2007

Optimum energy allocation for detection in wireless sensor networks

Krishna Kishore Gunturu
Louisiana State University and Agricultural and Mechanical College

OPTIMUM ENERGY ALLOCATION FOR DETECTION IN WIRELESS SENSOR NETWORKS

Thesis

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Master of Science in Electrical Engineering

in

The Department of Electrical Engineering

by
Krishna Kishore Gunturu
B.S. in Electronics and Communications Engineering., Andhra University, India 2005
December, 2007

Acknowledgments

I would like to express my gratitude to my advisor, Dr. Morteza Naraghi Pour for his guidance and motivation towards the completion of this thesis. His technical advice and suggestions helped me overcome hurdles and kept me enthusiastic and made this a wonderful learning experience. I would like to thank him for being very supportive and patient with me during my thesis and write up. I have learned a lot from him which will help me in my academic and professional career.

I would like to thank my committee members Dr. Jorge Aravena and Dr. Xue-Bin Liang for taking time out of their busy schedule and agreeing to be a part of my committee.

I would like to thank all my friends at LSU who made my stay at LSU pleasant and enjoyable. I would like to thank Chetan Chitnis for helping me with my write up.

Finally, I would like to thank my parents and sister for the support and unconditional love they provided me throughout my life. They have supported me in every decision I made in my life. I hope the completion of my degree will be a stepping stone for my academic and professional career.

Table of Contents

List of Tables

List of Figures

Abstract

The problem of binary hypothesis testing in a wireless sensor network is studied in the presence of noisy channels and for non-identical sensors. We have designed a mathematically tractable fusion rule for which optimal energy allocation for individual sensors can be achieved. In this thesis we considered two methods for transmitting the sensor observations; binary modulation and M-ary modulation. In binary modulation we are able to allocate the energy among the sensors and protect the individual quantized bits where as the M-ary modulation provides optimum energy allocation only among the sensors. The goal is to design a fusion rule and an energy allocation for the nodes subject to a limit on the total energy of all the nodes so as to optimize a cost function. Two cost functions were considered; the probability of error and the J-divergence distance measure. Probability of error is the most natural criteria used for binary hypothesis testing problem. Distance measure is applied when it is difficult to obtain a closed form for the error probability. Results of optimal energy allocation and the resulting probability of error are presented for the two cost functions. Comparisons are drawn between the two cost functions regarding the fusion rule, energy allocations and the error probability.

Chapter 1
Introduction

1.1 Overview

With the development of micro electromechanical systems (MEMS), sensors can be made smaller and cheaper [3]. This along with the advances in low power VLSI, digital signal processing and low manufacturing costs have lead to the development of wireless sensor networks (WSN) [17], [15]. These technologies allow for development of small and inexpensive wireless sensor nodes, which can be easily distributed over a large geographic area. The nodes can collect information and relay that information to a center where the information is processed to make an appropriate decision. In this way wireless sensor networks can be used for environmental monitoring (temperature, pressure, and pollution levels), situation awareness, intrusion detection and denial of access, to name a few. They can be the first line of defense in many applications where access is limited such as detection of biological hazards, chemical spills, health monitoring, fire detection, etc.

In many of the above detection problems, a decision needs to be made among a set of possibilities. In particular in many applications we are interested in a decision between two alternatives (e.g, presence or absence of a chemicals, an intruder, etc). In detection, such a problem is referred to as binary hypothesis testing, where a decision must be made between two alternatives, H_1 and H_0. Hypothesis H_1 represents the presence of the target while H_0 corresponds to the absence of the target. In more general cases with multiple alternatives, we have a multi-hypothesis testing problem where a decision is made in favor of M hypothesis $H_0, H_1, \cdots, H_{M-1}$ whose prior probabilities are denoted by $P(H = H_i) = p_i,\ i = 0, 1, \cdots, M-1$.

1.2 Topologies

In wireless sensor networks the measurements are made at the sensors whereas the decisions are made at the fusion center. Wireless sensor networks can be organized in a number of configurations depending on the distribution of the sensors and the fusion center. The three major topologies used for distributed detection are parallel, serial and tree topologies. In a parallel configuration (Fig.1.1), there is no communication between the individual sensors . In the parallel topology , the sensor i passes the information b_i to the fusion center where the decision is made by the fusion center based on the received information $(b_1, b_2, \cdots, b_L)$.

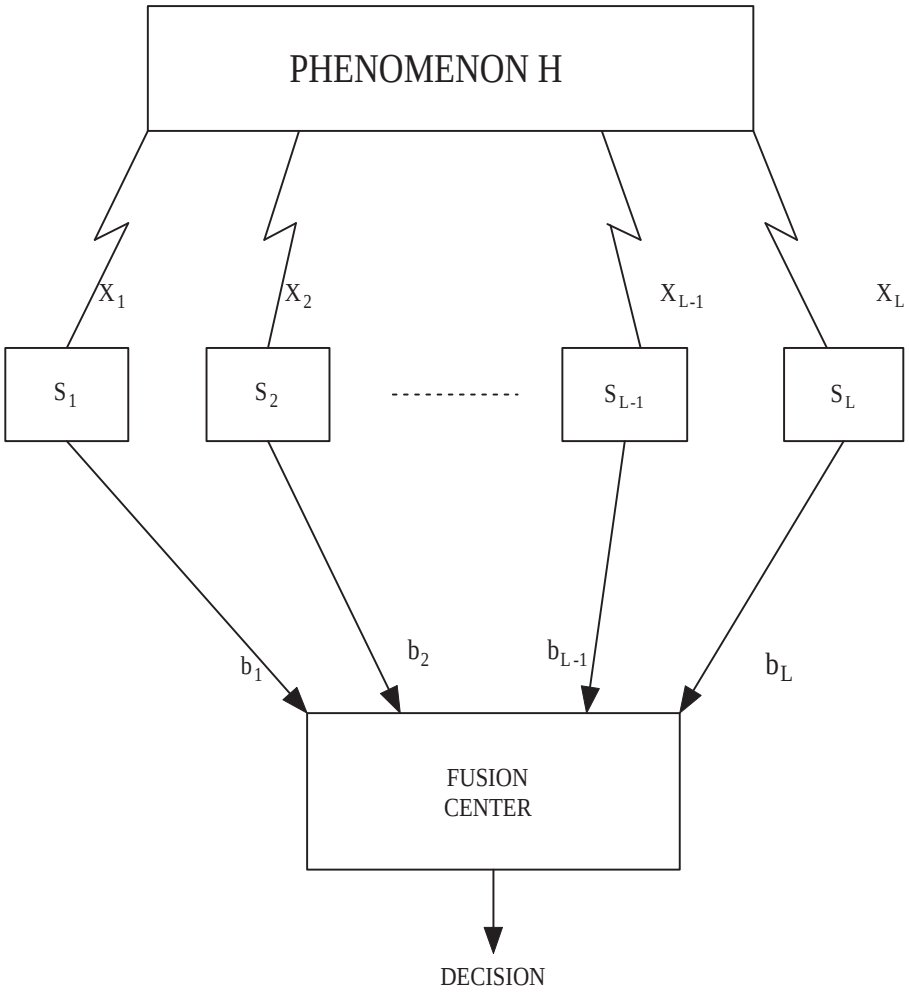

FIGURE 1.1. Parallel Topology of WSN

In a serial or tandem topology (Fig.1.2), all the sensors are connected in series and make observations of the common phenomenon. Each sensor makes a decision about the hypothesis based on the information it receives and forwards that decision to the next sensor. The decision of the first sensor is completely based on its own observation. This decision is transmitted to the second sensor which uses it along with its own observation to make a decision and transmits it to the adjacent sensor. The output of the last sensor is the final decision about the observed phenomenon.

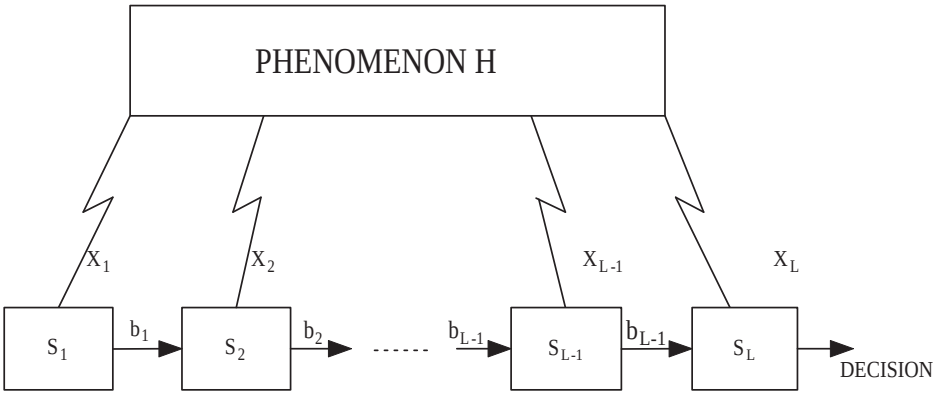

FIGURE 1.2. Series Topology of WSN

In the tree topology (Fig.1.3), the decision at each sensor is made based on its own observation and the decisions from its immediate predecessor. This decision is transmitted to its immediate successor. The information from the sensors flows on a unique path to the final center, which forms the root of the tree. Work has been done on detection for the tree topology in [25]. In the topologies discussed so far , the information is transmitted only in the direction of the fusion center. The sensors observe the phenomenon and transmit their decision toward s the fusion center.

