{1 + 2\sigma^2} \right)$.

One can see that this fusion rule which is called LRTCS in [14] only requires channel statistics and probability of detection and false alarm of the local sensors.

2.2.2 Sub-optimum Fusion Rule with Channel Statistics for Wireless Sensor Networks with Parallel Topology

Again we will examine high and low SNR behavior of the optimum rule with CS. First we analyze high-SNR behavior and we write the fusion rule in (2.13) as

$$
\Lambda_{CS}\left(\mathbf{r}\right) = \sum_{r_j < 0} \log \frac{1 + \left[P_{D_j} - Q\left(ar_j\right) \right] \sqrt{2\pi} a r_j e^{\frac{\left(ar_j\right)^2}{2}}}{1 + \left[P_{F_j} - Q\left(ar_j\right) \right] \sqrt{2\pi} a r_j e^{\frac{\left(ar_j\right)^2}{2}}} + \sum_{r_j > 0} \log \frac{1 + \left[P_{D_j} - Q\left(ar_j\right) \right] \sqrt{2\pi} a r_j e^{\frac{\left(ar_j\right)^2}{2}}}{1 + \left[P_{F_j} - Q\left(ar_j\right) \right] \sqrt{2\pi} a r_j e^{\frac{\left(ar_j\right)^2}{2}}} \tag{2.14}
$$

When $\sigma^2 \to 0$ and $a \to \infty$, for $r_j < 0$ $Q\left(ar_j\right) \to 1$ and $\left| ar_j e^{\frac{(ar_j)^2}{2}} \right| \gg 1$; for $r_j > 0$, $Q\left(ar_j\right) \to 0$ and $ar_j e^{\frac{(ar_j)^2}{2}} \gg 1$, consequently we obtain

$$\lim_{\sigma^2 \to 0} \Lambda_{cs}\left(\mathbf{r}\right) = \sum_{r_j<0} \log \frac{1-P_{D_j}}{1-P_{F_j}} + \sum_{r_j>0} \log \frac{P_{D_j}}{P_{F_j}} = \Lambda_{highSNR}\left(\mathbf{r}\right) \tag{2.15}$$

We can see that for high SNR channels the optimum rule using only CS reduces to high SNR approximation of optimum rule using CSI. This is a reasonable result because when SNR is too high, the fading coefficient becomes unimportant to use during detection. For low SNR namely $\sigma^2 \to \infty$, $a \to 0$. Using Taylor expansion

$$Q\left(x\right) = \frac{1}{2} - \frac{1}{\sqrt{2\pi}} \left(x - \frac{x^3}{2} + \frac{x^5}{2\times 4} - \frac{x^7}{2\times 4\times 8} + \ldots \right) \tag{2.16}$$

Hence using $\lim_{a\to 0} Q\left(ar_j\right) \approx 1/2 - ar_j / \sqrt{2\pi}$ and $e^{\left(ar_j\right)^2/2} \approx 1 + a^2 r_j^2 / 2$ we can conclude as in [14]

$$\lim_{\sigma^2 \to \infty} \Lambda_{cs}\left(\mathbf{r}\right) = \sum_{j=1}^{N} \left(P_{D_j} - P_{F_j}\right) \sqrt{2\pi}\, ar_j \tag{2.17}$$

If we assume local sensors are identical we obtain equal gain combiner (EGC) like fusion statistics

$$\Lambda_{EGC}\left(\mathbf{r}\right) = \sum_{j=1}^{N} r_j \tag{2.18}$$

2.3 Simulation Results

In this chapter, we will compare the performance of the fusion rules derived in the previous sections. For this experiment, there are 8 local sensors with the same detection probability $P_D = 0.5$ and false alarm probability $P_F = 0.05$. All communications are modulated by BPSK modulation, channels are subjected to unit power Rayleigh fading and AWGN therefore channel SNR is defined as $1/\sigma^2$. First, the average channel SNR is taken as 5dB and receiver operating characteristic (ROC) curves are obtained for derived fusion rules which are shown in Figure 2.2. It can be seen from figure, LRT has the best performance and LRTCS slightly worse performance than LRT. It is interesting that for this SNR, EGC outperforms both Chair-Varshney and MRC although it does not need any prior information.

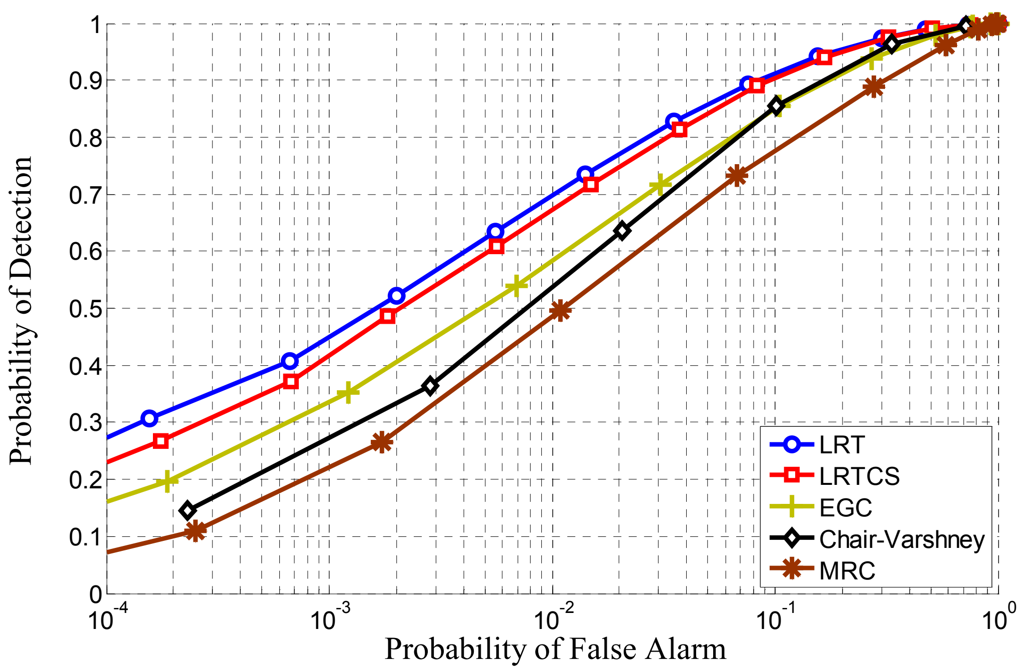

Figure 2.2 ROC curves for different fusion rules with average channel $SNR = 5dB$ where there are 8 local sensors with $P_F = 0.05$ and $P_D = 0.5$

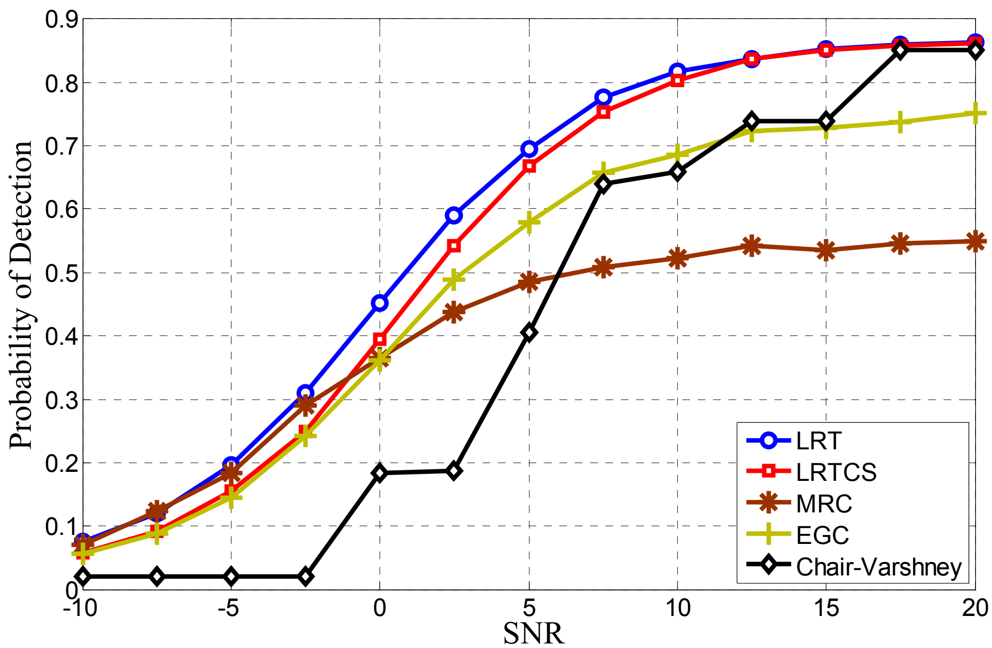

Figure 2.3 Global detection probability of different fusion rules as a function of *SNR* when $P_{F_0} = 0.01$.

As a second experiment, detection probability at the fusion center as function of *SNR* is obtained when the global false alarm probability is fixed at $P_{F_0} = 0.01$.Figure 2.3 shows that for high-SNR regimes we can see that LRT, LRTCS and Chair-Varshney become identical in detection performance which is consistent with our theoretical analysis. Also, for low-SNR, MRC and EGC behave like LRT and LRTCS respectively. LRTCS fusion rule outperforms suboptimum rules above 0dB *SNR*. Detection performance of EGC can be considered most robust among suboptimum fusion rules.

3. BINARY DISTRIBUTED DETECTION STRATEGIES in HIERARCHICAL WIRELESS SENSOR NETWORKS

In previous chapter we have analyzed decision fusion techniques under Neyman-Pearson criterion for noisy and Rayleigh faded WSN by considering parallel topology. However, the parallel topology is not convenient for sensors with relatively small transmission ranges. To increase the transmission range of the sensors, the hierarchical topology is preferable in which the local sensors send their decisions to the local fusion centers called cluster heads (CLH) and the cluster heads fuse these local sensor decisions and based on these, they make their decisions accordingly to be sent to the global fusion center. According to our knowledge, however, there is no work done in literature regarding to the distributed detection and decision fusion problem for the hierarchically configured WSN with fading and noise. To fill this gap, in this chapter, we develop the optimal and sub-optimal distributed detection and fusion rules for the hierarchical topology and investigate their performances through numerical simulations.

3.1 System Model

Two hypothesis, H_0 and H_1 must be considered while deciding whether a phenomenon present or not at region of interest. Each local sensor node makes an observation from environment, quantizes its observation to a value 0 or 1 and sends it to a cluster head through wireless channel with BPSK modulation and cluster heads make decisions to send to the global fusion center as depicted in Figure 3.1. Therefore, the signal received by the cluster head *"m"* from j^{th} sensor and by the global fusion center from m^{th} cluster head can be expressed respectively as:

$$r_j^m = h_j^m x_j^m + n_j^m$$
$$y_m = g_m s_m + n_m \tag{3.1}$$

x_j^m and s_m are the BPSK modulated signal which takes -1 and 1 values, n_j^m and n_m are the additive white Gaussian noise sample with zero mean and σ^2 variance and uncorrelated from channel to channel and h_j^m and g_m are the Rayleigh distributed fading channel gain. Since the variance is σ^2 and BPSK is used SNR of each channel is $1/\sigma^2$ assuming a unit power Rayleigh fading channel. Performance of a local sensor node can be defined in terms of false alarm and detection probabilities which are given respectively as

$$P_{F_j}^m = P\left(x_j^m = 1 \middle| H_0\right)$$
$$P_{D_j}^m = P\left(x_j^m = 1 \middle| H_1\right) \tag{3.2}$$

This notation shows that the performance indexes of j^{th} sensor connected to m^{th} cluster head.