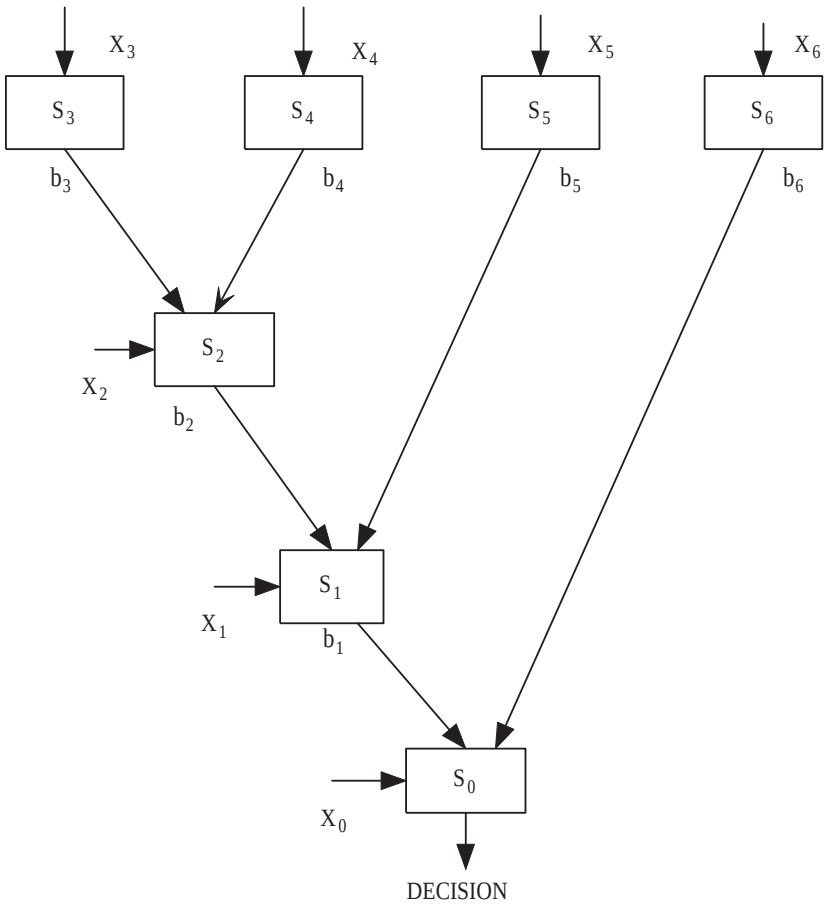

FIGURE 1.3. Tree Topology of WSN

1.3 Literature

In our thesis we will be considering the parallel configuration with a fusion center. The observations made at the sensor nodes are transmitted to the fusion center for the final decision. The local observation at the node is often a real valued measurement corrupted by noise. In classical theory of sensor signal processing [12], the detection is performed based on the real valued observation transmitted by the sensors. This type of detection is referred as centralized detection.

An alternative approach called decentralized detection is introduced in [31]. In this case, the sensors quantize the real valued measurements before transmission to the fusion center in order to reduce the communication bandwidth requirements. Such quantization degrades the performance of the fusion decision rule and thus leads to a trade off between the quality of the final decisions and the communication costs. Thus decentralized scheme is proffered in cases where bandwidth is limited. A quantizer (Fig.1.4) maps the real valued measurement into a finite set of output values $\{q_j\}$ based on the threshold values $\{t_j\}$. Based on the distance between the thresholds, quantizers are classified into uniform and non-uniform quantizer. In a uniform quantizer, the thresholds are equally spaced and the output values are at the center of the intervals, while in a non uniform quantizer the thresholds and output values are optimized based on the optimization of a cost function. A quantizer is defined using the quantizer outputs q_i and thresholds t_i as follows:

$$q: \ \Re \longrightarrow \{q_1, q_2, \cdots, q_M\}$$

Where

$$Q(X) = q_i, \text{ if } t_{i-1} \leq X \geq t_i$$

where $t_0 = -\infty, \ t_{M+1} = \infty$.

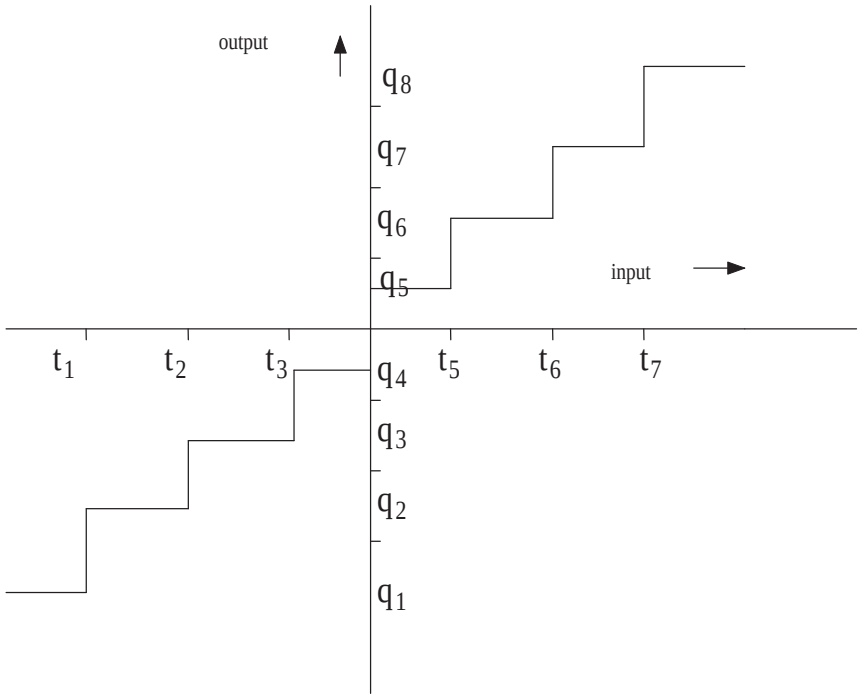

FIGURE 1.4. Quantizer

In [31], [32], [34], [36] and [33], the observations made at the sensors are quantized to a single bit which may be considered as the sensors local decisions. In [33] the author shows that using likelihood-ratio test as the local decision at the sensor nodes is optimal when the observations are conditionally independent given the hypothesis. The authors consider identical local decisions for all the sensors. Shannon-Gallagher-Berklekamp lower bound is applied to prove that using identical local decision rules at the sensors is optimal when the observations are independent and identically distributed [33]. However counter examples for which non-identical local decisions are optimal have also been identified [32].

In a binary hypothesis testing a sensor transmits a one if its decision is in favor of H_1 and a zero if its decides in favor of H_0. The fusion center then makes a final decision based on the local decisions transmitted by the sensors. Assuming that the two hypothesis are equally likely, the optimal decision rule at the sensors is a

majority rule on the decisions of the sensors. The decision is made in favor of H_1 if more ones are received and in favor of H_0 otherwise. In [31]-[33], the sensor nodes and the fusion center are designed to minimize the probability of error for the final decision at the fusion center.

Designing multi-level quantizer for binary hypothesis testing is difficult because the probability of error does not yield a tractable solution. In such cases, Ali-Silvey Distance measures [1] such as the J-divergence, Matsusita distance or the Bhattacharya distances have been employed as the cost function because of the theorems relating the maximum distance values to the minimum probability of error. In [26], multi-level quantizer for binary hypothesis testing is designed based on the maximization of the distance measures. To be specific, four-level quantizer is designed for the detection of signals corrupted by additive Gaussian noise.

The work done on decentralized detection until now has not taken into account important features of the sensor networks and of the wireless channel between the sensors and the fusion center. Perfect reception of the sensor output is assumed at the fusion center neglecting the effects of the wireless channel. There has been great deal of interest among the research community with regards to the inclusion of resource constraints such as power, cost and spectral bandwidth into the binary hypothesis problem.

The assumption of reliable transmission fails with the incorporation of wireless channel between the sensors and the fusion center. This limitation is made worse by the consideration of the stringent delay constraints. In [27], the author proposes a scheme where the information from neighboring nodes is combined through inter network signal processing to improve the reliability of the network. This paper ([27]) considers the use of feedback, retesting and rebroadcasting of the updated decisions to make the sensors arrive at a particular consensus. Fast and optimum

are the two modes of operation in [27]. In a fast mode, a decision is reached rapidly while in a optimum mode, consensus is reached after several rounds of information sharing.

In [4], the authors include noisy channels between the sensors and the fusion center for a binary hypothesis testing problem. The sensors have identical local decisions because the observations made by them are assumed as independent and identically distributions (i.i.d) [33]. The channel between a sensor and the fusion center is modeled by a binary symmetric channel. In [4], the sensor nodes and the fusion center are designed by minimizing the probability of error.

The assumption of independent observations makes the problem convenient and tractable for analysis but may not hold for arbitrary systems. For example, the sensor observations are correlated when they are located close to each other. In [35], the authors worked on the problem of decentralized detection with correlated sensors. The sensor nodes are placed in a straight line with a constraint on power per unit distance. The observations at the sensors are samples of a stochastic process and are assessed using the theory of large deviations at the fusion center to arrive at a final decision.

The scattered nature of the sensors will cause their respective communication channels to have different mean path gains, with certain nodes having much better connection than others. The quality of the wireless channel is also affected by the changes in the environment, interference and motion of the sensors. Thus, it is advisable to consider the impact of fading on the performance of wireless sensor networks.

Fading and noisy channels are introduced between the sensors and the fusion center for a binary hypothesis testing problem in [8]. The authors designed the optimum likelihood-ratio based fusion rule for wireless sensor networks with a

fading channel. This rule requires perfect knowledge of the local decision statistics and state of the communication channels. They also proposed the maximum ratio combining statistic and two stage approach using Chair-Varshney fusion rule to alleviate the above requirements. To further robustify the fusion rule, a statistic analogous to equal gain combiner that requires minimum a priori information is proposed [8]. The cost of calculating the instantaneous channel information can be reduced by the use of a fusion rule which requires only channel statistics [30].

The need for spectral bandwidth can be reduced by using a scheme where there is no fusion center and the sensors transmit their observations to their selected neighbors. All the sensors have the same a prior probabilities and update their decision when they make an observation or when they receive information from their neighbors. This process continues until a consensus is reached about the hypothesis among the sensors [7]. This work has been heavily influenced by the work done on distributed estimation [5].

1.4 Energy Constraint

The nodes in wireless sensor networks are powered by batteries for which replacement, if at all possible, is very difficult and expensive. Thus in many scenarios, wireless sensor nodes are expected to operate without battery replacement for long periods or the life of the sensor. Consequently constraining the energy consumption in the nodes is an very important design consideration. The life of the sensor battery can be prolonged by using energy harvesting radios as described in [28].

Research has been done on different aspects of energy efficient wireless sensor networks such as estimation, TCP/IP layer algorithms, modulation, detection etc. In [9], modulation strategy required to send a given number of bits under the energy constraint is analyzed. The total energy of the network includes the transmission

energy and the circuit energy consumption. Thus by optimizing the modulation and the transmission parameters, energy consumption can be minimized. The author states that 80% energy savings is achievable over non-optimized systems for uncoded systems. They also show the variation of the benefits of coding with the transmission distance and underlying modulation schemes.

In [22], quantization of sensor data and energy allocation for the purpose of estimation under energy constraint is considered. Due to bandwidth and energy constraint, the sensor transmits a finite number of bits to the fusion center, where the unknown parameter is estimated. The authors uses the mean reconstruction error as the cost function for optimizing the system parameters which includes the number of levels of quantizer and energy fractions at the nodes. Estimation in a wireless sensor network with correlated sensor nodes under energy constraint is considered in [19].

Energy efficient algorithms in each layer of wireless sensor networks are designed in [11]. Medium access control (MAC) and routing algorithms under energy constraint are considered in [16] and [24], respectively. The detection process with a constraint on the expected cost arising from transmission and measurements is introduced in [29] and [2]. The sensor node consists of four units: sensing unit, microprocessor, a communication unit and a power supply. The work proposes that energy can be saved by switching some of the components off periodically. They propose three modes of operation: active, mute and sleeping. The sensor node is active when all of its units are powered up. It can be in mute with its communication unit off when it plays no role in detection process. The sensor node is sleeping when all its units are switched off. In [29], the authors suggest that energy can be saved if nodes communicate with the rest of the network only when necessary where as in [2], the author proposes that significant energy can be saved

by switching off the node completely whenever the information content of its next observation is likely to be small.

In this thesis we consider the problem of binary hypothesis testing using wireless sensor networks under energy constraint. Traditionally, the decentralized detection problem has been investigated assuming identical sensor nodes. For example the work reported in [36], considers identically distributed observations for all the sensor nodes and error-free transmissions from the nodes to the fusion center. In this thesis we do not assume identically distributed observations. In particular the observation noise experienced by each sensor may be different. Furthermore, the wireless channels between the sensor nodes and the fusion center are assumed to be a noisy channels. Our goal is to design a fusion rule and an energy allocation for the nodes so as to minimize a cost function subject to a limit on the total energy of all the nodes. We consider two types of cost functions. The probability of error at the fusion center as well as the divergence distance measure.

The remainder of this thesis is organized as follows. The problem of energy allocation for the probability of error and the J-divergence cost functions is studied in Chapter 2 and 3, respectively. The results are presented in Chapter 4 and the conclusions are drawn in Chapter 5.

Chapter 2
Error Probability

2.1 Introduction

Consider a binary hypothesis testing problem where a decision is made between two hypothesis, H_1 and H_0 with prior probabilities $P(H = H_1) = p_1$ and $P(H = H_0) = p_0$. Each sensor makes an observation, quantizes it and transmits it to the fusion center where a decision is made between H_0 and H_1 based on the received data from all the sensors. Traditionally an ideal channel has been assumed between the sensors and the fusion center. In this thesis a noisy channel is assumed between each sensor and the fusion center.

The goal is to design a fusion rule and an energy allocation for the nodes subject to a limit on the total energy of all the nodes so as to optimize a cost function. The cost function is defined in terms of the performance of the fusion rule. Two of the mostly used cost functions are the distance measures [26] and the probability of error [36]. Probability of error is the most natural criteria used for the decision making process. However, probability of error calculation requires the knowledge of the prior probabilities. In cases where obtaining the prior probabilities is difficult, Neyman-Pearson criteria is used. In Neyman-Pearson criteria the probability of detection is maximized subject to a limit on the probability of false alarm [34]. In this chapter, we will use the error probability criteria to optimize the system parameters. We denote by E_T the total energy of the L sensors. The fraction of energy allocated to sensor i is denoted by θ_i. This implies that the energy of sensor i for transmission of its observation is given by $\theta_i E_T$ and $\sum_{i=1}^{L} \theta_i = 1$.

2.2 System Model

Consider a network of L wireless sensors with sensor i acquiring a measurement X_i about the observed phenomenon. Each observation consists of two signals s_0 and s_1, where $s_1 = -s_0 = d$, corrupted by additive Gaussian noise with mean zero and variance ${\sigma_i}^2$. Although Gaussian noise is assumed here, the results can be extended to other cases. With this assumption we have

$$\begin{aligned} p_{X_i}(x|H_0) &\sim \mathcal{N}(-d, \sigma_i^2) \\ p_{X_i}(x|H_1) &\sim \mathcal{N}(d, \sigma_i^2) \end{aligned} \tag{2.1}$$

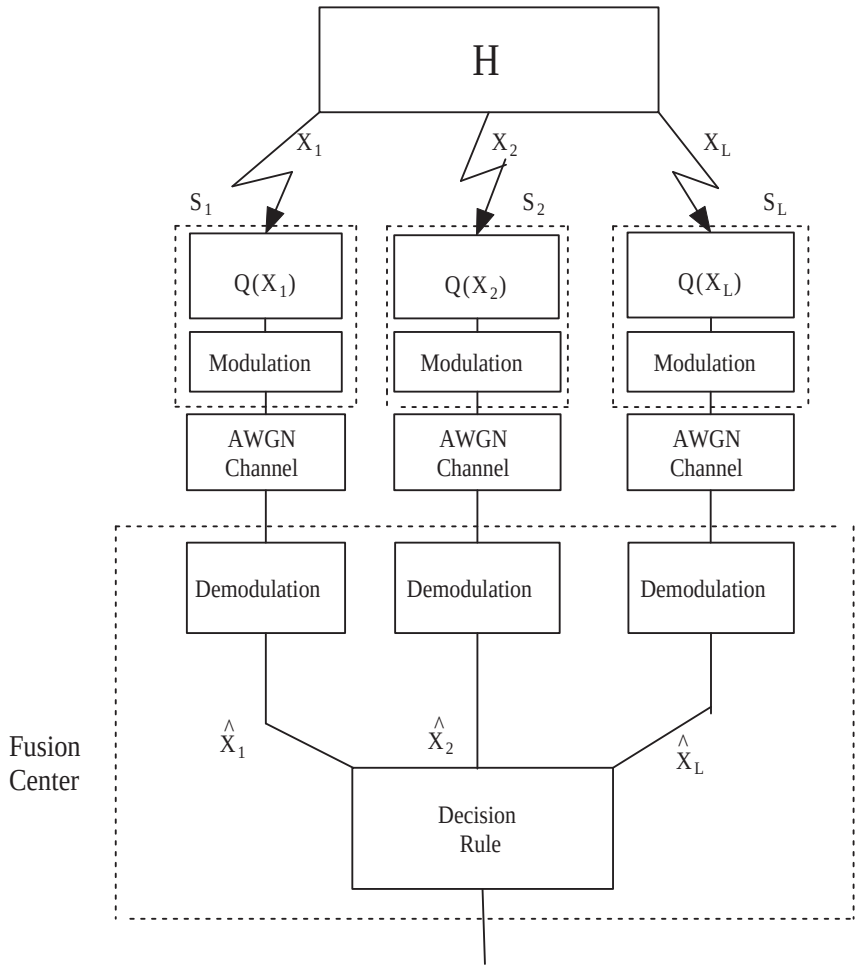

FIGURE 2.1. System Block diagram

Each sensor quantizes its real-valued measurement before transmission to the fusion center in order to reduce the communication bandwidth requirements. The quantizer for sensor i uses output values $\{q_{ij}\}_{j=1}^{M}$ and the thresholds $\{t_{ij}\}_{j=0}^{M}$, where $t_{i0} = -\infty$ and $t_{iM} = \infty$ for all i. Each quantizer output is subsequently encoded into bits $b_{i1}, b_{i2}, \cdots, b_{iN}$, where $N = \log_2 M$.

The channel between the sensors and the fusion center is assumed to be an additive white Gaussian noise channel (AWGN). A modulation scheme is assumed for transmission of the quantizer output bits across the AWGN channel. The system block diagram is shown in Figure 2.1. We denote the output of the ith demodulator by $(z_{i1}, z_{i2}, \cdots, z_{iN})$. These bits are used to construct an estimate of the observation signal X_i, which we denote by $\hat{X}_i$. The fusion center uses the vector $\hat{\mathbf{X}} = (\hat{X}_1, \hat{X}_2, \cdots, \hat{X}_L)$ to make a decision regrading the hypothesis H.

2.3 Bitwise Energy Allocation

In this section the quantized bits $b_{i1}, b_{i2}, \cdots, b_{iN}$ are modulated using a binary modulation scheme such as BPSK or BFSK. Consequently the channel between sensor i and the fusion center can be modeled by a binary symmetric channel. Let β_{ij} denote the fraction of energy of sensor i used to transmit the jth bit. Then cross over probability ϵ_{ij} is given in terms of channel noise power spectral density, $\frac{N_0}{2}$, and the bit energy E_b. For example for BPSK modulation $\epsilon_{ij} = Q(\sqrt{\frac{2E_b}{N_0}}) = Q(\sqrt{\frac{2\beta_{ij}\theta_i E_T}{N_0}})$ and for BFSK modulation $\epsilon_{ij} = Q(\sqrt{\frac{E_b}{N_0}}) = Q(\sqrt{\frac{\beta_{ij}\theta_i E_T}{N_0}})$. The demodulated bits

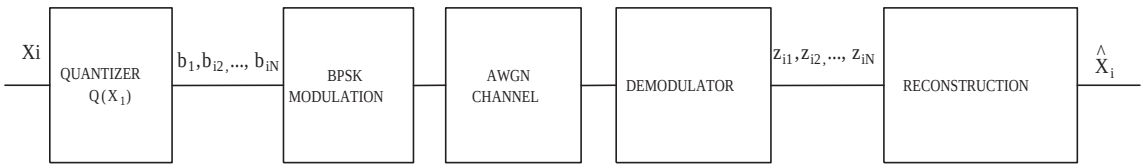

FIGURE 2.2. System Block Diagram for Bitwise allocation

$z_{i1}, z_{i2}, \cdots, z_{iN}$ are used to reconstruct the observation signal $\hat{X}_i$. The fusion center

receives the reconstructed signal sequence $\hat{X_1}, \hat{X_2}, \cdots, \hat{X_L}$ and must decide on the value of H.