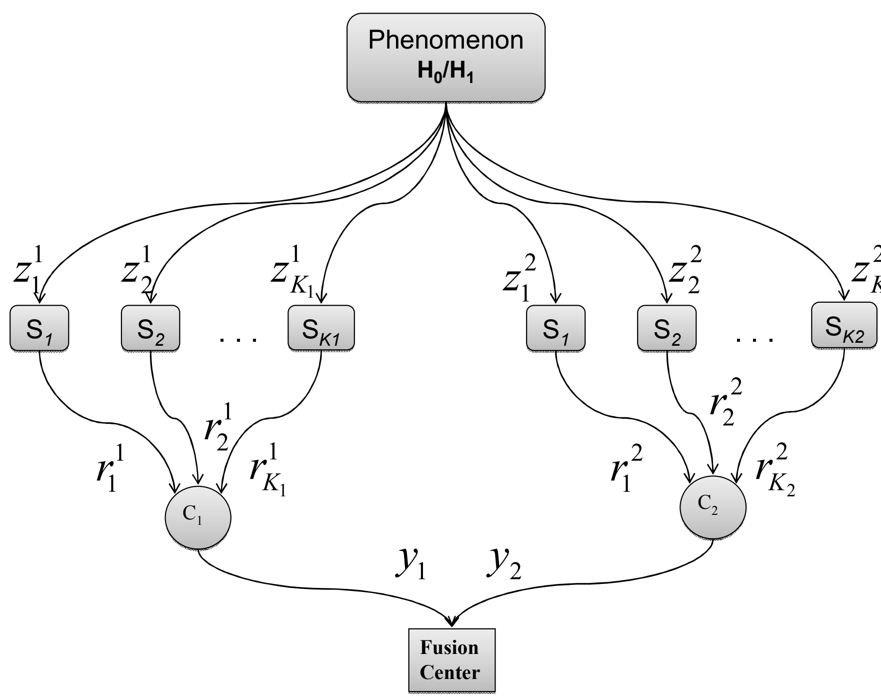

Figure 3.1 A wireless sensor network with hierarchical topology. Two cluster heads and a global fusion center

3.2 Fusion Rules Based on Channel State Information

When all performance indices namely detection and false alarm probability of local sensor nodes, and channel state information (CSI) of channels from cluster heads to global fusion center is known Likelihood Ratio Test (LRT) based optimum fusion rule is given by [9].

$$\Lambda^0(\mathbf{y}) = \frac{f(\mathbf{y}|H_1)}{f(\mathbf{y}|H_0)} = \prod_{m=1}^{N} \frac{P_{D_m}\, e^{-\frac{(y_m-g_m)^2}{2\sigma^2}} + (1-P_{D_m})\, e^{-\frac{(y_m+g_m)^2}{2\sigma^2}}}{P_{F_m}\, e^{-\frac{(y_m-g_m)^2}{2\sigma^2}} + (1-P_{F_m})\, e^{-\frac{(y_m+g_m)^2}{2\sigma^2}}} \tag{3.3}$$

y_m is the received signal from m^{th} cluster head, P_{D_m} and P_{F_m} are detection and false alarm probabilities of the cluster heads and depends on fusion rules at the clusters centers, since the optimum rule is observed it is assumed that the cluster heads have the information about performance indexes of local sensor node nodes and the CSI of the channels. Log-LRT based fusion rule at the clusters is given by

$$\Lambda^m(r) = \sum_{j=1}^{K_m} \log\left[\frac{P_{D_j}^m\, e^{-\frac{(r_j^m-h_j^m)^2}{2\sigma^2}} + (1-P_{D_j}^m)\, e^{-\frac{(r_j^m+h_j^m)^2}{2\sigma^2}}}{P_{F_j}^m\, e^{-\frac{(r_j^m-h_j^m)^2}{2\sigma^2}} + (1-P_{F_j}^m)\, e^{-\frac{(r_j^m+h_j^m)^2}{2\sigma^2}}}\right] \tag{3.4}$$

Therefore P_{D_m} and P_{F_m} in (3.3) are defined as

$$\begin{aligned} P_{F_m} &= P\left(\Lambda^m(r) > t_m \middle| H_0\right) = P\left(\sum_{j=1}^{K_m} \psi\left(r_j^m\right) > t_m \middle| H_0\right) \\ P_{D_m} &= P\left(\Lambda^m(r) > t_m \middle| H_1\right) = P\left(\sum_{j=1}^{K_m} \psi\left(r_j^m\right) > t_m \middle| H_1\right) \end{aligned} \tag{3.5}$$

where $\mathbf{r}^m = [r_1^m, r_2^m, \cdots, r_{K_m}^m]$ is the signal vector received by the m^{th} cluster head and $\psi\left(r_j^m\right)$ is the log likelihood ratio (LLR) corresponding to the signal r_j^m which can be expressed as

$$\psi\left(r_j^m\right) = \log\left(\frac{P_{D_j}^m \, \mathrm{e}^{-\frac{\left(r_j^m - h_j^m\right)^2}{2\sigma^2}} + \left(1 - P_{D_j}^m\right) \mathrm{e}^{-\frac{\left(r_j^m + h_j^m\right)^2}{2\sigma^2}}}{P_{F_j}^m \, \mathrm{e}^{-\frac{\left(r_j^m - h_j^m\right)^2}{2\sigma^2}} + \left(1 - P_{F_j}^m\right) \mathrm{e}^{-\frac{\left(r_j^m + h_j^m\right)^2}{2\sigma^2}}} \right) \tag{3.6}$$

Since we have 2^{K_m} different possible decision vectors $\mathbf{u}^m = [u_1^m, \cdots, u_{K_m}^m]$ for the m^{th} cluster head we can re-express (3.5) by using the total probability theorem as

$$\begin{aligned} P_{F_m} &= 1 - \sum_{i=0}^{2^{K_m}-1} \mathrm{P}\left(\mathbf{u} = \mathbf{u}_i^m \middle| H_0\right) F_{\Lambda^m}\left(t_m\right) \\ P_{D_m} &= 1 - \sum_{i=0}^{2^{K_m}-1} \mathrm{P}\left(\mathbf{u} = \mathbf{u}_i^m \middle| H_0\right) F_{\Lambda^m}\left(t_m\right) \end{aligned} \tag{3.7}$$

$F_{\Lambda^m}\left(t_m\right)$ is cumulative distribution function (CDF) of the fusion rule at the m^{th} cluster head given that sensor decision $\mathbf{u}$ vector equals to i^{th} realization of $\mathbf{u}^m$. Therefore, the fusion rule becomes

$$\Lambda^0\left(y\right) = \sum_{m=1}^{N} \log \left[\frac{\left(1 - \sum_{i=0}^{2^{K_m}-1} \mathrm{P}\left(\mathbf{u} = \mathbf{u}_i^m \middle| H_0\right) F_{\Lambda^m}\left(t_m\right)\right) \mathrm{e}^{-\frac{\left(y_m - g_m\right)^2}{2\sigma^2}} + \left(\sum_{i=0}^{2^{K_m}-1} \mathrm{P}\left(\mathbf{u} = \mathbf{u}_i^m \middle| H_0\right) F_{\Lambda^m}\left(t_m\right)\right) \mathrm{e}^{-\frac{\left(y_m + g_m\right)^2}{2\sigma^2}}}{\left(1 - \sum_{i=0}^{2^{K_m}-1} \mathrm{P}\left(\mathbf{u} = \mathbf{u}_i^m \middle| H_0\right)\right) F_{\Lambda^m}\left(t_m\right) \mathrm{e}^{-\frac{\left(y_m - g_m\right)^2}{2\sigma^2}} + \left(\sum_{i=0}^{2^{K_m}-1} \mathrm{P}\left(\mathbf{u} = \mathbf{u}_i^m \middle| H_0\right) F_{\Lambda^m}\left(t_m\right)\right) \mathrm{e}^{-\frac{\left(y_m + g_m\right)^2}{2\sigma^2}}} \right] \tag{3.8}$$

The terms, $P\left(\mathbf{u}=\mathbf{u}_i^m \middle| H_0\right)$ and $P\left(\mathbf{u}=\mathbf{u}_i^m \middle| H_1\right)$, in (3.8) can be determined as

$$\begin{aligned} \mathrm{P}(\mathbf{u}=\mathbf{u}_i^m \mid H_0) &= \prod_{j=1,\mathbf{u}=\mathbf{u}_i^m}^{K_m} \left(P_{F_j}^m\right)^{u_j} (1-P_{F_j}^m)^{1-u_j} \\ \mathrm{P}(\mathbf{u}=\mathbf{u}_i^m \mid H_1) &= \prod_{j=1,\mathbf{u}=\mathbf{u}_i^m}^{K_m} \left(P_{D_j}^m\right)^{u_j} (1-P_{D_j}^m)^{1-u_j} \end{aligned} \tag{3.9}$$

Where u_j is the j^{th} element of decision vector $\mathbf{u}_i^m$, this optimum rule has the best performance by means of detection but requires too much prior information about channels and sensors. Some reduction should be done to form a simpler rule practically applicable. If we define $S_0=\{m: y_m<0\}$ and $S_1=\{m: y_m>0\}$ and divide the fusion rule by using this interval the fusion rule can be written as

$$\Lambda^0(y)=\prod_{m\in S_o}\frac{P_{D_m}+\left(1-P_{D_m}\right)\mathrm{e}^{-\frac{2y_m g_m}{\sigma^2}}}{P_{F_m}+\left(1-P_{F_m}\right)\mathrm{e}^{-\frac{2y_m g_m}{\sigma^2}}}\times\prod_{m\in S_1}\frac{P_{D_m}\,\mathrm{e}^{\frac{2y_m g_m}{\sigma^2}}+\left(1-P_{D_m}\right)}{P_{F_m}\,\mathrm{e}^{\frac{2y_m g_m}{\sigma^2}}+\left(1-P_{F_m}\right)} \tag{3.10}$$

In the high SNR regime ($\sigma^2\to 0$) terms with $\mathrm{e}^{\pm\frac{2y_m g_m}{\sigma^2}}$ become greater than other terms so fusion rule become as in [9].

$$\Lambda_{HighSNR}^0=\lim_{\sigma^2\to 0}\log\left(\Lambda^0(y)\right)=\sum_{m\in S_0}\log\left(\frac{1-P_{D_m}}{1-P_{F_m}}\right)+\sum_{m\in S_1}\log\left(\frac{P_{D_m}}{P_{F_m}}\right) \tag{3.11}$$

To apply the high SNR approximation global fusion rule, the knowledge on P_{D_m} and P_{F_m} is needed. For simplicity, we assume also that SNR is high for the cluster head fusion. This means that we can also apply the high SNR approximation of (3.4) and obtain a rule similar to (3.11) which is known Chair-Varshney rule [4] in literature. Assuming all sensors within each clusters are identical, i.e., $P_{F_j}^m=P_F^m$ and $P_{D_j}^m=P_D^m$, P_{D_m} and P_{F_m} are derived by using Chair-Varshney rule in [14] as

$$
\begin{aligned}
P_{F_m} &= \sum_{i=K_\tau^m}^{K_m} \binom{K_m}{i} p_0{}^i \left(1-p_0\right)^{K_m-i} \\
P_{D_m} &= \sum_{i=K_\tau^m}^{K_m} \binom{K_m}{i} p_1{}^i \left(1-p_1\right)^{K_m-i}
\end{aligned}
\tag{3.12}
$$

where K_τ^m is the decision threshold of the m^{th} cluster head in terms of number of sensors and p_0^m and p_1^m are the probabilities of a nonnegative cluster observation r_j^m under H_0 and H_1 respectively which are given as [14]

$$
\begin{aligned}
p_0^m &= \frac{1}{2} + \frac{P_{F_m} - \frac{1}{2}}{\sqrt{1+2\sigma^2}} \\
p_1^m &= \frac{1}{2} + \frac{P_{D_m} - \frac{1}{2}}{\sqrt{1+2\sigma^2}}
\end{aligned}
\tag{3.13}
$$

Similarly, to obtain a simplified global fusion rule, we can also consider low SNR regime (i.e. $\sigma^2 \to \infty$). To do that, we can re-express the global fusion rule in (3.3) as

$$
\Lambda^0\left(\mathbf{y}\right) = \prod_{m=1}^{N} \frac{P_{D_m} + \left(1-P_{D_m}\right) \mathrm{e}^{-\frac{2y_m g_m}{\sigma^2}}}{P_{F_m} + \left(1-P_{F_m}\right) \mathrm{e}^{-\frac{2y_m g_m}{\sigma^2}}} \tag{3.14}
$$

To obtain a low SNR approximation of the global fusion rule, the first order Taylor series approximation of $\mathrm{e}^{-\frac{2y_m g_m}{\sigma^2}}$ is employed as $\mathrm{e}^{-\frac{2y_m g_m}{\sigma^2}} \approx \left(1 - \frac{2y_m g_m}{\sigma^2}\right)$. Therefore, the low-SNR global fusion can be obtained similar to [9]as

$$
\Lambda_{lowSNR}^0(\mathbf{y}) = \sum_{m=1}^{N} \left(P_{D_m} - P_{F_m}\right) g_m y_m \tag{3.15}
$$

If cluster heads are identical in terms of detection and false alarm probabilities, the global fusion rule can be simplified further as

$$\Lambda_{lowSNR}^{0}(\mathbf{y}) = \sum_{m=1}^{N} g_m y_m \tag{3.16}$$

which is equivalent up to a scale factor to the Maximum Ratio Combining (MRC) fusion rule. Therefore, we do not need to know the detection and false alarm probabilities of the cluster heads for the low-SNR fusion rule by assuming they are identical.

Parallel to this, fusion rule for m^{th} cluster with identical local sensor node nodes under low-SNR can be determined in as

$$\Lambda_{lowSNR}^{m}\left(\mathbf{r}\right) = \sum_{j=1}^{K_m} h_j^m r_j^m \tag{3.17}$$

Hence, cluster heads do not also need to know performance indices of local sensor nodes assuming they are identical, although the performance of the low-SNR fusion rule can be analyzed analytically by deriving the probability distribution function of (3.17), if we define

$$\Omega = h_j^m r_j^m \tag{3.18}$$

Since we assume h_j^m is known to cluster heads we can write pdf of r_j^m as

$$f\left(r \middle| h_j^m, H_1\right) = P_{D_j}^m N\left(r, h_j^m, \sigma^2\right) + \left(1 - P_{D_j}^m\right) N\left(r, -h_j^m, \sigma^2\right). \tag{3.19}$$

Where $N\left(x, \mu, \sigma^2\right)$ denotes normal distribution with mean μ, and σ^2 variance for the random variable X. Therefore when h_j^m is known r_j^m is a normal distributed random variable with mean $2P_{D_j}^m h_j^m - h_j^m$ and σ^2 variance. Then distribution of (3.18) becomes

$$f_{\Omega}\left(\omega\middle|h_j^m, H_1\right) = N\left(\omega, 2P_{D_j}^m h_j^{m^2} - h_j^{m^2}, h_j^{m^2}\sigma^2\right) \tag{3.20}$$

Probability distribution of (3.17) can be derived convolving the (3.20) K_m times for channels in the cluster.