Finding the optimum decision rule for H based on $\hat{\mathbf{X}}$ is mathematically intractable. The optimal decision rule requires the conditional distribution of $\hat{\mathbf{X}}$ given the hypothesis H. This is difficult to calculate because of the complexity introduced by the quantizer operation as well as the effects of the channel. Suppose the fusion rule has access to the exact values of the sequence $(X_1X_2, \cdots, X_L)$. This ignores the quantizer noise and the channel errors. It is well known that the optimum decision rule based on the observation $(X_1, X_2, \cdots, X_L)$ is given by

$$\xi(x_1, x_2, \cdots, x_L) = \begin{cases} H_1, & \sum_{i=1}^{L} \alpha_i x_i \geq \tau \\ H_0, & \sum_{i=1}^{L} \alpha_i x_i < \tau \end{cases} \tag{2.2}$$

where $\alpha_i = \frac{1}{\sigma_i^2}$, $i = 1, 2, \cdots, L$ and τ is the threshold used at the fusion center.

With this in mind, and assuming that the quantizer noise is small and the channel error effects are insignificant we opt for the following fusion rule for the received sequence $\hat{\mathbf{X}} = (\hat{X_1}, \hat{X_2}, \cdots, \hat{X_L})$. Let

$$\xi(\hat{x_1}, \hat{x_2}, \cdots, \hat{x_L}) = \begin{cases} H_1, & \sum_{i=1}^{L} \alpha_i \hat{x_i} \geq \tau \\ H_0, & \sum_{i=1}^{L} \alpha_i \hat{x_i} < \tau \end{cases} \tag{2.3}$$

Evaluation of the performance of this rule requires the distribution of $Y = \sum \alpha_i \hat{X_i}$. Assuming a large number of sensors ($L \rightarrow \infty$), we use an asymptotic conditional distribution of Y given the value of H. For this we invoke a form of central limit theorem which is derived from Lyapanov's theorem on the limiting distribution of the sum of non-identically distributed random variables [10]. Specifically, it is shown in [10] that the distribution of Y converges to that of a Gaussian random variable provided that

$$E(\hat{X}_i) \leq \infty,\ var(\hat{X}_i) \leq \infty,$$

and the ratio of the coefficients α_i remains bounded , i.e,

$$|\frac{\alpha_i}{\alpha_j}| \leq\ B\ < \infty,\ \forall\ i, j.$$

In view of this, to compute the conditional distribution of Y given H we only need to compute the conditional mean and variance of Y given H (since the conditional distribution is Gaussian). The conditional moments of Y are then evaluated as follows.

$$E(Y|H_\ell) = \sum \alpha_i E(\hat{X}_i|H_\ell), \quad \text{var}(Y|H_\ell) = \sum {\alpha_i}^2\ \text{var}(\hat{X}_i|H_\ell), \quad \ell = 0, 1 \quad (2.4)$$

The reconstructed signal $(\hat{X}_i)$ is a discrete random variable which takes the quantizer output values $(q_{i1}, q_{i2}, \cdots, q_{iM})$ based on the values of the received bits $z_{i1}, z_{i2}, \cdots, z_{iN}$. The conditional moment of the reconstructed signal $\hat{X}_i$ is given by:

$$E(\hat{X}_i|H_0) = \sum_{j=1}^{M} q_{ij} P(\hat{X}_i = q_{ij}|H_0) \quad (2.5)$$

Let $E(\hat{X}_i|H_0) = \omega_{i0}$. Then

$$\text{var}(\hat{X}_i|H_0) = \sum_{j=1}^{M} {q_{ij}}^2 P(\hat{X}_i = q_{ij}|H_0) - {\omega_{i0}}^2 \quad (2.6)$$

Similarly for hypothesis H_1, we have

$$E(\hat{X}_i|H_1) = \sum_{j=1}^{M} q_{ij} P(\hat{X}_i = q_{ij}|H_1) \quad (2.7)$$

Let $E(\hat{X}_i|H_1) = \omega_{i1}$. Then

$$\text{var}(\hat{X}_i|H_1) = \sum_{j=1}^{M} {q_{ij}}^2 P(\hat{X}_i = q_{ij}|H_1) - {\omega_{i1}}^2 \quad (2.8)$$

For $\ell = 0, 1$, let $\text{var}(\hat{X}_i|H_\ell) = {\gamma_{i\ell}}^2$.

In order to calculate the conditional moments we need to determine the probability distribution of reconstructed signal $\hat{X}_i$. Suppose the quantizer output value q_{ij} is mapped to the binary vector $z_{i1}, z_{i2}, \cdots, z_{iN}$. Then having received $(z_{i1}, z_{i2}, \cdots, z_{iN})$, the reconstruction of $\hat{X}_i$ produces $\hat{X}_i = q_{ij}$.

The probability distribution function of $\hat{X}_i$ is given as follows for $\ell = 0, 1$.

$$\begin{aligned} P(\hat{X}_i = q_{ij})|H_\ell) &= P(z_{i1}, z_{i2}, \cdots, z_{iN}|H_\ell) \\ &= \sum_{k_1=0}^{1} \cdots \sum_{k_N=0}^{1} P(z_{i1}, \cdots, z_{iN}|b_{i1} = k_1, \cdots, b_{iN} = k_N, H_\ell) \\ &\quad P(b_{i1} = k_1, \cdots, b_{iN} = k_N|H_\ell) \\ &= \sum_{k_1=0}^{1}, .., \sum_{k_N=0}^{1} \prod_{j=1}^{N} P(z_{ij}|b_{ij} = k_j) \\ &\quad P(b_{i1} = k_1, \cdots, b_{iN} = k_N|H_\ell) \end{aligned} \tag{2.9}$$

The transition probability $P(z_{ij} = 1|b_{ij} = 0)$ is given by ϵ_{ij}. The probability $P(b_{i1} = k_1, \cdots, b_{iN} = k_N|H_\ell)$ is determined by the distribution of X_i and the quantizer thresholds. For example, if $b_{i1} = k_1, b_{i2} = k_2, \cdots, b_{iN} = k_N$ represents the quantizer output q_{ij}, then

$$P(b_{i1} = k_1, \cdots, b_{iN} = k_N|H_\ell) = P(X_i = q_{ij}|H_\ell) = P(t_{ij} \leq X_i \leq t_{ij+1}|H_\ell) \tag{2.10}$$

From (2.9) and (2.10), we can obtain the probability mass function of the reconstructed variable $\hat{X}_i$. Thus, the conditional moments of the signal Y can be determined.

In Binary hypothesis testing problem, if a decision is taken in favor of H_1 when H_0 is true, the error is called false alarm and the associated conditional probability of error is the probability of false alarm, denoted as P_F. A miss is said to occur, if a decision is in favor of H_0 when H_1 is true. The corresponding conditional probability of error is the probability of miss denoted by P_M. Probability of detection

(P_D) is the conditional probability that the decision is made in favor of H_1 when H_1 is true.

Based on the fusion rule in (2.3), the probability of false alarm is given by

$$P_F = P(Y \geq \tau | H_0), \tag{2.11}$$

Since the distribution of $Y = \sum \alpha_i \hat{X}_i$ is approximated by a Gaussian distribution with conditional moments given by (2.4),

$$P_F = Q\left(\frac{\tau - \sum_{i=1}^{L} \alpha_i \omega_{i0}}{\sqrt{\sum_{i=1}^{L} {\alpha_i}^2 {\gamma_{i0}}^2}}\right) \tag{2.12}$$

and the probability of detection is given by

$$\begin{aligned} P_D &= P(Y \geq \tau | H_1) \\ &= Q(\frac{\tau - \sum_{i=1}^{L} \alpha_i \omega_{i1}}{\sqrt{\sum_{i=1}^{L} {\alpha_i}^2 {\gamma_{i1}}^2}}) \end{aligned} \tag{2.13}$$

Thus the probability of miss is given by:

$$\begin{aligned} P_M &= P(Y \leq \tau | H_1) \\ &= 1 - P_D \end{aligned} \tag{2.14}$$

Finally the probability of error is given as

$$\begin{aligned} P_e &= p(H = H_0) P(Y \geq \tau | H_0) + P(H = H_1) P(Y \leq \tau | H_1) \\ &= p_0 P_F + p_1 P_M \\ &= p_0 P_F + p_1 (1 - P_D) \\ &= p_0 Q\left(\frac{\tau - \sum_{i=1}^{L} \alpha_i \omega_{i0}}{\sqrt{\sum_{i=1}^{L} {\alpha_i}^2 {\gamma_{i0}}^2}}\right) + p_1 [1 - Q\left(\frac{\tau - \sum_{i=1}^{L} \alpha_i \omega_{i1}}{\sqrt{\sum_{i=1}^{L} {\alpha_i}^2 {\gamma_{i1}}^2}}\right) \end{aligned} \tag{2.15}$$

Let $\underline{\theta} = (\theta_1, \theta_2, \cdots, \theta_L)$, $\underline{\beta_{\mathbf{i}}} = (\beta_{i1}, \beta_{i2}, \cdots, \beta_{iN})$. Also let $\mathbf{y} = (\alpha_1, \alpha_2, \cdots, \alpha_L)$. Our goal is to minimize the probability of error by designing the fusion rule and

allocating energies to the each transmitted bit of each sensor subject to a limit on the total energy of the sensor network.The optimization problem can now be stated as follows.

$$\begin{aligned} \text{Minimize} \quad & P_e(\tau, \underline{\theta}, \{\underline{\beta_{\mathbf{i}}}\}_{i=1}^{L}, \mathbf{y}, t_{ij}, q_{ij}) \\ \text{Subject to} \quad & \textstyle\sum_{i=1}^{L} \theta_i = 1 \\ & \textstyle\sum_{j=1}^{N} \beta_{ij} = 1, \ i = 1, 2, \cdots, N \\ & \theta_i \geq 0 \ \& \ \beta_{ij} \geq 0, \ i = 1, 2, \cdots, L. \ j = 1, 2, \cdots, N \end{aligned} \tag{2.16}$$

where we have written $P_e(\tau, \underline{\theta}, \{\underline{\beta_{\mathbf{i}}}\}_{i=1}^{L}, \mathbf{y}, t_{ij}, q_{ij})$ for the probability of error P_e in order to show its dependence on the set of parameters.