3.3 Fusion Rules with Known Channel Fading Statistics

Thus far, except high-SNR approximation the other two fusion rule uses channel state information (CSI) to make a global decision, since obtaining CSI is a power consuming issue and reducing power consumption is a crucial issue for WSN's we will try to find a fusion rule that uses only channel statistic (CS) namely probability density function of channel gain. Channel fading statistics based fusion rule is given in [14] and if we rewrite it for hierarchical scenario we have

$$\Lambda_{CS}^0(\mathbf{y}) = \sum_{m=1}^{N} \log\left[\frac{1+\left[P_{D_m} - Q(ay_m)\right]\sqrt{2\pi}ay_m e^{\frac{(ay_m)^2}{2}}}{1+\left[P_{F_m} - Q(ay_m)\right]\sqrt{2\pi}ay_m e^{\frac{(ay_m)^2}{2}}}\right] \tag{3.21}$$

in general $a = \left(\dfrac{\sigma_g}{\sigma\sqrt{\sigma_g^2 + \sigma^2}}\right)$ and $2\sigma_g^2$ is the mean square value of the g_m CSI for m^{th} channel. The channel statistics based fusion rule at the cluster head is

$$\Lambda_{CS}^m(\mathbf{r}) = \sum_{j=1}^{K_m} \log\left[\frac{1+\left[P_{D_j}^m - Q(ar_j^m)\right]\sqrt{2\pi}ar_j^m e^{\frac{\left(ar_j^m\right)^2}{2}}}{1+\left[P_{F_j}^m - Q(ar_j^m)\right]\sqrt{2\pi}ar_j^m e^{\frac{\left(ar_j^m\right)^2}{2}}}\right] \tag{3.22}$$

Therefore false alarm and detection probability at the cluster heads are given respectively as

$$
\begin{aligned}
P_{F_m} &= P\left(\Lambda_{CS}^{m}(\mathbf{r}) > t_m \middle| H_0\right) = P\left(\sum_{j=1}^{K_m} \psi_{CS}\left(r_j^m\right) > t_m \middle| H_0\right) \\
P_{D_m} &= P\left(\Lambda_{CS}^{m}(\mathbf{r}) > t_m \middle| H_1\right) = P\left(\sum_{j=1}^{K_m} \psi_{CS}\left(r_j^m\right) > t_m \middle| H_1\right)
\end{aligned}
\tag{3.23}
$$

To obtain these probabilities we have to find the probability distribution of $\psi_{CS}\left(r_j^m\right) = \log\left[\frac{1+\left[P_{D_j}^m - Q(ar_j^m)\right]\sqrt{2\pi}ar_j^m e^{\frac{\left(ar_j^m\right)^2}{2}}}{1+\left[P_{F_j}^m - Q(ar_j^m)\right]\sqrt{2\pi}ar_j^m e^{\frac{\left(ar_j^m\right)^2}{2}}}\right]$ and to achieve this distribution, we first look at the following approximation.

$$
Q(x) \cong \frac{1}{\sqrt{2\pi}x} e^{-x^2/2} \quad x > 3 \tag{3.24}
$$

Therefore $e^{-x^2/2} = x\sqrt{2\pi}Q(x) + C(x) \quad x > 3$ where $C(x)$ is the correction term. Hence

$$
C(x) = e^{-x^2/2} - x\sqrt{2\pi}Q(x) \tag{3.25}
$$

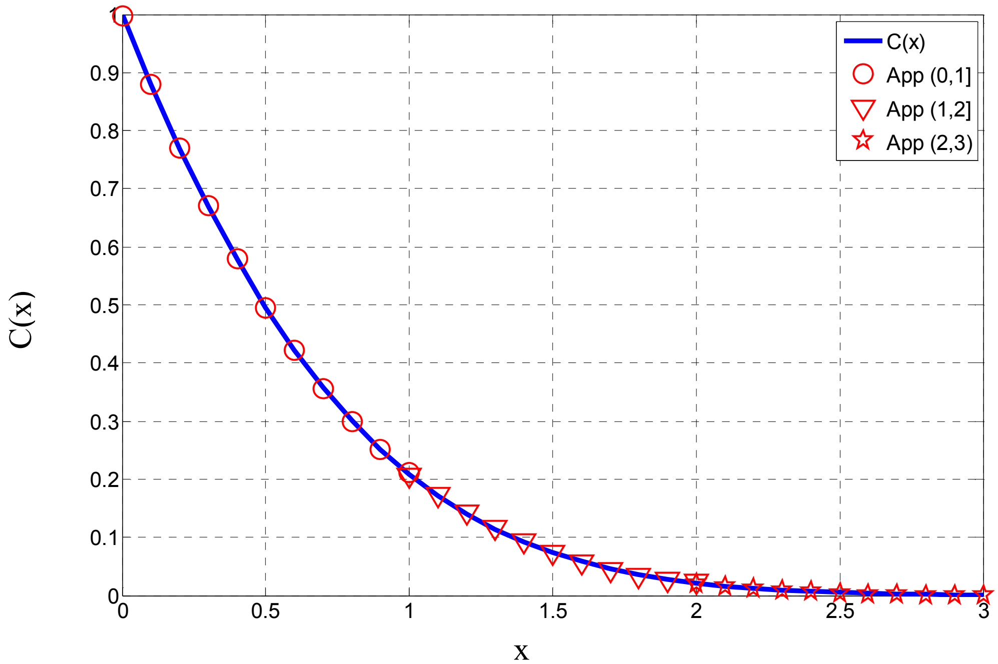

Figure 3.2 Correction function $C(x)$ and its approximation

Figure 3.2 shows $C(x)$ which is the correction term and can be approximated partially by a quadratic polynomial, and App's in the figure shows the partial approximation of the correction term , so $\psi_{cs}(r)$ is written for where $x = ar$

$$\psi_{cs}(r_j^m) = \log\left[\frac{1+\dfrac{\left[P_{D_j}^m - Q(x)\right]\sqrt{2\pi x}}{\sqrt{2\pi}xQ(x)+C(x)}}{1+\dfrac{\left[P_{F_j}^m - Q(x)\right]\sqrt{2\pi x}}{\sqrt{2\pi}xQ(x)+C(x)}}\right]$$
$$\psi_{cs}(r_j^m) = \log\left[\frac{x\sqrt{2\pi}P_{D_j}^m + C(x)}{x\sqrt{2\pi}P_{F_j}^m + C(x)}\right] \tag{3.26}$$

Since correction function is approximated by quadratic polynomial $C(x) = \alpha x^2 + \beta x + \chi$, (3.26) turns into

$$e^{\psi_{CS}} = \frac{x\sqrt{2\pi}P_{D_j}^m + \alpha x^2 + \beta x + \chi}{x\sqrt{2\pi}P_{F_j}^m + \alpha x^2 + \beta x + \chi} \tag{3.27}$$

and where the optimal values of coefficients of α, β, χ are obtained for intervals (0,1], (1,2] and (2,3] and tabulated in Table 3.1.

Table 3.1 Coefficients of quadratic approximations of $C(x)$ for different intervals

	$(0,1]$	$(1,2]$	$(2,3]$
α	0.4353	0.1664	0.02439
β	-1.221	-0.681	-0.1409
χ	0.9977	0.7205	0.2047

After collecting the terms in (3.27) together and changing x into ar we get a quadratic equation and the roots are

$$r_{1,2} = \frac{-\left(\sqrt{2\pi}P_{D_j}^m + \beta - e^{\psi_{CS}}\left(\sqrt{2\pi}P_{F_j}^m + \beta\right)\right) \pm \sqrt{\left(\sqrt{2\pi}P_{D_j}^m + \beta - e^{\psi_{CS}}\left(\sqrt{2\pi}P_{F_j}^m + \beta\right)\right)^2 - 4\alpha\chi\left(e^{\psi_{CS}} - 1\right)^2}}{2\alpha\left(e^{\psi_{CS}} - 1\right)a} \tag{3.28}$$

The distribution of $\psi_{cs}(r)$ can be found using the formula in [15].

$$f_{\psi_{CS}}\left(\psi_{cs}\right) = \sum_k \left.\frac{f_r(r)}{\left|d\psi_{cs}/dr\right|}\right|_{r=r_k} \tag{3.29}$$

An approximation of Q function will be used to find solutions of $\psi_{CS} = g(r)$. Coefficients are derived by collecting the terms of equation (3.27). The probability distribution of $\psi_{CS}(r)$ is found from (3.29),and $f_{\Lambda_{CS}}(r)$ can be found by convolving this pdf K_m times since they are i.i.d when σ_h is same for every channel. To obtain $f_\psi(r)$ we have to know $f_r(r)$; but because of Neyman-Pearson assumption we do not know the prior probabilities of H_0 and H_1 therefore conditional pdfs $f_\psi(r|H_0)$ and $f_\psi(r|H_1)$ can be calculated from $f_r(r|H_0)$ and $f_r(r|H_1)$ they are given in [14]

$$
\begin{aligned}
f_r(r|H_0) &= \frac{\sigma}{\sqrt{2\pi}\left(\sigma_c^2+\sigma\right)} e^{-\left(\frac{r^2}{2\sigma^2}\right)} \times \left[1+\left[P_{F_j}^m - Q(ar_j^m)\right]\sqrt{2\pi}ar_j^m e^{\frac{\left(ar_j^m\right)^2}{2}}\right] \\
f_r(r|H_1) &= \frac{\sigma}{\sqrt{2\pi}\left(\sigma_c^2+\sigma\right)} e^{-\left(\frac{r^2}{2\sigma^2}\right)} \times \left[1+\left[P_{D_j}^m - Q(ar_j^m)\right]\sqrt{2\pi}ar_j^m e^{\frac{\left(ar_j^m\right)^2}{2}}\right]
\end{aligned} \tag{3.30}
$$

The accuracy of this method is tested by a simulation experiment that the results can be seen in figure.

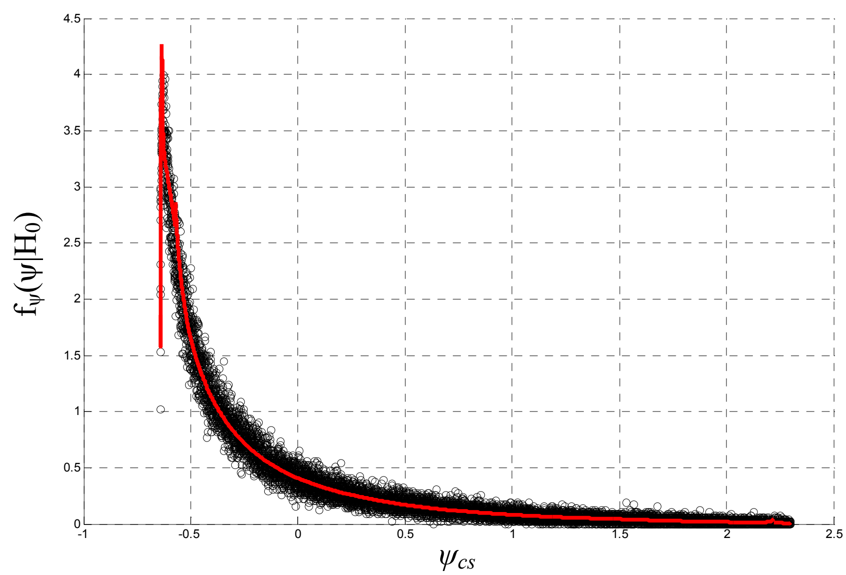

Figure 3.3 Probability Density Function of $\psi_{cs}(r)$ for P_{Fj}^m=0.05 and P_{Dj}^m=0.5 under H_0

In this experiment there is one local sensor node which sends its decision to the fusion center and fusion center calculates the $\psi_{CS}\left(r\right)$ value and the pdf is obtained from this simulation using "hist" comment MATLAB. False alarm (P_{Fj}^{m}) and detection probabilities (P_{Dj}^{m}) are chosen 0.05 and 0.5 respectively and the SNR ($1/\sigma^2$) is 5dB by assuming unit power Rayleigh fading. The figure shows that our approximation is valid.

In the high SNR regime $\lim_{\sigma^2 \to 0} a \to \infty$ and fusion rule with channel statistics turn into Chair-Varshney fusion rule and it can be seen easily by rewriting the fusion rule as

$$\begin{aligned}\Lambda_{CS}^{0}(y) = &\sum_{m\in S_0} \log\left[\frac{1+\left[P_{D_m} - Q(ay_m)\right]\sqrt{2\pi}ay_m e^{\frac{(ay_m)^2}{2}}}{1+\left[P_{F_m} - Q(ay_m)\right]\sqrt{2\pi}ay_m e^{\frac{(ay_m)^2}{2}}}\right] \\ &+\sum_{m\in S_1} \log\left[\frac{1+\left[P_{D_m} - Q(ay_m)\right]\sqrt{2\pi}ay_m e^{\frac{(ay_m)^2}{2}}}{1+\left[P_{F_m} - Q(ay_m)\right]\sqrt{2\pi}ay_m e^{\frac{(ay_m)^2}{2}}}\right]\end{aligned} \tag{3.31}$$

If $m \in S_0$ when $a \to \infty$ $Q(ay_m) \to 1$ and $\left|ay_m e^{\frac{(ay_m)^2}{2}}\right| \gg 1$ and if $m \in S_1$ $Q(ay_m) \to 0$ $ay_m e^{\frac{(ay_m)^2}{2}} \gg 1$ therefore we can conclude

$$\sum_{m\in S_0} \log\left(\frac{1-P_{D_m}}{1-P_{F_m}}\right) + \sum_{m\in S_1} \log\left(\frac{P_{D_m}}{P_{F_m}}\right) = \Lambda_{HighSNR}^{0} \tag{3.32}$$

One can realize that high-SNR approximation for LRTCS (3.32) is same with high-SNR approximation for LRT (3.11) which means when SNR is decreased difference between detection performances of fusion rules decrease.

In the low high SNR regime when $\lim_{\sigma^2 \to \infty} a \to 0$ setting $Q(ay_m) \approx 1/2 - ay_m / \sqrt{2\pi}$ and $e^{\frac{(ay_m)^2}{2}} \approx 1 + (ay_m)^2 / 2$ we can achieve EGC fusion rule as in [14]

$$\Lambda_3^0(y) = \frac{1}{N} \sum_{m=1}^{N} y_m \tag{3.33}$$

3.4 Performance Evaluation

In this section, we compare the performances of the proposed optimal and sub-optimal global fusion rules through numerical simulations. Throughout our discussion, we consider a hierarchical WSN with 2 clusters and each cluster head communicates with 4 local sensor node nodes, i.e., N=2, K_1=K_2=4, and we assume that sensors are identical in terms of their false alarm and detection probabilities as $P_{F_j}^m = 0.05$ and $P_{D_j}^m = 0.5$. We first obtain the receiver operating characteristics (ROC) curves of the global fusion rules as in Figure 3.4 for *SNR* of 5 dB. It can be seen from this figure that the optimal LRT based global fusion rule considerably outperforms LRTCS, the high-SNR and low-SNR approximations of fusion rules for various global false alarm probabilities. As one can also realize from Figure 3.4, that the performance curve of the high-SNR fusion rule is not drawn for all global false alarm rates since an integer threshold K_τ^m which satisfies (3.12) cannot be found for the false alarm probabilities less than $3\text{x}10^{-2}$. Also, a notable point is low-SNR approximation for LRTCS has better performance than low-SNR approximation for LRT for 5dB *SNR* although it needs less prior information.

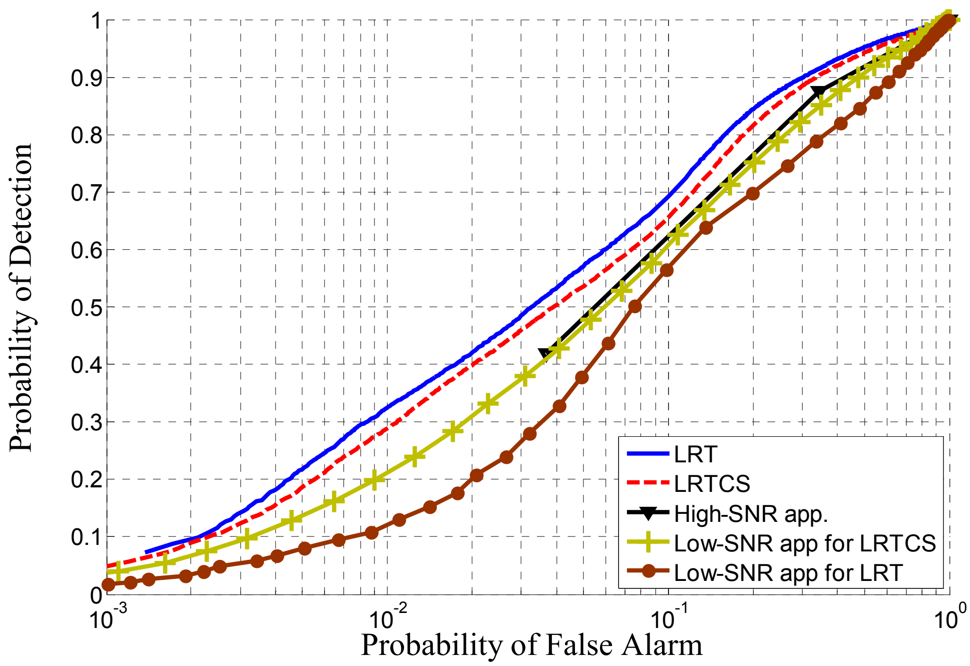

Figure 3.4 ROC curves for different fusion rules for $SNR = 5dB$ .There are 2 clusters and 4 local sensor node in each cluster with $P_{D_j}^{m} = 0.5$, $P_{F_j}^{m} = 0.05$

Since low-SNR approximation resembles MRC one can think that it should have better performance than EGC like high SNR approximation but MRC is preferable when there is identical input into the multiple fading channels. When inputs of channels are not identical it is not guaranteed to have better performance.

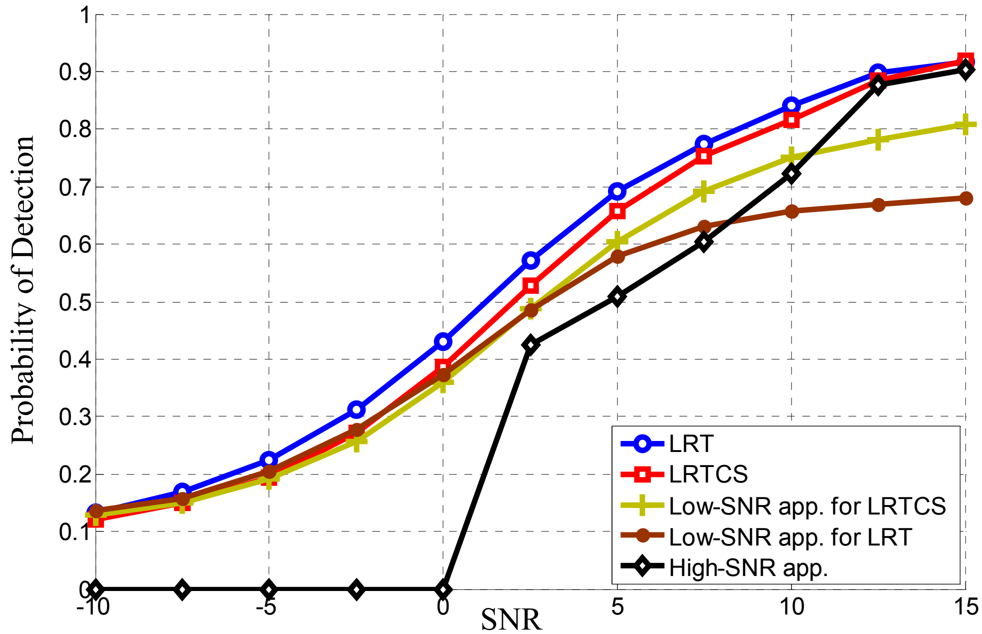

Figure 3.5 Global detection probability of different fusion rules as a function of *SNR* when $P_F^0 = 0.1$.

Secondly, we set the global false probability P_F^0 as 0.1 and obtain the global detection probabilities of the proposed fusion rules as a function of different *SNR* values as seen in Figure 3.5. Again, the LRT shows the best performance whereas the high-SNR and low-SNR approximations to LRT and LRTCS give satisfactory results at low-noise and high-noise regimes respectively as expected. Again, integer threshold K_τ^m could not be found for *SNR's* smaller than 2.5 dB high-SNR approximation.

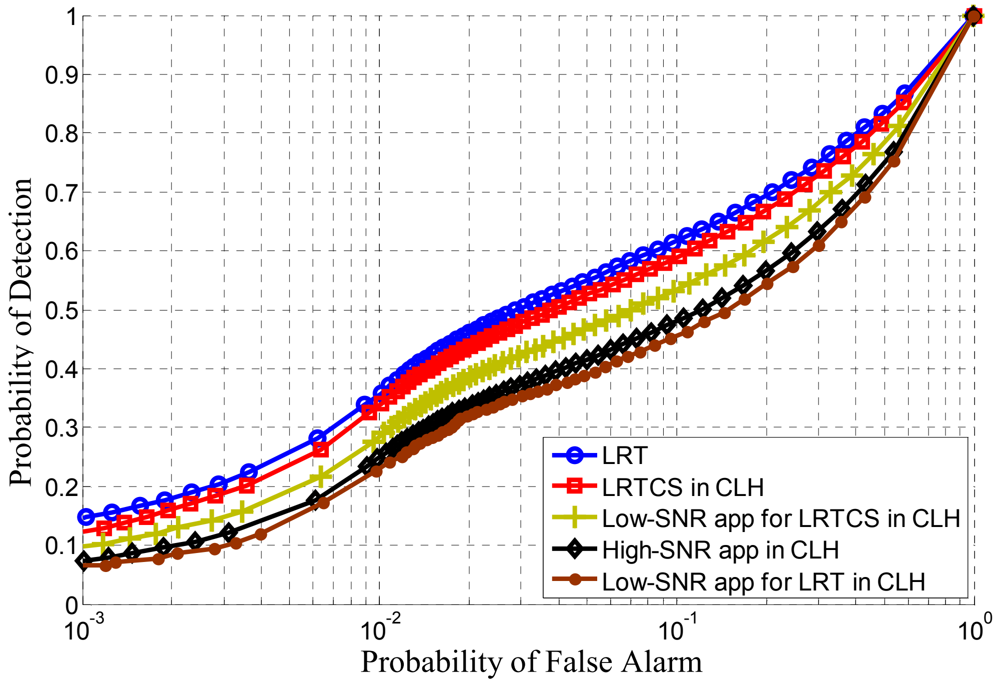

Figure 3.6 ROC curves when fusion center uses LRT and cluster heads use various fusion rules for $SNR = 5dB$ .There are 2 clusters and 4 local sensor node node in each cluster with $P_{D_j}^{m} = 0.5$, $P_{F_j}^{m} = 0.05$

For WSNs it is assumed that fusion center has no power limitation therefore optimal and most power consuming rule LRT can be used in fusion center. Figure 3.6 shows the detection performance when fusion center uses LRT and cluster heads (CLH) uses various fusion rules. It can be seen from figure when LRTCS is used in CLHs global detection performance is very close to optimal one. Also, for this 5 dB avarage SNR scenario Low-SNR approximation for LRTCS has better performance than high SNR approximation.

4. DISTRIBUTED DETECTION with MULTI-BIT DECISION in PARALLEL and HIERARCHICAL WIRELESS SENSOR NETWORKS

So far, we analyzed fusion strategies when local sensor node and cluster heads quantized their decision with one threshold namely their decisions are single-bit. In this chapter we investigate fusion strategies using multi-bit decisions in both local sensor nodes and cluster heads in WNS's with hierarchical topology. Making multi-bit decisions improves the detection performance because local sensor node and cluster heads sends much more information to global fusion center compared to single-bit decision. We first give some information about two-bit decision as an example of multi-bit decision and derive optimum and suboptimum fusion rules for multi-bit decision.