This is a non-linear programming problem that can be solved using the method of Lagrange multipliers. Lagrange multiplier converts a constrained problem of n variables with k constraints into an unconstrained problem of $n+k$ variables. The method introduces a new parameter called lagrange multiplier for each constraint in order to convert the problem into an unconstrained problem [6].

For our problem a Lagrangian is formulated as follows:

$$L(\tau, \underline{\theta}, \{\underline{\beta_{\mathbf{i}}}\}_{i=1}^{L}, \mathbf{y}, t_{ij}, q_{ij}, \{\kappa_i\}, \mu, \{\psi_{ij}\}, \chi_i) =$$

$$P_e(\tau, \underline{\theta}, \{\underline{\beta_{\mathbf{i}}}\}_{i=1}^{L}, \mathbf{y}, t_{ij}, q_{ij}) + \sum_{i=1}^{L} \kappa_i \theta_i + \mu\left(\sum_{i=1}^{L} \theta_i - 1\right) + \sum_{i=1}^{L} \chi_i \sum_{j=1}^{N} (\beta_{ij} - 1) + \sum_{i=1}^{L} \sum_{j=1}^{N} \psi_{ij} \beta_{ij} \tag{2.17}$$

The optimal solution of the Lagragian is obtained by applying the Karush-kuhn-Tucker (KKT) conditions.The Karush-kuhn-Tucker (KKT) conditions [18], [20] is a generalization of the method of Lagrange multipliers. It gives the necessary conditions for the solution to be optimal. The Karush-kuhn-Tucker (KKT) [18], [20] conditions dictate that there must exist $\gamma \geq 0$, $\{\kappa_i \geq 0\}_{i=1}^{L}$, $\{\psi_{ij} \geq 0, \ \forall i, j\}$,

$\{\chi_i\}_{i=1}^{L}$ and μ such that

$$\theta_i \geq 0,\ \kappa_i\theta_i = 0,\ i = 1, 2, .., L,\ \psi_{ij}\beta_{ij} \geq 0,\ j = 1, 2, .., N.$$

$$\sum_{i=1}^{L} \theta_i = 1$$

$$\sum_{j=1}^{N} \beta_{ij} = 1,\ \forall i$$

$$\gamma \nabla P_e(\tau, \underline{\theta}, \{\underline{\beta_{\mathbf{i}}}\}_{i=1}^{L}, \mathbf{y}, t_{ij}, q_{ij}) + \nabla \sum_{i=1}^{L} \kappa_i\theta_i$$
$$+\nabla\mu(\sum_{i=1}^{L} \theta_i - 1) + \nabla \sum_{i=1}^{L}\sum_{j=1}^{N} \psi_{ij}\beta_{ij} + \nabla \sum_{i=1}^{L} \chi_i(\sum_{j=1}^{N} \beta_{ij} - 1) = 0 \quad (2.18)$$

Where ∇ denotes gradient.

According to the necessary condition of KKT ([20], [18]), there exists a local minimum at point (x) if the objective function and the constraints are continuously differentiable at a point x. The sufficient condition of KKT state that there is a feasible global optimum satisfying the above equation (2.18) if the objective function and the non-equality constraints are convex and the equality constraints are affine. By solving the constrained problem (2.16), we can obtain the optimal energy allocations $(\underline{\theta}, \underline{\beta_{\mathbf{i}}})$ and the fusion rule $(\mathbf{y}, \tau)$. To get a better understanding of the above problem, we present the results for $N = 3$ and $N = 1$ in Chapter 4.

2.4 Multiple Bits with Equal Energy Allocation

In this section we use M-ary modulation to transmit the quantized observation across the channel. The quantizer outputs $q_{i1}, q_{i2}, \cdots, q_{iM}$ are mapped to symbols $l_1, l_2, \cdots, l_M$ respectively. The transmission symbol u_i takes the value of one of the mapped symbols depending on the distribution of the observation signal X_i. The value of u_i is transmitted across the AWGN channel and z_i denotes the demodulated symbol. The demodulated symbol z_i is used to get an estimate of the

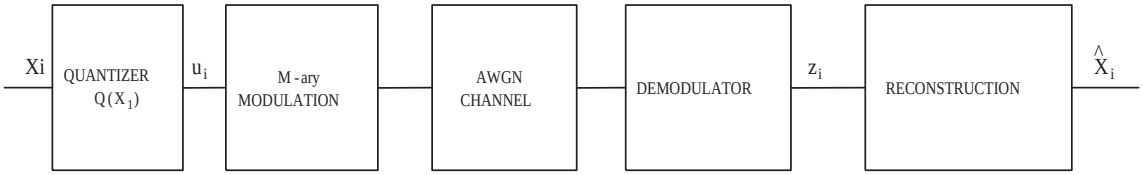

FIGURE 2.3. System Block diagram for M-ary Modulation

observation signal X_i, which is denoted by $\hat{X}_i$. The system model is given by Figure 2.3.

The fusion center receives the reconstructed signal sequence $\hat{\mathbf{X}} = (\hat{X}_1, \hat{X}_2, \cdots, \hat{X}_L)$ and must decide on the value of H. In order to make the fusion rule optimal and tractable we opt for the fusion rule given by (2.3). Evaluation of this rule requires the distribution of $Y = \sum \alpha_i \hat{X}_i$. Assuming a large number of sensors ($L \to \infty$), we invoke the central limit theorem [10] on the conditional distribution of Y given the value of H.

The conditional moments of Y are given by (2.4). In order to obtain the conditional moments of Y, we need to determine the conditional moments of $\hat{X}_i$ which are given by Equations (2.5, 2.6, 2.7 and 2.8). The reconstructed signal $\hat{X}_i$ is a discrete random variable which takes the quantizer output values $q_{i1}, q_{i2}, \cdots, q_{iM}$ depending on the received symbol z_i.

Suppose the quantizer output value q_{ij} is mapped to the symbol l_j. Then having received $z_i = l_j$, the reconstruction of $\hat{X}_i$ produces an estimate $\hat{X}_i = q_{ij}$.

The conditional distribution of $\hat{X}_i$ is given by:

$$\begin{aligned} P(\hat{X}_i = q_{ij}|H_\ell) &= P(z_i = l_j|H_\ell) \\ &\quad \sum_{k=1}^{M} P(z_i = l_j|u_i = l_k, H_\ell) P(u_i = l_k|H_\ell) \\ &\quad \sum_{k=1}^{M} P(z_i = l_j|u_i = l_k) P(u_i = l_k|H_\ell) \end{aligned} \tag{2.19}$$

where $P(z_i = l_j|u_i = l_k)$ is the transition probability that the symbol l_j is received when the symbol l_k is transmitted across the channel.

The transition probability is determined as follows.

Let $s_j(t)$ be the modulated signal for the jth symbol (l_j) and $r(t)$ denote the output of the channel. Lets denote the kth orthonormal basis function by $\delta_k(t)$.

For an AWGN channel, the conditional distribution of the received signal is given in [14] as follows

$$P(r|u_i = l_j) = (\pi N_0)^{\frac{N}{2}} \exp[-\frac{1}{N_0}\sum_{k=1}^{N}(r_k - s_{jk})^2] \tag{2.20}$$

where $s_{jk} = \int_{t=0}^{T} s_j(t)\delta_k(t)dt$ and $r_k = \int_{t=0}^{T} r(t)\delta_k(t)dt$. Thus the transition probabilities for the discrete memoryless channel are given by:

$$P(z_i = l_m|l_j) = \int_{R_m} (\pi N_0)^{\frac{N}{2}} \exp[-\frac{1}{N_0}\sum_{k=1}^{N}(r_k - s_{jk})^2]dr_k. \tag{2.21}$$

where R_m is the decision region used by the demodulator for the symbol l_m.

The transition probabilities $P(z_i|u_i)$ depend on the energy $E_T\theta_i$ required to transmit the symbol across the channel. The probability $P(u_i = l_j|H_\ell)$ is determined by the distribution of X_i and the quantizer thresholds. For example, if the symbol l_j represents the quantizer output q_{ij}, then

$$P(u_i = l_j|H_\ell) = P(Q(X_i) = q_{ij}|H_\ell) = P(t_{ij} \leq X_i \leq t_{ij+1}|H_\ell) \tag{2.22}$$

From (2.21) and (2.22) the conditional distribution of the reconstructed signal $\hat{X}_i$ is calculated. Thus, the conditional moments of $\hat{\mathbf{X}}$ can be determined (4).