4.1 Two-Bit Decision

In conventional one bit decision, decision makers (local sensor node and/or cluster head) quantize its information about phenomenon after comparing a threshold. To make multi-bit decision one needs $2^M - 1$ threshold where M is the number of bits used for sending decision. Therefore, for a 2 bit decision one needs three thresholds, and j^{th} local sensor Figure 2.1make the following decision

$$\begin{aligned}
&\Lambda\left(z_j\right) > t_{1j} \rightarrow \mathbf{x}_j = [1,1]\left(H_1\right) \\
&t_{1j} > \Lambda\left(z_j\right) > t_{2j} \rightarrow \mathbf{x}_j = [1,-1] \\
&t_{2j} > \Lambda\left(z_j\right) > t_{3j} \rightarrow \mathbf{x}_j = [-1,1] \\
&t_{3j} > \Lambda\left(z_j\right) \rightarrow \mathbf{x}_j = [-1,-1]\left(H_0\right)
\end{aligned} \tag{4.1}$$

Where z_j is the observation of local sensor node, t_{ij}'s are the thresholds and $\mathbf{x}_j$ is the vector of BPSK modulated decisions and in general it has M elements. If observation is greater than t_{1j} or smaller than t_{3j} local sensor node decides H_1 or H_0 respectively but if it is between these values local sensor node does not give a strict decision and sends information about which phenomenon is more likely by comparing t_{2j}. Performance indices of local sensor node can be expressed as

$$\begin{aligned}
P_{F_j} &= P\left\{\Lambda\left(z_j\right) > t_{1j} \left| H_0\right.\right\} \\
P_{10H_{0j}} &= P\left\{t_{1j} > \Lambda\left(z_j\right) > t_{2j} \left| H_0\right.\right\} \\
P_{01H_{0j}} &= P\left\{t_{2j} > \Lambda\left(z_j\right) > t_{3j} \left| H_0\right.\right\} \\
P_{T_j} &= P\left\{\Lambda\left(z_j\right) < t_{3j} \left| H_0\right.\right\} \\
P_{D_j} &= P\left\{\Lambda\left(z_j\right) > t_{1j} \left| H_1\right.\right\} \\
P_{10H_{1j}} &= P\left\{t_{1j} > \Lambda\left(z_j\right) > t_{2j} \left| H_1\right.\right\} \\
P_{01H_{1j}} &= P\left\{t_{2j} > \Lambda\left(z_j\right) > t_{3j} \left| H_1\right.\right\} \\
P_{M_j} &= P\left\{\Lambda\left(z_j\right) < t_{3j} \left| H_1\right.\right\}
\end{aligned} \tag{4.2}$$

Where P_{F_j}, P_{D_j}, P_{M_j} are the false alarm ,detection and miss probability respectively, P_{T_j} is the probability of deciding H_0 when null hypothesis is true, $P_{10H_{ij}}$ and $P_{01H_{ij}}$ where $i \in \{0,1\}$ are about which phenomenon is more likely conditioned on hypotheses.

4.2 Optimum Fusion Rule for Multi-bit Decision

In the next sections we will derive optimum fusion rules for WSNs with parallel and hierarchical topology.

4.2.1 Optimum Fusion Rule for Multi-bit Decision with Parallel Topology

In hierarchical topology, we can model local sensor nodes-cluster head and cluster heads-fusion center connections as two cascade parallel network topology. Therefore in the first step we will derive fusion rule for parallel topology and then we will achieve fusion rules for hierarchical topology. For parallel topology in Figure 2.1 using Neyman-Pearson lemma if we write the likelihood ratio based optimum fusion rule will be like in (3.3) we get

$$\Lambda(\mathbf{r}) = \frac{f(\mathbf{r}|H_1)}{f(\mathbf{r}|H_0)} = \prod_{m=1}^{N} \frac{f\left(r_j^0, r_j^1 \middle| H_1, h_j^0, h_j^1\right)}{f\left(r_j^0, r_j^1 \middle| H_0, h_j^0, h_j^1\right)} \tag{4.3}$$

Using conditionally independent assumption we can write

$$f\left(r_j^0, r_j^1 \middle| H_1, h_j^0, h_j^1\right) = \sum_{x_j^0}\sum_{x_j^1} f\left(r_j^0 \middle| H_1, h_j^0\right) f\left(r_j^1 \middle| H_1, h_m^1\right) p\left(x_j^0, x_j^1 \middle| H_1\right) \tag{4.4}$$

Assuming channel gains and performance indices in (4.2) are known to global fusion center previous equation becomes

$$\begin{aligned} f\left(r_j^0, r_j^1 \middle| H_1, h_j^0, h_j^1\right) &= P_{D_j} N\left(r_j^0, h_j^0, \sigma_j^{\ 2}\right) N\left(r_j^1, h_j^1, \sigma_j^{\ 2}\right) + P_{10H_{1j}} N\left(r_j^0, h_j^0, \sigma_j^{\ 2}\right) N\left(r_j^1, -h_j^1, \sigma_j^{\ 2}\right) \\ & P_{01H_{1j}} N\left(r_j^0, -h_j^0, \sigma_j^{\ 2}\right) N\left(r_j^1, h_j^1, \sigma_j^{\ 2}\right) + P_{M_j} N\left(r_j^0, -h_j^0, \sigma_j^{\ 2}\right) N\left(r_j^1, -h_j^1, \sigma_j^{\ 2}\right) \end{aligned} \tag{4.5}$$

Where $N\left(x, \mu, \sigma^2\right)$ denotes normal distribution with mean μ, and σ^2 variance for the random variable X. Therefore optimum global fusion rule is

$$\Lambda(\mathbf{y}) = \sum_{j=1}^{N} \log \left[\frac{ P_{D_j} e^{\frac{-\left(r_j^0 - h_j^0\right)^2 - \left(r_j^1 - h_j^1\right)^2}{2\sigma^2}} + P_{10H_{1j}} e^{\frac{-\left(r_j^0 - h_j^0\right)^2 - \left(r_j^1 + h_j^1\right)^2}{2\sigma^2}} + P_{01H_{1j}} e^{\frac{-\left(r_j^0 + h_j^0\right)^2 - \left(r_j^1 - h_j^1\right)^2}{2\sigma^2}} + P_{M_j} e^{\frac{-\left(r_j^0 + h_j^0\right)^2 - \left(r_j^1 + h_j^1\right)^2}{2\sigma^2}} }{ P_{F_j} e^{\frac{-\left(r_j^0 - h_j^0\right)^2 - \left(r_j^1 - h_j^1\right)^2}{2\sigma^2}} + P_{10H_{0j}} e^{\frac{-\left(r_j^0 - h_j^0\right)^2 - \left(r_j^1 + h_j^1\right)^2}{2\sigma^2}} + P_{01H_{0j}} e^{\frac{-\left(r_j^0 + h_j^0\right)^2 - \left(r_j^1 - h_j^1\right)^2}{2\sigma^2}} + P_{T_j} e^{\frac{-\left(r_j^0 + h_j^0\right)^2 - \left(r_j^1 + h_j^1\right)^2}{2\sigma^2}} } \right] \tag{4.6}$$

4.2.2 Optimum Fusion Rule for Multi-bit Decision with Hierarchical Topology

Since we explore the optimum rule for global fusion in hierarchical topology we have to assume that cluster heads have the information of performance indices and channel state information then fusion rule for m^{th} cluster head will be in the same form with (4.6)

$$\Lambda^m(\mathbf{r}) = \sum_{j=1}^{K_m} \log \left[\frac{ P_{D_j}^m e^{\frac{-\left(r_{1j}^m - h_{1j}^m\right)^2 - \left(r_{2j}^m - h_{2j}^m\right)^2}{2\sigma^2}} + P_{10H_{1j}}^m e^{\frac{-\left(r_{1j}^m - h_{1j}^m\right)^2 - \left(r_{2j}^m + h_{2j}^m\right)^2}{2\sigma^2}} + P_{01H_{1j}}^m e^{\frac{-\left(r_{1j}^m + h_{1j}^m\right)^2 - \left(r_{2j}^m - h_{2j}^m\right)^2}{2\sigma^2}} + P_{M_j}^m e^{\frac{-\left(r_{1j}^m + h_{1j}^m\right)^2 - \left(r_{2j}^m + h_{2j}^m\right)^2}{2\sigma^2}} }{ P_{F_j}^m e^{\frac{-\left(r_{1j}^m - h_{1j}^m\right)^2 - \left(r_{2j}^m - h_{2j}^m\right)^2}{2\sigma^2}} + P_{10H_{0j}}^m e^{\frac{-\left(r_{1j}^m - h_{1j}^m\right)^2 - \left(r_{2j}^m + h_{2j}^m\right)^2}{2\sigma^2}} + P_{01H_{0j}}^m e^{\frac{-\left(r_{1j}^m + h_{1j}^m\right)^2 - \left(r_{2j}^m - h_{2j}^m\right)^2}{2\sigma^2}} + P_{T_j}^m e^{\frac{-\left(r_{1j}^m + h_{1j}^m\right)^2 - \left(r_{2j}^m + h_{2j}^m\right)^2}{2\sigma^2}} } \right] \tag{4.7}$$

And optimum rule at the global fusion center will be

$$\Lambda^0(\mathbf{y}) = \sum_{m=1}^{N} \log \left[\frac{\begin{array}{l} P_{D_m} e^{\frac{-\left(y_m^0 - g_m^0\right)^2 - \left(y_m^1 - g_m^1\right)^2}{2\sigma^2}} + P_{10H_{1m}} e^{\frac{-\left(y_m^0 - g_m^0\right)^2 - \left(y_m^1 + g_m^1\right)^2}{2\sigma^2}} \\ + P_{01H_{1m}} e^{\frac{-\left(y_m^0 + g_m^0\right)^2 - \left(y_m^1 - g_m^1\right)^2}{2\sigma^2}} + P_{M_m} e^{\frac{-\left(y_m^0 + g_m^0\right)^2 - \left(y_m^1 + g_m^1\right)^2}{2\sigma^2}} \end{array}}{\begin{array}{l} P_{F_m} e^{\frac{-\left(y_m^0 + g_m^0\right)^2 - \left(y_m^1 + g_m^1\right)^2}{2\sigma^2}} + P_{10H_{0m}} e^{\frac{-\left(y_m^0 - g_m^0\right)^2 - \left(y_m^1 + g_m^1\right)^2}{2\sigma^2}} \\ + P_{01H_{0m}} e^{\frac{-\left(y_m^0 + g_m^0\right)^2 - \left(y_m^1 + g_m^1\right)^2}{2\sigma^2}} + P_{T_m} e^{\frac{-\left(y_m^0 + g_m^0\right)^2 - \left(y_m^1 + g_m^1\right)^2}{2\sigma^2}} \end{array}} \right] \tag{4.8}$$

Optimum distributed detection in hierarchical topology can be done using (4.8) in global fusion center and (4.7) in cluster heads.

4.3 Sub-Optimum Fusion Rule for Multi-bit Decision

Optimum fusion rules for cluster heads and global fusion center has the best detection performance, on the other hand it needs all sensors performance indices and channel gains. Therefore, we analyzed the asymptotic behavior of optimum fusion to achieve simpler and applicable fusion rules.

4.3.1 Sub-Optimum Fusion Rule for Multi-bit Decision in Parallel Topology

Global fusion rule in (4.6) is optimum in detection performance but it is very complex and needs very much prior information , we now try to simplify it by analyzing high and low SNR behaviors.

For this purpose, we rewrite (4.6) as

$$\Lambda(\mathbf{r}) = \prod_{r_j^0, r_j^1 < 0} \frac{P_{D_j} + P_{10H_{1j}} e^{\frac{2r_j^1 h_j^1}{\sigma^2}} + P_{01H_{1j}} e^{\frac{-2r_j^0 h_j^0}{\sigma^2}} + P_{M_j} e^{\frac{-2r_j^0 h_j^0 - 2r_j^1 h_j^1}{\sigma^2}}}{P_{F_j} + P_{10H_{0j}} e^{\frac{2r_j^1 h_j^1}{\sigma^2}} + P_{01H_{0j}} e^{\frac{-2r_j^0 h_j^0}{\sigma^2}} + P_{T_j} e^{\frac{-2r_j^0 h_j^0 - 2r_j^1 h_j^1}{\sigma^2}}}$$

$$\times \prod_{r_j^0, r_j^1 > 0} \frac{P_{D_j} e^{\frac{2r_j^0 h_j^0 + 2r_j^1 h_j^1}{\sigma^2}} + P_{10H_{1j}} e^{\frac{2r_j^0 h_j^0}{\sigma^2}} + P_{01H_{1j}} e^{\frac{2r_j^1 h_j^1}{\sigma^2}} + P_{M_j}}{P_{F_j} e^{\frac{2r_j^0 h_j^0 + 2r_j^1 h_j^1}{\sigma^2}} + P_{10H_{0j}} e^{\frac{2r_j^0 h_j^0}{\sigma^2}} + P_{01H_{0j}} e^{\frac{2r_j^1 h_j^1}{\sigma^2}} + P_{T_j}}$$

$$\times \prod_{r_j^0 > 0, r_j^1 < 0} \frac{P_{D_j} e^{\frac{2r_j^0 h_j^0}{\sigma^2}} + P_{10H_{1j}} e^{\frac{2r_j^0 h_j^0 - 2r_j^1 h_j^1}{\sigma^2}} + P_{01H_{1j}} + P_{M_j} e^{\frac{-2r_j^1 h_j^1}{\sigma^2}}}{P_{F_j} e^{\frac{2r_j^0 h_j^0}{\sigma^2}} + P_{10H_{0j}} e^{\frac{2r_j^0 h_j^0 - 2r_j^1 h_j^1}{\sigma^2}} + P_{01H_{0j}} + P_{T_j} e^{\frac{-2r_j^1 h_j^1}{\sigma^2}}} \qquad (4.9)$$

$$\times \prod_{r_j^0 < 0, r_j^1 > 0} \frac{P_{D_j} e^{\frac{2r_j^1 h_j^1}{\sigma^2}} + P_{10H_{1j}} + P_{01H_{1j}} e^{\frac{-2r_j^0 h_j^0 + 2r_j^1 h_j^1}{\sigma^2}} + P_{M_j} e^{\frac{-2r_j^0 h_j^0}{\sigma^2}}}{P_{F_j} e^{\frac{2r_j^1 h_j^1}{\sigma^2}} + P_{10H_{0j}} + P_{01H_{0j}} e^{\frac{-2r_j^0 h_j^0 + 2r_j^1 h_j^1}{\sigma^2}} + P_{T_j} e^{\frac{-2r_j^0 h_j^0}{\sigma^2}}}$$