The probability of false alarm is given by (2.12) and the probability of detection by (2.13). Finally the probability of error is given by:

$$P_e = p_0 P_F + p_1 P_D \tag{2.23}$$

The optimization problem can now be stated as follows.

$$\begin{aligned} \text{minimize} \quad & P_e(\tau, \underline{\theta}, \mathbf{y}, t_{ij}, q_{ij}) \\ \text{Subject to} \quad & \textstyle\sum_{i=1}^{L} \theta_i = 1 \\ & \theta_i \geq 0 \quad i = 1, 2, \cdots, L \end{aligned} \tag{2.24}$$

As discussed in the previous section, the optimization problem can be solved using Lagrange multipliers. For this a Lagragian is formulated as follows

$$L(\tau, \underline{\theta}, \mathbf{y}, t_{ij}, q_{ij}, \{\kappa_i\}, \mu) =$$

$$P_e(\tau, \underline{\theta}, \mathbf{y}, t_{ij}, q_{ij}) + \sum_{i=1}^{L} \kappa_i \theta_i + \mu(\sum_{i=1}^{L} \theta_i - 1) \tag{2.25}$$

The optimal solution of the Lagragian is obtained by applying the Karush-kuhn-Tucker (KKT) conditions. The Karush-kuhn-Tucker (KKT) condition [18], [20] is a generalization of the method of Lagrange multipliers. The Karush-kuhn-Tucker (KKT) [18], [20] conditions dictate that there must exist $\gamma \geq 0$, $\{\kappa_i \geq 0\}_{i=1}^{L}$, and μ such that

$$\theta_i \geq 0, \ \kappa_i \theta_i = 0, \ i = 1, 2, .., L.$$

$$\sum_{i=1}^{L} \theta_i = 1$$

$$\gamma \nabla P_e(\tau, \underline{\theta}, \mathbf{y}, t_{ij}, q_{ij}) + \nabla \sum_{i=1}^{L} \kappa_i \theta_i + \nabla \mu(\sum_{i=1}^{L} \theta_i - 1) = 0 \tag{2.26}$$

where ∇ denotes gradient.

The optimal energy allocations ($\underline{\theta}$) and the fusion rule ($\mathbf{y}, \tau$) can be obtained by solving the lagrange multiplier problem (2.26) and (2.26).

2.4.1 Two Bit Case with QPSK Modulation

To get a better understanding of the problem, we solved the above problem using a four-level quantizer. The observations made at the senors are quantized into finite output values $q_{i1}, q_{i2}, q_{i3}, q_{i4}$ and subsequently mapped to symbols l_1, l_2, l_3, l_4

respectively. Let u_i be the symbol transmitted by the sensor i across the channel and received as z_i. The channel is modeled by a discrete memoryless channel. The received symbol z_i is processed to get an estimate of the observation signal X_i, denoted by $\hat{X}_i$. The fusion center receives the reconstructed sequence $\hat{\mathbf{X}} = (\hat{X}_1, \hat{X}_2, \cdots, \hat{X}_L)$ and makes a decision on the value of H using the fusion rule (2.3).

In order to evaluate the performance of the this rule, we need the distribution of $Y = \sum \alpha_i \hat{X}_i$. The conditional distribution of Y is given by (2.4). To determine the conditional distribution of the signal Y we need to determine conditional distribution of $\hat{X}_i$. The conditional distribution of $\hat{X}_i$ is given by (2.19). This requires the knowledge of the transition probabilities for the channel.

The transition probabilities of the channel are determined as follows:

Let $s_j(t)$ be the modulated signal for the jth symbol (l_j) and $r(t)$ denote the output of the channel. Lets denote the kth orthonormal basis function by $\delta_k(t)$.

The conditional distribution of the received signal is given by [14]

$$P(r|u_i = l_j) = \pi N_0 \exp[-\frac{1}{N_0} \sum_{k=1}^{2} (r_k - s_{jk})^2] \tag{2.27}$$

where $s_{jk} = \int_{t=0}^{T} s_j(t)\delta_k(t)dt$ and $r_k = \int_{t=0}^{T} r(t)\delta_k(t)dt$. For a QPSK modulation, they are given by $(\{s_{j1}, s_{j2}\} = \{(\pm\sqrt{\frac{E_T\theta_i}{2}}, \pm\sqrt{\frac{E_T\theta_i}{2}}), \forall j)$ for symbols l_1, l_2, l_3, l_4 respectively.

The transition probabilities of the channel are given by

$$P(z_i = l_m|l_j) = \int_{R_m} P(r|u_i = l_j)dr_k. \tag{2.28}$$

The decision regions for the four symbols l_1, l_2, l_3, l_4 are given by $\{(r_1 \geq 0, r_2 \geq 0), (r_1 \leq 0, r_2 \geq 0), (r_1 \geq 0, r_2 \leq 0), (r_1 \leq 0, r_2 \leq 0)\}$ respectively in a two dimensional plane. Thus we can determine the transition probabilities (2.28). The

conditional distribution of Y can now be determined (4). Applying the conditional moments of Y, we finally arrive at the probability of error (2.15).

The optimization problem is now given by (2.24). The optimization problem can be solved using Lagragian multipliers. For this a Lagragian is formulated by (2.25). The Karush-kuhn-Tucker (KKT) [18], [20] conditions dictate that there must exist $\gamma \geq 0$, $\{\kappa_i \geq 0\}_{i=1}^{L}$, and μ such that

$$\theta_i \geq 0,\ \kappa_i \theta_i = 0,\ i = 1, 2, .., L.$$

$$\sum_{i=1}^{L} \theta_i = 1$$

$$\gamma \nabla P_e(\tau, \underline{\theta}, \mathbf{y}, t_{ij}, q_{ij}) + \nabla \sum_{i=1}^{L} \kappa_i \theta_i + \nabla \mu (\sum_{i=1}^{L} \theta_i - 1) = 0 \tag{2.29}$$

where ∇ denotes gradient. The optimal energy allocations ($\underline{\theta}$) and the fusion rule ($\mathbf{y}, \tau$) can be obtained by solving the constrained problem.

Chapter 3
Distance Measure

3.1 Introduction

In general we would like to perform the energy allocation using the error probability as the cost function. However, as noted in the previous chapter obtaining the optimal detection rule may be intractable. In addition, it is difficult to obtain the closed form expression for the error probability if we are using an M-level quantizer to quantize the sensor observations. To overcome this problem, we use the Ali-Silvey class of distance measures [1]. This class of Ali-silvey distance measures is defined between the probability distributions $P_0 = P(x|H_0)$ and $P_1 = P(x|H_1)$ and is written as follows:

$$d(P_0, P_1) \;=\; f\{E_0\{C(L)\}\} \tag{3.1}$$

where $d(P_0, P_1)$ is the distance between the probability distributions, f is an increasing function, C is a convex function and L is the likelihood ratio $\frac{dP_1}{dP_0}$ and E_0 denotes expectation w.r.t P_0. The Three examples of the Ali-Silvey distance measures are given in the following:

- J-Divergence [21] : $E_0\{(L-1)\ \log L\}$. This expression fulfills all the requirements of the Ali-silvey distance measures. It can be expanded as follows to make it easier for analysis.

$$\begin{aligned} J-Divergence \;&=\; E_0\{(L-1)\ \log L\} \\ &= \int \frac{dP_1 - dP_0}{dP_0}\ \log L\ dP_0 \\ &= \int \log L\ dP_1 - \log L\ dP_0 \\ &= E_1(\log L) - E_0(\log L) \end{aligned} \tag{3.2}$$

This form of J-divergence (3.2) is widely used in the detection problems [26].

- Matsusita's Distance [23] : $d = [E_0\{(\sqrt{L}-1)^2\}]^2$. In this case the convex function $C(L)$ is given by $(\sqrt{(}L)-1))^2$ and the increasing function $f(d)$ by $\sqrt{d}$. Both the functions satisfy the requirements of Ali-Silvey distance measures (3.1).

- Bhattacharya distance : $B = -\log(1-d^2)$ (d is Matsusita's distance). The bhattacharya distance has the same convex function as the matsusita distance while the increasing function $f(d)$ is given by $\log(1-d)$.

These distance measures are frequently used in detection problems because of the theorems relating the maximum distance to the minimum probability of error [26]. Lets apply the theorem to the classic binary hypothesis testing problem where a decision is made at the fusion between two hypothesis, H_1 and H_0. The fusion center makes a decision based on the decision vector $\hat{\mathbf{X}}$. A solution to the problem is obtained by classifying the sample space into two complementary regions R_1 and R_2 and allocating to H_ℓ if $\hat{X} \in R_\ell$.

Let P_1 be the distribution of $\hat{X}$ under hypothesis H_1 and P_0 under hypothesis H_0.

The probability of error for this case is given by:

$$p_e = \pi_1 \int_{R_1} P_1 d\hat{x} + \pi_0 \int_{R_0} P_0 d\hat{x} \tag{3.3}$$

where π_1 and π_0 are the prior probabilities of the hypothesis,H_1 and H_0 respectively.

If R_1 is chosen to minimize p_e, we have [1]

$$R_1 = \{\hat{x} : P_1/P_0 < \pi_1/\pi_0\},$$

and consequently

$$R_0 = \{\hat{x} : P_1/P_0 \geq \pi_1/\pi_0\},$$

In [1], it is shown that

$$\begin{aligned} 1 - p_e &= \pi_0 + \pi_1 \int_{L > \frac{\pi_1}{\pi_0}} [L - \frac{\pi_1}{\pi_0}] d\hat{x} \\ &= \pi_1 E_0 |L - \frac{\pi_1}{\pi_0}| \end{aligned} \tag{3.4}$$

Thus we have

$$1 - 2p_e = \pi_1 E_0 |L - \frac{\pi_1}{\pi_0}| \tag{3.5}$$

In the Equation 3.5, the coefficient $1 - 2p_e$ reelects the distance between the probability distributions. Thus this relates the minimum error probability to the maximum distance measures.

This has triggered the application of Ali-Silvey distance measures in signal detection theory. They are applied in [26] to design a generalized quantizer for binary hypothesis testing problem. They have been applied to signal selection and radar technology in [13]. We apply them here to derive a tractable design procedure for the binary hypothesis testing. The design of the application depends on the type of distance measure being used. Therefore, the selection of a suitable distance measure is important. In our thesis, we use the J-divergence distance measure in particular because it gives a tractable solution to the energy allocation problem. Our goal is to design a fusion rule and an energy allocation for the nodes so as to maximize the J-divergence with a limit on the total energy of all the nodes. In the case of J-Divergence, we will desgin a generalized quantizer. Generalized quantizer involves the design of the thresholds only. The quantizer outputs can be obtained from the thresholds.