For high SNR namely $\sigma^2 \to 0$ the terms with $e^{\frac{\pm 2r_j^0 h_j^0 \pm 2r_j^1 h_j^1}{\sigma^2}}$ dominates other terms and the previous equation become

$$\Lambda_{highSNR}(\mathbf{r}) = \sum_{r_j^0, r_j^1 < 0} \log \frac{P_{M_j}}{P_{T_j}} + \sum_{r_j^0, r_j^1 > 0} \log \frac{P_{D_j}}{P_{F_j}} + \sum_{r_j^0 < 0, r_j^1 > 0} \log \frac{P_{01H_{1j}}}{P_{01H_{0j}}} + \sum_{r_j^0 > 0, r_j^1 < 0} \log \frac{P_{10H_{1j}}}{P_{10H_{0j}}} \quad (4.10)$$

One can see that this rule is a modified version of Chair –Varshney [4] rule for 2-bit decision. For low-SNR analysis of global fusion rule we start from

$$\Lambda\ (\mathbf{r}) = \prod_{j=1}^{N} \frac{P_{D_j} + P_{10H_{1j}} e^{\frac{-2r_j^1 h_j^1}{\sigma^2}} + P_{01H_{1j}} e^{\frac{-2r_j^0 h_j^0}{\sigma^2}} + P_{M_j} e^{\frac{-2r_j^0 h_j^0}{\sigma^2}} e^{\frac{-2r_j^1 h_j^1}{\sigma^2}}}{P_{F_j} + P_{10H_{0j}} e^{\frac{-2r_j^1 h_j^1}{\sigma^2}} + P_{01H_{0j}} e^{\frac{-2r_j^0 h_j^0}{\sigma^2}} + P_{T_j} e^{\frac{-2r_j^0 h_j^0}{\sigma^2}} e^{\frac{-2r_j^1 h_j^1}{\sigma^2}}} \qquad (4.11)$$

Since low-SNR assumption requires $\sigma^2 \to \infty$ we can write $e^{\frac{-2r_j^0 h_j^0}{\sigma^2}} \approx \left(1 - \frac{2r_j^0 h_j^0}{\sigma^2}\right)$ as first order Taylor expansion, using $P_{D_j} = 1 - P_{10H_{1j}} - P_{01H_{1j}} - P_{M_j}$ and $P_{F_j} = 1 - P_{10H_{0j}} - P_{01H_{0j}} - P_{T_j}$, (4.11) becomes

$$\Lambda(\mathbf{r}) = \prod_{j=1}^{N} \frac{\begin{array}{c}\left(1 - P_{10H_{1j}} - P_{01H_{1j}} - P_{M_j}\right) + P_{10H_{1j}}\left(1 - \frac{2r_j^1 h_j^1}{\sigma^2}\right) \\ + P_{01H_{1j}}\left(1 - \frac{2r_j^0 h_j^0}{\sigma^2}\right) + P_{M_m}\left(1 - \frac{2r_j^0 h_j^0}{\sigma^2}\right)\left(1 - \frac{2r_j^1 h_j^1}{\sigma^2}\right)\end{array}}{\begin{array}{c}\left(1 - P_{10H_{0j}} - P_{01H_{0j}} - P_{T_j}\right) + P_{10H_{0j}}\left(1 - \frac{2r_j^1 h_j^1}{\sigma^2}\right) \\ + P_{01H_{0j}}\left(1 - \frac{2r_j^0 h_j^0}{\sigma^2}\right) + P_{T_j}\left(1 - \frac{2r_j^0 h_j^0}{\sigma^2}\right)\left(1 - \frac{2r_j^1 h_j^1}{\sigma^2}\right)\end{array}} \tag{4.12}$$

Taking logarithm of both sides we have

$$\lim_{\sigma^2 \to \infty} \Lambda(\mathbf{r}) = \lim_{\sigma^2 \to \infty} \left[\sum_{j=1}^{N} \begin{array}{c} \left(P_{10H_{0j}} - P_{10H_{1j}}\right)\frac{2r_j^1 h_j^1}{\sigma^2} + \left(P_{01H_{0j}} - P_{01H_{1j}}\right)\frac{2r_j^0 h_j^0}{\sigma^2} \\ + \left(P_{T_j} - P_{M_j}\right)\left(\frac{2r_j^0 h_j^0}{\sigma^2} + \frac{2r_j^1 h_j^1}{\sigma^2} - \frac{2r_j^0 h_j^0}{\sigma^2}\frac{2r_j^1 h_j^1}{\sigma^2}\right) \end{array} \right] \tag{4.13}$$

If we assume performance indices of local sensor nodes are same global fusion rule is

$$\Lambda_{lowSNR}(\mathbf{r}) = \sum_{j=1}^{N} r_j^0 h_j^0 + \sum_{j=1}^{N} r_j^1 h_j^1 \tag{4.14}$$

4.3.2 Sub-Optimum Fusion Rule for Multi-bit Decision in Hierarchical Topology

Since we explore the asymptotic behavior of fusion rules we will assume high or low SNR for all wireless channels in WSN. If high-SNR assumption is valid between local sensor nodes and cluster heads fusion rule in m^{th} cluster head will be

$$\Lambda_{highSNR}^{m}(\mathbf{r}) = \sum_{r_{1j}^m, r_{2j}^m < 0} \log \frac{P_{M_j}^m}{P_{T_j}^m} + \sum_{r_{1j}^m, r_{2j}^m > 0} \log \frac{P_{D_j}^m}{P_{F_j}^m} + \sum_{r_{1j}^m < 0, r_{2j}^m > 0} \log \frac{P_{01H_{1j}}^m}{P_{01H_{0j}}^m} + \sum_{r_{1j}^m > 0, r_{2j}^m < 0} \log \frac{P_{10H_{1j}}^m}{P_{10H_{0j}}^m} \quad (4.15)$$

Again fusion rule in global fusion center is

$$\Lambda_{highSNR}^{0}(\mathbf{y}) = \sum_{y_m^0, y_m^1 < 0} \log \frac{P_{M_m}}{P_{T_m}} + \sum_{y_m^0, y_m^1 > 0} \log \frac{P_{D_m}}{P_{F_m}} + \sum_{y_m^0 < 0, y_m^1 > 0} \log \frac{P_{01H_{1m}}}{P_{01H_{0m}}} + \sum_{y_m^0 > 0, y_m^1 < 0} \log \frac{P_{10H_{1m}}}{P_{10H_{0m}}} \quad (4.16)$$

Assuming low-SNR for all channels in the WSN in m^{th} cluster head fusion rule is given as

$$\Lambda_{lowSNR}^{m}(\mathbf{r}) = \sum_{j=1}^{K_m} r_{1j}^m h_{1j}^m + \sum_{j=1}^{K_m} r_{2j}^m h_{2j}^m \quad (4.17)$$

and low-SNR fusion rule in global fusion center

$$\Lambda_{lowSNR}^{0}(\mathbf{y}) = \sum_{m=1}^{N} y_m^0 g_m^0 + \sum_{m=1}^{N} y_m^1 g_m^1 \quad (4.18)$$

In this low-SNR case global fusion center and cluster head should use (4.17) and (4.18) respectively to fuse the information came from previous component of network.

4.4 Optimum Fusion Rule for Multi-bit Decision only with Channel Statistics

Achieving channel gain information in a WSN spends precious resources as power, to keep away from consuming power we in this section we try to

4.4.1 Optimum Fusion Rule for Multi-bit Decision in Parallel Topology only with Channel Statistics

As mentioned in section 3.3 to decrease the required prior information at the fusion center we will try to obtain optimum fusion rule with CS. Likelihood ratio test in global fusion center is

$$\frac{f\left(\mathbf{r}\middle|H_1\right)}{f\left(\mathbf{r}\middle|H_0\right)}=\prod_{j=1}^{N}\frac{f\left(r_j^0,r_j^1\middle|H_1\right)}{f\left(r_j^0,r_j^1\middle|H_0\right)} \tag{4.19}$$

Where

$$f\left(r_j^0,r_j^1\middle|H_i\right)=\sum_{x_j^0,x_j^1}p\left(x_j^0,x_j^1\middle|H_i\right)f\left(r_j^0,r_j^1\middle|x_j^0,x_j^1\right)\qquad i\in\{0,1\} \tag{4.20}$$

When x_j^0,x_j^1 are given r_j^0,r_j^1 are independent therefore distribution becomes

$$f\left(r_j^0,r_j^1\middle|x_j^0,x_j^1\right)=f\left(r_j^0\middle|x_j^0\right)f\left(r_j^1\middle|x_j^1\right) \tag{4.21}$$

Pdf's are given in [14] as

$$f\left(r_j^k\middle|x_j^k=\pm 1\right)=\frac{2\sigma}{\sqrt{2\pi}\left(1+2\sigma^2\right)}e^{-\frac{\left(r_j^k\right)^2}{2\sigma^2}}\left[1\pm\sqrt{2\pi}\left(\frac{1}{\sigma\sqrt{\left(1+2\sigma^2\right)}}\right)e^{-\frac{\left(\frac{r_j^k}{\sigma\sqrt{\left(1+2\sigma^2\right)}}\right)^2}{2\sigma^2}}Q\left(\frac{\mp r_j^k}{\sigma\sqrt{\left(1+2\sigma^2\right)}}\right)\right] \tag{4.22}$$

Global fusion rule can be obtained putting(4.22),(4.21) and(4.20) in (4.19)

$$\Lambda_{CS} = \sum_{j=1}^{N} \log \frac{P_{D_j} A^2 + \left(P_{01H_{1j}} + P_{01H_{1j}}\right) AB + P_{M_j} B^2}{P_{F_j} A^2 + \left(P_{01H_{0j}} + P_{01H_{0j}}\right) AB + P_{T_j} B^2} \tag{4.23}$$

Where

$$A = \left[1 + \sqrt{2\pi} \left(\frac{1}{\sigma\sqrt{\left(1+2\sigma^2\right)}} \right) e^{-\frac{\left(\frac{r_j^k}{\sigma\sqrt{\left(1+2\sigma^2\right)}}\right)^2}{2\sigma^2}} Q\left(\frac{-r_j^k}{\sigma\sqrt{\left(1+2\sigma^2\right)}} \right) \right]$$

$$B = \left[1 - \sqrt{2\pi} \left(\frac{1}{\sigma\sqrt{\left(1+2\sigma^2\right)}} \right) e^{-\frac{\left(\frac{r_j^k}{\sigma\sqrt{\left(1+2\sigma^2\right)}}\right)^2}{2\sigma^2}} Q\left(\frac{r_j^k}{\sigma\sqrt{\left(1+2\sigma^2\right)}} \right) \right] \tag{4.24}$$

Finally we can use a heuristic equal gain combiner like fusion rule when channel state information is not available and it can be given as

$$\Lambda_{EGC}\left(\mathbf{y}\right) = \sum_{m=1}^{N} r_j^0 + \sum_{m=1}^{N} r_j^1 \tag{4.25}$$

4.4.2 Optimum Fusion Rule for Multi-bit Decision in Hierarchical Topology only with Channel Statistics

Optimum cluster head fusion rule for cluster heads have only statistics information about channel will be in the same form with (4.23)

$$\Lambda_{CS}^m = \sum_{j=1}^{K_m} \log \frac{P_{D_j}^m C^2 + \left(P_{01H_{1j}}^m + P_{10H_{1j}}^m\right) CD + P_{M_j}^m D^2}{P_{F_j}^m C^2 + \left(P_{01H_{0j}}^m + P_{10H_{0j}}^m\right) CD + P_{T_j}^m D^2} \tag{4.26}$$

Where

$$C=\left[1+\sqrt{2\pi}\left(\frac{1}{\sigma\sqrt{\left(1+2\sigma^2\right)}}\right)e^{-\frac{\left(\frac{r_{k,j}^m}{\sigma\sqrt{\left(1+2\sigma^2\right)}}\right)^2}{2\sigma^2}}Q\left(\frac{-r_{k,j}^m}{\sigma\sqrt{\left(1+2\sigma^2\right)}}\right)\right]$$

$$D=\left[1-\sqrt{2\pi}\left(\frac{1}{\sigma\sqrt{\left(1+2\sigma^2\right)}}\right)e^{-\frac{\left(\frac{r_{k,j}^m}{\sigma\sqrt{\left(1+2\sigma^2\right)}}\right)^2}{2\sigma^2}}Q\left(\frac{r_{k,j}^m}{\sigma\sqrt{\left(1+2\sigma^2\right)}}\right)\right] \quad (4.27)$$