3.2 Bitwise Energy Allocation

The observations made by the sensors are quantized into bits $(b_{i1}, \cdots, b_{iN})$ before being transmitted to the fusion center in order to reduce the bandwidth requirements. The bits are modulated using a binary modulation scheme such as BPSK and BFSK. Consequently the channel between the sensor i and the fusion center can be modeled by a binary symmetric channel. The received signal is demodulated by the ith demodulator into bits $z_{i1}, z_{i2}, \cdots, z_{iN}$ which are used to estimate the observation signal, which is denoted by $\hat{X}_i$. The reconstructed signal $\hat{X}_i$ takes the quantizer output values $\{q_{i1}, q_{i2}, \cdots, q_{iM}\}$ depending on the received bits. The reconstructed signal sequence $\hat{\mathbf{X}} = (\hat{X}_1, \hat{X}_2, \cdots, \hat{X}_L)$ is used by the fusion center to make a decision on the observed hypothesis H. The optimal fusion rule is a likelihood- ratio test [34] .

$$\xi(\hat{x_1}, \hat{x_2}, \cdots, \hat{x_L}) = \begin{cases} H_1, & \ln \frac{p(\hat{\mathbf{X}}|H_1)}{p(\hat{\mathbf{X}}|H_0)} \geq \tau \\ H_0, & \ln \frac{p(\hat{\mathbf{X}}|H_1)}{p(\hat{\mathbf{X}}|H_0)} < \tau \end{cases} \tag{3.6}$$

We can write

$$T(\hat{\mathbf{X}}) = \ln \frac{p(\hat{\mathbf{X}}|H_1)}{p(\hat{\mathbf{X}}|H_0)} = \sum_{i=1}^{L} \ln \frac{p(\hat{X}_i|H_1)}{p(\hat{X}_i|H_0)} \tag{3.7}$$

Our goal is to to design a fusion rule and an energy allocation for the nodes subject to a limit on the total energy of all the nodes so as to maximize the distance measure. Probability of error is the most natural criteria used for decision making process. In order to calculate the probability of error , we need to determine the probability of false alarm (P_F) and probability of detection (P_D) for the fusion rule. The probability of false alarm is given by

$$P_F = P(T(\hat{\mathbf{X}}) \geq \tau | H_0) \tag{3.8}$$

and the probability of detection is given by

$$P_D = P(T(\hat{\mathbf{X}}) \geq \tau | H_1) \tag{3.9}$$

It is difficult to obtain a closed form for the above equations (3.8)and (3.9). The fusion rule does not give a tractable design solution if we are using error probability criteria as the cost function. Therefore, in this chapter we opt for an alternative cost function, namely the J-divergence distance measure which belongs to the class of Ali-Silvey distance measures [1] between probability measures. The J-divergence distance measure gives us an tractable solution even though it may not be an optimal solution.

Let $\underline{\theta} = (\theta_1, \theta_2, \cdots, \theta_L)$ and $\underline{\beta_{\mathbf{i}}} = (\beta_{i1}, \beta_{i2}, \cdots, \beta_{iN},\ \forall i)$.

The J-divergence distance measure is given by (3.2):

$$\begin{aligned} J(\underline{\theta}, \{\underline{\beta_i}\}_{i=1}^{L}, t_{ij}) &= E_{H_1}\left[T(\hat{\mathbf{X}})\right] - E_{H_0}\left[T(\hat{\mathbf{X}})\right] \\ &= E_{H_1}\left[\ln \frac{p(\hat{\mathbf{X}}|H_1)}{p(\hat{\mathbf{X}}|H_0)}\right] - E_{H_0}\left[\ln \frac{p(\hat{\mathbf{X}}|H_1)}{p(\hat{\mathbf{X}}|H_0)}\right] \\ &= E_{H_1}\left[\sum_{i=1}^{L} \ln \frac{p(\hat{X}_i|H_1)}{p(\hat{X}_i|H_0)}\right] - E_{H_0}\left[\sum_{i=1}^{L} \ln \frac{p(\hat{X}_i|H_1)}{p(M_i|H_0)}\right] \\ &= \sum_{i=1}^{L} E_{H_1}\left[\ln \frac{p(\hat{X}_i|H_1)}{p(\hat{X}_i|H_0)}\right] - E_{H_0}\left[\ln \frac{p(\hat{X}_i|H_1)}{p(\hat{X}_i|H_0)}\right] \end{aligned} \tag{3.10}$$

Thus $J(\underline{\theta}, \{\underline{\beta_i}\}_{i=1}^{L}, t_{ij}) = \sum_{i=1}^{L}\ J(\theta_i, \underline{\beta_i}, \{t_{ij}\}_{j=1}^{M-1}, \{q_{ij}\}_{j=1}^{M})$, where

$$\begin{aligned} J(\theta_i, \underline{\beta_i}, \{t_{ij}\}_{j=1}^{M-1}) &= E_{H_1}[\ln \frac{p(\hat{X}_i|H_1)}{p(\hat{X}_i|H_0)}] - E_{H_0}[\ln \frac{p(\hat{X}_i|H_1)}{p(\hat{X}_i|H_0)} \\ &= \sum_{j=1}^{M} p(\hat{X}_i = q_{ij}|H_1)\left[\ln \frac{p(\hat{X}_i = q_{ij}|H_1)}{p(\hat{X}_i = q_{ij}|H_0)}\right] \\ &\quad -p(\hat{X}_i = q_{ij}|H_0)\left[\ln \frac{p(\hat{X}_i = q_{ij}|H_1)}{p(\hat{X}_i = q_{ij}|H_0)}\right] \end{aligned} \tag{3.11}$$

In order to determine the J-divergence , we need the probability mass function of the reconstructed signal $\hat{X}_i$. This is given by (2.9).

The optimization problem can now be formulated as follows.

$$\begin{aligned} \text{Maximize} \quad & \textstyle\sum_{i=1}^{L} J(\theta_i, \underline{\beta_i}, \{t_{ij}\}_{j=1}^{M-1}) \\ \text{Subject to} \quad & \textstyle\sum_{i=1}^{L} \theta_i = 1 \\ & \textstyle\sum_{j=1}^{N} \beta_{ij} = 1,\ i = 1, 2, \cdots, L \\ & \theta_i \geq 0\ \&\ , \beta_{ij} \geq 0,\ i = 1, 2, \cdots, L.\ j = 1, 2, \cdots, N. \end{aligned} \tag{3.12}$$

For this a lagrangian is formulated as follows:

$$L(\underline{\theta}, \{\underline{\beta_i}\}_{i=1}^{L}, t_{ij}, \{\kappa_i\}, \mu, \{\psi_{ij}\}, \chi_i) =$$
$$-\sum_{i=1}^{L} J(\theta_i, \underline{\beta_i}, \{t_{ij}\}_{j=1}^{M-1}) + \sum_{i=1}^{L} \kappa_i \theta_i + \mu(\sum_{i=1}^{L} \theta_i - 1) + \sum_{i=1}^{L} \chi_i \sum_{j=1}^{N} (\beta_{ij} - 1) + \sum_{i=1}^{L} \sum_{j=1}^{N} \psi_{ij} \beta_{ij} \tag{3.13}$$

The Karush-kuhn-Tucker(KKT) [18], [20] is a generalization of the method of Lagrange multipliers. They state the necessary conditions for the solution to be optimal. The Karush-kuhn-Tucke(KKT) [18], [20] conditions dictate that there must exist $\gamma \geq 0$, $\{\kappa_i \geq 0\}_{i=1}^{L}$, $\{\psi_{ij} \geq 0,\ \forall i, j\}$, χ_i and μ such that

$$\theta_i \geq 0,\ \kappa_i \theta_i = 0,\ i = 1, 2, .., L,\ \psi_{ij} \beta_{ij} \geq 0,\ j = 1, 2, .., N.$$
$$\sum_{i=1}^{L} \theta_i = 1$$
$$\sum_{j=1}^{N} \beta_{ij} = 1,\ \forall i$$
$$\begin{aligned} & -\gamma \nabla \sum_{i=1}^{L} J(\underline{\theta_i}, \underline{\beta_i}, \{t_{ij}\}_{j=1}^{M-1}) + \nabla \sum_{i=1}^{L} \kappa_i \theta_i \\ & + \nabla \mu (\sum_{i=1}^{L} \theta_i - 1) + \nabla \sum_{i=1}^{L} \sum_{j=1}^{N} \psi_{ij} \beta_{ij} + \nabla \sum_{i=1}^{L} \chi_i (\sum_{j=1}^{N} \beta_{ij} - 1) = 0 \end{aligned} \tag{3.14}$$

where ∇ denotes gradient.

The solution of the constrained problem (Eq.3.12) gives the optimum system design for the wireless sensor networks under energy constraint. The results for both $N = 3$ and $N = 2$ are presented in the Chapter 4.

3.3 Multiple Bit with Equal Energy Allocation

In this section the quantizer outputs q_{ij} are modulated using M-ary modulation before transmission across the channel. The quantizer outputs $q_{i1}, q_{i2}, \cdots, q_{iM}$ are mapped to symbols $l_1, l_2, \cdots, l_M$ respectively. The transmitted symbol u_i takes the value of one of the mapped symbols depending on the distribution of the observation signal X_i. The channel is modeled as discrete memoryless channel whose transition probabilities are given by (2.21). The value of u_i is transmitted across this channel and z_i denotes the received symbol. The received symbol z_i is used to get an estimate of the observation signal X_i, which is denoted by $\hat{X}_i$. The system model is given by Figure 2.3. The fusion center takes a decision based on the reconstructed signal sequence $\hat{\mathbf{X}} = \{\hat{X}_1, \hat{X}_2, \cdots, \hat{X}_L\}$. The optimal fusion rule is given by (3.6).