We can conclude that optimum performance without exact channel state information can be achieved using (4.27) in cluster heads and following equation in global fusion center

$$\Lambda_{CS}^0=\sum_{m=1}^{N}\log\frac{P_{D_m}A^2+\left(P_{01H_{1m}}+P_{01H_{1m}}\right)AB+P_{M_m}B^2}{P_{F_m}A^2+\left(P_{01H_{0m}}+P_{01H_{0m}}\right)AB+P_{T_m}B^2} \quad (4.28)$$

where

$$A=\left[1+\sqrt{2\pi}\left(\frac{1}{\sigma\sqrt{\left(1+2\sigma^2\right)}}\right)e^{-\frac{\left(\frac{y_m^k}{\sigma\sqrt{\left(1+2\sigma^2\right)}}\right)^2}{2\sigma^2}}Q\left(\frac{-y_m^k}{\sigma\sqrt{\left(1+2\sigma^2\right)}}\right)\right]$$

$$B=\left[1-\sqrt{2\pi}\left(\frac{1}{\sigma\sqrt{\left(1+2\sigma^2\right)}}\right)e^{-\frac{\left(\frac{y_m^k}{\sigma\sqrt{\left(1+2\sigma^2\right)}}\right)^2}{2\sigma^2}}Q\left(\frac{y_m^k}{\sigma\sqrt{\left(1+2\sigma^2\right)}}\right)\right] \quad (4.29).$$

Also, we can use our heuristic rule in also cluster heads to get a simple fusion rule as

$$\Lambda^{m}_{EGC}\left(\mathbf{r}\right)=\sum_{j=1}^{K_m} r_{1j}^{m}+\sum_{j=1}^{K_m} r_{2j}^{m} \tag{4.30}$$

In cluster heads and in global fusion center we can use

$$\Lambda^{0}_{EGC}\left(\mathbf{y}\right)=\sum_{m=1}^{N} y_{m}^{0}+\sum_{m=1}^{N} y_{m}^{1} \tag{4.31}$$

4.5 Performance Evaluation

In this section we will give some numerical simulation results which show the detection performance of derived fusion rules for multi-bit decision in local sensor nodes and cluster heads. 2-bit decision is assumed for this simulation therefore transmission power is divided by two for comparing single-bit decision strategies. 5dB average SNR is assumed in simulations for single-bit which means 2 dB for 2-bit decision. Figure 4.1 shows 2-bit LRT has the best detection performance and 2-bit LRT, 2-bit LRTCS, and 2-bit EGC outperforms single-bit LRT .Also for this scenario 2-bit high-SNR approximation has worse performance than 2-bit low-SNR approximation for LRT since the transmission power is halved SNR of per channel decreased 3 dB .

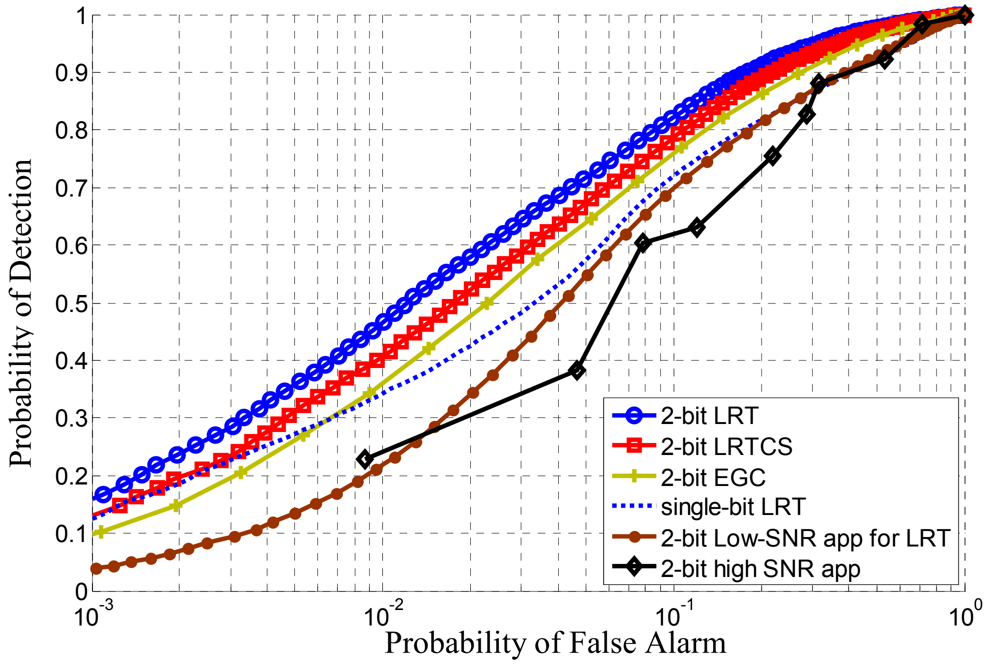

Figure 4.1 ROC curves of 2-bit fusion rules when $P_{D_j}^m = 0.5, P_{F_j}^m = 0.05$ and $P_{M_j}^m = 0.2$

5. DISTRIBUTED DETECTION USING FUSION CENTER DIVERSITY

In this chapter, we model the communication between local sensor nodes and fusion center as Multiple Input Multiple Output (MIMO) communication by deploying multiple antennas at the receiver namely fusion center. Using the same amount of transmit power WSN can achieve a better detection performance when the communication is more reliable. This motivates our work and we derive the optimum fusion rules for fusion center when CSI is available. We will analyze usage of MIMO in Distributed Detection and derive optimum fusion rule for this scenario, and give some numerical simulations and performance evaluation.

5.1 MIMO in Distributed Detection

MIMO is the use of multiple antennas at transmitter and receiver to improve communication performance [16]. It improves performance by increasing the spectral efficiency and link reliability. Communication channels in a WSN are not error free due to fading and noise and this decreases the detection performance. Fig. 2 shows simplest form of MIMO communication for WSN which includes 2 sensors and 2 antennas deployed in fusion center.

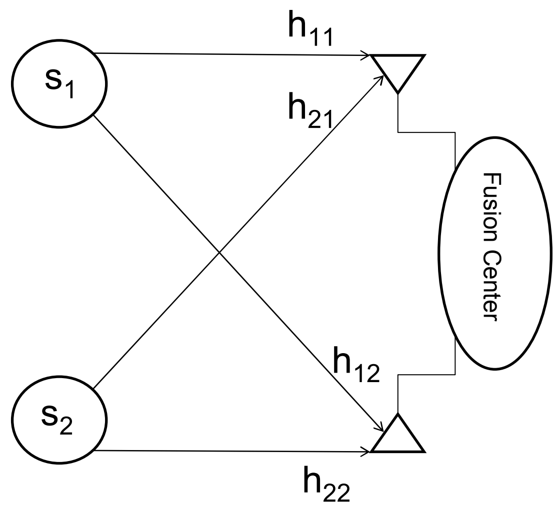

Figure 5.1 An example for MIMO communication in WSN

In Figure 5.1 s_1 and s_2 are local sensor nodes and h_{ij}'s are gains corresponding to channel between i^{th} sensor and j^{th} antenna at the fusion center. For a M_t transmitter and M_r receiver antenna MIMO system the received vector can be written as

$$\mathbf{r} = \mathbf{Hx} + \mathbf{n} \tag{5.1}$$

$\mathbf{r}$ is the received signal vector, $\mathbf{x}$ is the BPSK modulated signal vector depends on the decisions of the local sensor nodes, $\mathbf{H}$ is the $M_r \times M_t$ channel gain matrix consisting h_{ij}'s and $\mathbf{n}$ is the noise vector where n_j's are AWGN sample with zero mean and variance $\sigma_{n_i}^2$ at the j^{th} receiver antenna .

Optimum Likelihood Ratio Test (LRT) based fusion statistics is given as [9]

$$\Lambda(\mathbf{r}) = \frac{f(\mathbf{r}|H_1)}{f(\mathbf{r}|H_0)} \tag{5.2}$$

For a fusion center with M_r receiver antennas, assuming r_j's are conditionally independent when x_j's are given, fusion statistics in Eq.(5.2) become

$$\Lambda(\mathbf{r}) = \frac{\sum_{\mathbf{x}} f(r_1|H_1,\mathbf{x}) f(r_2|H_1,\mathbf{x}) ... f(r_{M_r}|H_1,\mathbf{x}) p(\mathbf{x}|H_1)}{\sum_{\mathbf{x}} f(r_1|H_0,\mathbf{x}) f(r_2|H_0,\mathbf{x}) ... f(r_{M_r}|H_0,\mathbf{x}) p(\mathbf{x}|H_0)} \tag{5.3}$$

When performance indices and Channel State Information (CSI) available at fusion center, distributions in Eq.(5.3) can be derived, as an example $f(r_1|H_1)$ is obtained as follows

$$f(r_1|H_1, x_1, x_2, ..., x_{M_t}) = \prod_{i=1,\mathbf{u}=\mathbf{u_k}}^{M_t} (P_{D_i})^{u_i} (1-P_{D_i})^{1-u_i} N\left(r_1; \sum_{j=1}^{M_t} (-1)^{u_i+1} h_{j,1}; \sigma_n\right) \tag{5.4}$$

$$f(y_1|H_1) = \sum_{k=1}^{2^{M_t}} \prod_{i=1,\mathbf{u}=\mathbf{u_k}}^{M_t} (P_{D_i})^{u_i} (1-P_{D_i})^{1-u_i} N\left(r_1; \sum_{j=1}^{M_t} (-1)^{u_i+1} h_{j,1}; \sigma_n\right), \tag{5.5}$$

where $\mathbf{u}$ one realization of local sensor node is decision among 2^{M_t} possible ones and u_i is the decision of i^{th} local sensor node. P_{D_i} 's are the detection probabilities of local sensor nodes and $N(x;\mu;\sigma)$ denotes normal distribution with μ mean and σ standard deviation with parameter x. Therefore optimum LRT fusion rule for M_t local sensor node and fusion center with M_r antennas is

$$\Lambda = \frac{\prod_{l=1}^{M_r} \sum_{k=1}^{2^{M_t}} \prod_{i=1,\mathbf{u}=\mathbf{u_k}}^{M_t} (P_{D_i})^{u_i} (1-P_{D_i})^{1-u_i} N\left(r_l; \sum_{j=1}^{M_t} (-1)^{u_i+1} h_{jl}; \sigma_l\right)}{\prod_{l=1}^{M_r} \sum_{k=1}^{2^{M_t}} \prod_{i=1,\mathbf{u}=\mathbf{u_k}}^{M_t} (P_{F_i})^{u_i} (1-P_{F_i})^{1-u_i} N\left(r_l; \sum_{j=1}^{M_t} (-1)^{u_i+1} h_{jl}; \sigma_l\right)}. \tag{5.6}$$

5.2 Performance Evaluation

In this section, we compare the performances of the optimal global fusion rules for MIMO and conventional LRT. Throughout our discussion, we assume that sensors are identical in terms of their false alarm and detection probabilities as $P_{F_j} = 0.05$ and $P_{D_j} = 0.5$. Since we normalize the average power of Rayleigh fading gains to 1, the signal-to-noise ratio (SNR) is defined as

$$SNR(\text{dB}) = 10\log_{10}\left(\frac{1}{\sigma^2}\right) \quad (5.7)$$

We first obtain the Receiver Operating Characteristics (ROC) curves of the global fusion center for the simplest cases when $M_t = 2$ and $M_r = 1,2,3$ cases that can be seen in Figure 5.2 when $SNR = 0$ dB and total transmission power is kept same.

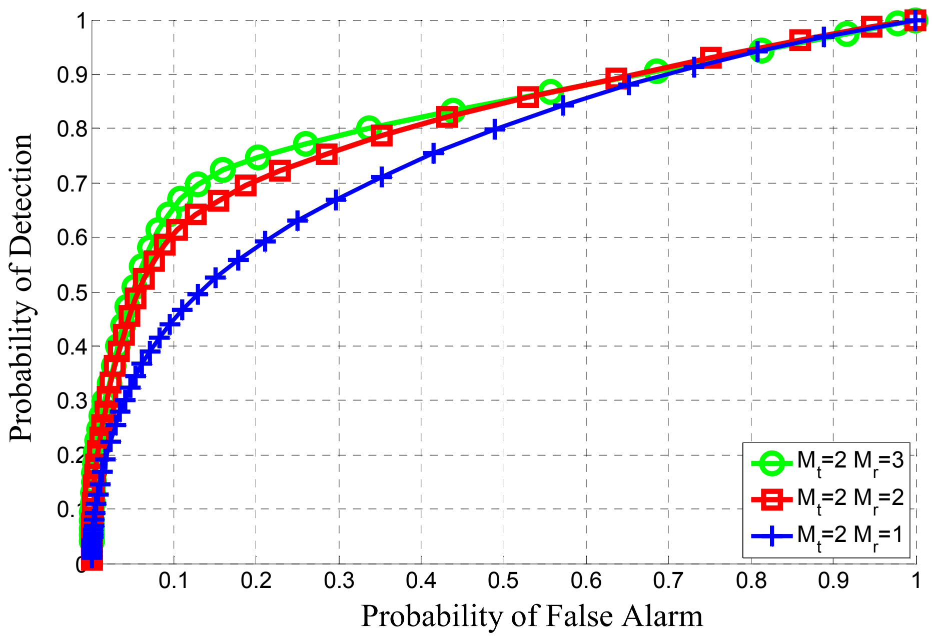

Figure 5.2 ROC curves for different scenarios with 2 sensors and 0 dB average SNR

Figure 5.2 shows that scenario with multiple antennas in the fusion center outperforms the conventional scenario when global False Alarm probability smaller than 0.6 and after 0.8 MIMO schemes have the same detection performance with 2 sensors and 1 receiver antenna case. Next, we explore the probability of detection behaviors of these different scenarios as a function of average SNR of the wireless channel. Figure 5.3 shows the detection probability of fusion when global false alarm probability is fixed at $P_{F_0} = 0.1$

depending on SNR. It can be seen from Figure 5.3 deploying antennas at the fusion center increases the performance more efficiently at low SNRs.

As the number of the local sensor nodes is increased global detection performance of WSN using MIMO increases dramatically. Difference between system performances can be seen in Figure 5.4. There are 8 local sensor nodes in this example and even deploying one extra antenna at the fusion center remarkably increases the detection probability, but deploying one more antenna makes only slight difference in system performance.

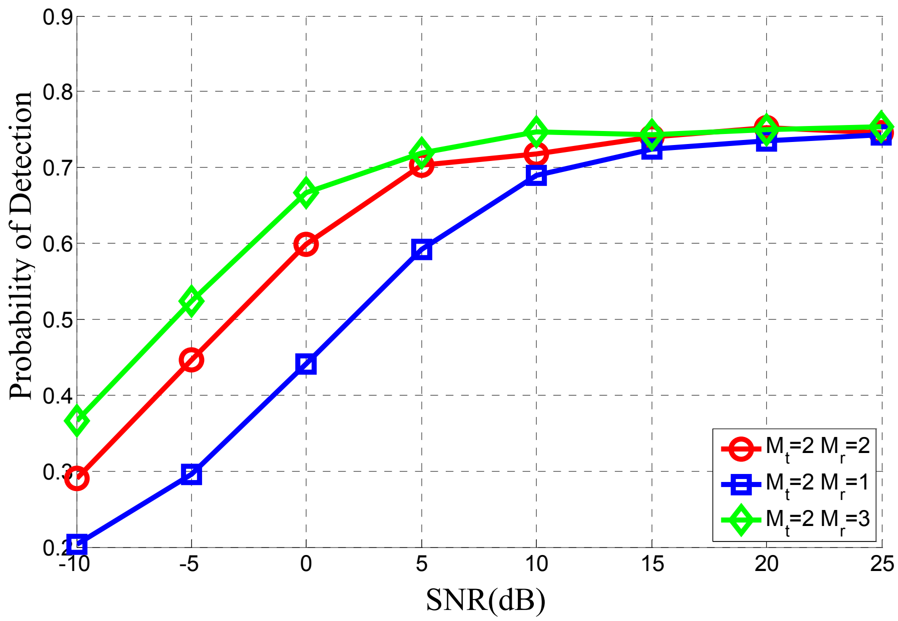

Figure 5.3 Probability of Detection as a function of average channel SNR when the global false alarm probability is fixed at $P_{F_0} = 0.1$

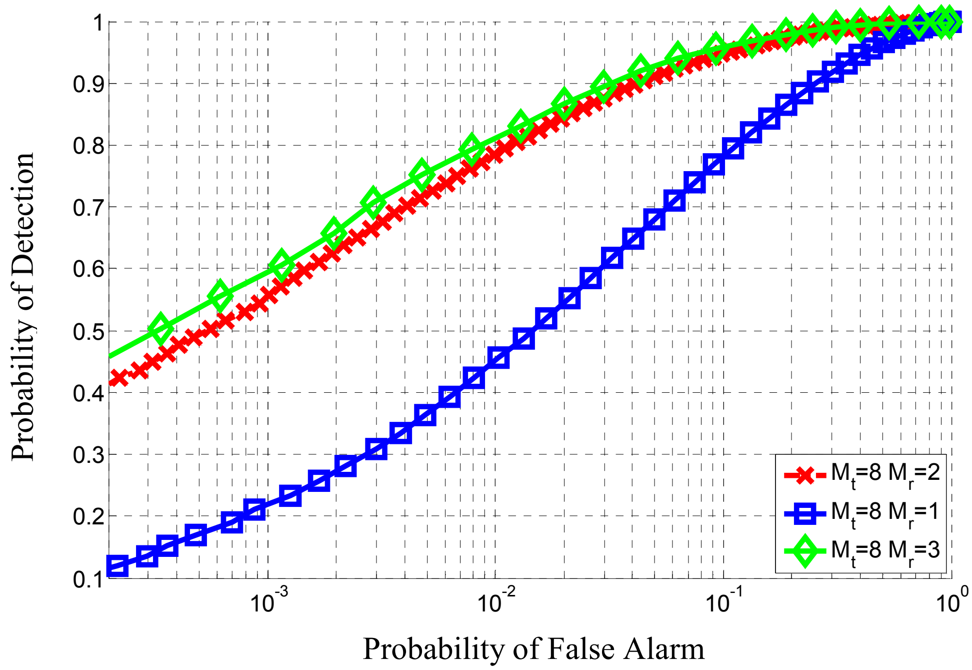

Figure 5.4 ROC curves for different scenarios with 8 sensors and 0 dB average SNR

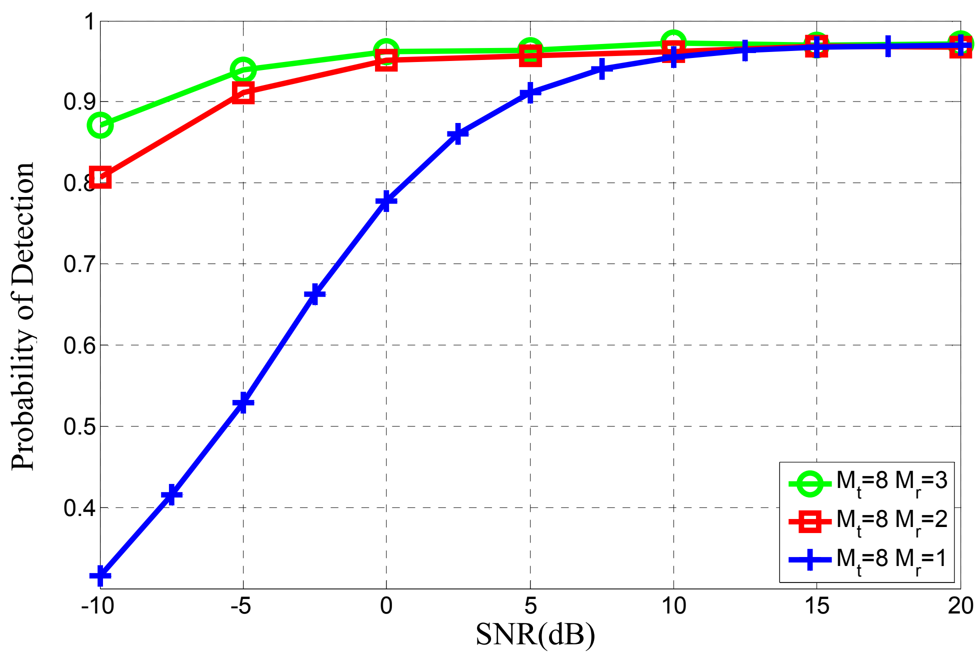

Figure 5.5 Probability of Detection as a function of average channel SNR when the global false alarm probability is fixed at $P_{F_0} = 0.1$

Again, Figure 5.5 shows the detection performance as a function of SNR (dB) for a WSN with 8 local sensor nodes and global false alarm probability is fixed at $P_{F_0} = 0.1$.For 8 local sensor node case detection probability for same SNR increases remarkably, and keeping the transmission power same (2 local sensor node case has 6 dB greater SNR) if we compare 8 local sensor node case and 2 local sensor node case, we can see that 8 local sensor node with multiple antenna outperforms 2 local sensor node case.

6. CONCLUSION and FUTURE WORK

Wireless sensor networks have promising applications in medical, environmental, commercial and military applications therefore it draws too much attention in recent days. Detection is one of the major jobs to be performed by a WSN. Due to restriction of communication sources such as power and bandwidth it is more convenient that a WSN should use distributed detection. In this thesis, various distributed detection strategies are investigated.

In the second chapter firstly distributed detection strategies under Neyman-Pearson criteria in the literature is briefed. We have derived optimum and suboptimum fusion rules when CSI or only CS is available at the fusion center. Numerical simulations are done to investigate the performance of these fusion rules.

In the third chapter we have investigated detection problem of a phenomenon in wireless sensor networks with hierarchical topology. We derived optimum and suboptimum rules for this topology by analyzing the high and low SNR behavior of the optimum rule in both knowing channel gains and knowing only the channel statistics. We have shown that MRC and EGC is low SNR approximation for LRT and LRTCS and Chair-Varsney rule is a high SNR approximation for LRT as in parallel topology. We also derived pdf for the LRTCS with a proper approximation. As in previous works this work shows the fusion rule which requires much prior information has better detection performance.

Multi-bit decisional distributed detection for WSNs that have hierarchical and parallel topology is studied in chapter four. Again, optimum and sub-optimum fusion rules are derived for both cases we have CSI or only CS at global fusion center and cluster heads.

Also, a heuristic fusion rule like EGC based on our knowledge from previous chapter is proposed. Performance of fusion rules are shown via numerical simulations. An outstanding remark is, besides expected ones, our heuristic fusion rule which does not need any prior information outperforms single-bit LRT too.

In the last chapter, the detection problem of a phenomenon with WSNs using multiple antennas at the fusion center is analyzed. Optimum fusion rule is derived by modeling the WSN as a MIMO system. Performance evaluation is done throughout numerical simulations and it is shown that deploying multiple antennas at the receiver increases the detection performance. Simulation results show us that if more than one antenna are deployed in the fusion center, improvement in detection performance is inversely proportional with average channel SNR. Beside that, we cannot obtain the same performance improvement when we continue increase the number of antennas at the fusion center.

For future work by letting cluster heads make their own observation from environment fusion rules based on their own observations and signals coming from local sensor nodes can be derived. Optimum rule for multi-bit decision with only CS can be analyzed for low and high SNR assumptions, analysis can end up with a new fusion rule or it can converge one of derived rules. Simplifying the optimum fusion rule proposed in the last chapter and determining optimum number of antennas at the fusion center as a function of number of local sensors can be interesting to research.

References

[1] Akyildiz I. F., Su W., Sankarasubramaniam Y., Cayirci E., "A survey on sensor networks,"*IEEE Communications Magazine,* pp. 102–114, August 2002.

[2] Jason Lester Hill. System Architecture for Wireless Sensor Networks. PhD thesis, University of California Berkeley, 2003. Available from: http://www.jlhlabs.com/jhill_cs/jhill_thesis.pdf.

[3] Varshney P.K.*, "Distributed Detection and Data Fusion,"*Springer-Verlag,1996.

[4] Chair Z., Varshney P.K.*, "Optimal data fusion in multiple sensor detection system,"* IEEE Trans. Aerospace Elect. Syst., vol. 22 pp.98-101, Jan 1986.

[5] S. C. A. Thomopoulos, R. Viswanathan, and D. K. Bougoulias, *"Optimal distributed decision fusion,"* IEEE Trans. Aerospace Elect. Syst., vol. 25, pp. 761-765, Sept 1989.

[6] Hoballah I.Y., Varshney P.K., "Distributed Bayesian signal detection," *IEEE Trans. Inform.Theory*, vol. 35, pp. 995–1000, Sept. 1989.

[7] Drakopoulos, E. and Lee, C. C., *"Optimal multisensor fusion of correlated local decisions,"* IEEE Trans. Aerospace Elect. Syst., vol. 27, pp. 593–605, July 1991.

[8] Thomopoulos, S. C. and Zhang, L., *"Distributed decision fusion with networking delays and channel errors,"* Inform. Sci., vol. 66, pp. 91–118, Dec. 1992.

[9]Chen B., Jiang R., Kasetkasem T, and Varshney P.K.*, " Channe aware decision fusion in wireless sensor networks,"* IEEE Trans. Signal Process. Vol.52, no12, pp.3454-3458, Dec 2004.

[10] Bahceci I., Altunbasak Y., AlRegib G., *"Parallel distributed detection for wireless sensor networks: performance analysis and design,"* IEEE GLOBECOM Conference,Volume 4, Page(s):2420 – 2424, 28 Nov.-2 Dec. 2005.

[11]MATLAB , http://www.mathworks.com

[12] Gander, W., Gautschi W., "*Adaptive Quadrature - Revisited*", BIT, Vol. 40, 2000, pp. 84-101.

[13] Forsythe, G. E., Malcolm M. A., Moler C. B., "*Computer Methods for Mathematical Computations*," Prentice-Hall, 1976

[14] Niu R., Chen B., Varshney P. K., *"Fusion of decisions transmitted over Rayleigh fading channels in wireless sensor networks,"* IEEE Trans.Signal Processing, vol. 54, no. 3, pp. 1018-1027, Mar. 2006.

[15] A. Leon-Garcia*," Probability and Random Processes for Electrical Engineering,"* 2nd ed. Reading, MA: Addison-Wesley, 1994.

[16] A. Goldsmith, "*Wireless Communications."* New York : Cambridge University Press,2005