It is difficult to obtain a closed form for the error probability for this fusion rule. Therefore, we opt for an alternative cost function, namely the J-divergence distance measure which belongs to the class of Ali-Silvey distance measures. The j-divergence is given as follows:

Let $\underline{\theta} = (\theta_1, \theta_2, \cdots, \theta_L)$

$$J(\underline{\theta}, \{t_{ij}\}_{j=1}^{M-1}) = \sum_{i=1}^{L} E_{H_1}\left[\ln \frac{p(\hat{X}_i|H_1)}{p(\hat{X}_i|H_0)}\right] - E_{H_0}\left[\ln \frac{p(\hat{X}_i|H_1)}{p(\hat{X}_i|H_0)}\right]$$

Thus $J(\underline{\theta}, t_{ij}) = \sum_{i=1}^{L} J(\theta_i, \{t_{ij}\}_{j=1}^{M-1})$, where

$$\begin{aligned} J(\theta_i, t_{ij}) &= E_{H_1}[\ln \frac{p(\hat{X}_i|H_1)}{p(\hat{X}_i|H_0)}] - E_{H_0}[\ln \frac{p(\hat{X}_i|H_1)}{p(\hat{X}_i|H_0)} \\ &= \sum_{j=1}^{M} p(\hat{X}_i = q_{ij}|H_1)[\ln \frac{p(\hat{X}_i = q_{ij}|H_1)}{p(\hat{X}_i = q_{ij}|H_0)}] \\ &\quad -p(\hat{X}_i = q_{ij}|H_0)[\ln \frac{p(\hat{X}_i = q_{ij}|H_1)}{p(\hat{X}_i = q_{ij}|H_0)}] \end{aligned} \tag{3.15}$$

where q_{ij} represents the quantizer output values.

The probability mass function of the reconstructed signal $\hat{X}_i$ is given by (2.19). The optimization problem for the J-divergence is formulated as follows.

$$\text{Maximize} \qquad \sum_{i=1}^{L} j(\theta_i, t_{ij}) \tag{3.16}$$

$$\text{Subject to} \qquad \sum_{i=1}^{L} \theta_i = 1 \tag{3.17}$$

$$\theta_i \geq 0 \tag{3.18}$$

The above non linear programming problem can be solved using the method of Lagrange multipliers. The Lagrangian is formulated as follows:

$$L(\underline{\theta}, t_{ij}\{\kappa_i\}, \mu) = -\sum_{i=1}^{L} J(\theta_i, \{t_{ij}\}_{j=1}^{M-1}) + \sum_{i=1}^{L} \kappa_i \theta_i + \mu(\sum_{i=1}^{L} \theta_i - 1) \tag{3.19}$$

The Karush-kuhn-Tucker(KKT) [18], [20] conditions dictate that there must exist $\gamma \geq 0$, $\{\kappa_i \geq 0\}_{i=1}^{L}$, and μ such that

$$\theta_i \geq 0, \kappa_i \theta_i = 0, i = 1, 2, .., L,$$

$$\sum_{i=1}^{L} \theta_i = 1$$

$$-\gamma \nabla \sum_{i=1}^{L} J(\theta_i, \{t_{ij}\}_{j=1}^{M-1}) + \nabla \sum_{i=1}^{L} \kappa_i \theta_i + \nabla \mu(\sum_{i=1}^{L} \theta_i - 1) = 0 \tag{3.20}$$

where ∇ denotes gradient.

The optimal and tractable system design can be obtained by solving the constrained problem (3.18). To get a better understanding of the problem , we will solve the problem for $N = 2$ where we use QPSK modulation to modulate the quantized bits. The results presented in chapter.4 will corroborate the discussion.

3.3.1 Single Bit Case

In this section , the observations made by the sensors are quantized using a binary quantizer. The output of the bi-level quantizer is considered as the local decision

made by the sensor. For the given observation (2.1), the sensor i computes a local binary decision u_i according to

$$u_i = \begin{cases} 1, & \text{if} \quad \ln(\frac{p_{X_i}(x|H_1)}{p_{X_i}(x|H_0)}) \geq \lambda_i \\ 0, & \text{if} \quad \ln(\frac{p_{X_i}(x|H_1)}{p_{X_i}(x|H_0)}) < \lambda_i \end{cases}$$

For the given distribution (2.1), the optimal value of λ_k is given by

$$\lambda_i = \frac{{\sigma_i}^2(\log_2 q_0 - \log_2 q_1)}{2d} \tag{3.21}$$

The channel between sensor i and the fusion center is modeled by a binary symmetric channel. The value of u_i is transmitted to the fusion center over this channel and z_i denotes the received bit. For the sake of concreteness we assume that the sensors use a BPSK modulation scheme. . The fusion center receives the sequence $\mathbf{Z} = (z_1, z_2, \cdots, z_L)$ and must decide on the state of H. The fusion rule is given by (3.6). In this case obtaining an expression for the error probability that is suitable for energy allocation is difficult. Therefore, in this section we opt for an alternative cost function, namely the J-divergence distance measure.

$$J(\underline{\theta}) = E_{H_1}\left[T(\mathbf{z})\right] - E_{H_0}\left[T(\mathbf{z})\right] \tag{3.22}$$

where $T(\mathbf{z})$ is the log-likelihood ratio function given by $T(Z) = \ln \frac{p(\mathbf{z}|H_1)}{p(\mathbf{z}|H_0)}$ and E_{H_ℓ} is expectation operation under the hypothesis H_ℓ. We can write

$$T(\mathbf{z}) = \ln \frac{p(\mathbf{z}|H_1)}{p(\mathbf{z}|H_0)} = \sum_{i=1}^{L} \ln \frac{p(z_i|H_1)}{p(z_i|H_0)} \tag{3.23}$$

From (3.10), we have $J(\underline{\theta}) = \sum_{i=1}^{L} j(\theta_i)$.

In order to calculate the J-divergence, we need to determine the conditional distribution of z_i. The conditional distribution of z_i is given by

$$\begin{aligned} P(z_i = 1|H_0) &= P(z_i = 1|u_i = 0, H_0)P(u_i = 0|H_0) \\ &\quad + P(z_i = 1|u_i = 1, H_0)P(u_i = 1|H_0) \\ &= \epsilon_i \left[1 - Q\left(\frac{\lambda_i \sigma_i}{2d} + \frac{d}{\sigma_i}\right)\right] + (1 - \epsilon_i)Q\left(\frac{\lambda_i \sigma_i}{2d} + \frac{d}{\sigma_i}\right) \quad (3.24) \end{aligned}$$

where $\epsilon_i = Q(\frac{2E_T\theta_i}{N_0})$.

Similarly for hypothesis H_1, we have

$$\begin{aligned} P(z_i = 1|H_1) &= P(z_i = 1|u_i = 0, H_1)P(u_i = 0|H_1) \\ &\quad + P(z_i = 1|u_i = 1, H_1)P(u_i = 1|H_1) \\ &= \epsilon_i \left[1 - Q\left(\frac{\lambda_i \sigma_i}{2d} - \frac{d}{\sigma_i}\right)\right] + (1 - \epsilon_i)Q\left(\frac{\lambda_i \sigma_i}{2d} - \frac{d}{\sigma_i}\right) \quad (3.25) \end{aligned}$$

Thus we can determine the J-divergence.

The optimization problem for this case is given by (3.18). The lagragian is given by (3.19). The optimal solution for the lagragian can be obtained by using the KKT conditions [18], [20]. The KKT conditions dictate that there must exist $\{\kappa\}_{i=1}^{L}$ and μ such that:

$$\theta_i \geq 0, \kappa_i \geq 0, \kappa\theta_i = 0, i = 1, 2, ..., L$$

$$\sum_{i=1}^{L} \theta_i = 0$$

$$-\nabla(\sum_{i=1}^{L} J(\theta_i) + \nabla(\sum_{i=1}^{L} \kappa_i\theta_i) + \nabla(\mu(\sum_{i=1}^{L} \theta_i - 1)) = 0 \quad (3.26)$$

By solving this problem we can obtain the optimal energy allocation scheme.

Chapter 4
Results

4.1 Introduction

In this thesis, we consider a binary hypothesis testing problem where a decision is made between two hypothesis, H_1 and H_0 with prior probabilities $P(H = H_1) = p_1$ and $P(H = H_0) = p_0$. The decision is made at the fusion center based on the quantized observations transmitted by the sensor. The decision is made using a fusion rule which is designed by optimizing the cost function. In this thesis we used two different cost functions; the error probability criteria and J-divergence distance measure. Apart from designing the fusion rule, we are also interested in the energy allocation for the nodes with a limit on the total energy of all the nodes.

Previously in the Chapter 2 and 3, we presented the fusion rule used by the fusion center for both the error probability and J-divergence distance measure. We derived the optimization problem for both the cost functions. We also presented the analytical formulation required to solve the optimization problem. In this chapter we will apply them to obtain at results which will provide us with an optimal fusion rule and energy allocation for wireless sensor networks with a limit on the total energy for all the nodes.

4.2 M-ary Modulation

In this section, we will present the results for the case where an M-ary modulation is used to modulate the quantizer outputs. The quantizer outputs $q_{i1}, q_{i2}, \cdots, q_{iM}$ are mapped to symbols $l_1, l_2, \cdots, l_M$ before modulation. Since the quantized observations are transmitted as symbols, we will have energy allocation only among the sensors. All the bits representing the symbol are allocated equal energy. The

received symbols at the fusion center are processed to get an estimate of the observation signal X_i, denoted by $\hat{X}_i$. This reconstructed signal $\hat{X}_i$ is used by the fusion center to make a decision on the value of H.

In this thesis, we will be designing the optimum fusion rule and energy allocation with a limit on the total network energy so as to optimize the cost function. The two different cost functions used in this thesis are error probability and the distance measure. In this section we will present the results obtained by solving the analytical formulations presented in Chapter 2 and 3 for the M-ary modulation.

4.2.1 Error Probability Criteria

The optimal fusion rule and the energy allocations obtained by minimizing the error probability are presented in this section. The fusion rule for the error probability criteria is considered to be optimal for $\alpha_i = \frac{1}{{\sigma_i}^2}$ (2.3). To show the efficacy of this prediction rule (2.3), we consider the case where $N = 1$. We also assume error free channel between the sensors and the fusion center. By using $N = 1$, we reduce the number of variables that determine the error probability but see an increase in the quantization noise.

TABLE 4.1. WSN Configuration for L=5.

Sensor index (i)	1	2	3	4	5
${\sigma_i}^2$	1	2	3	4	5

The analytical formulation for this case is presented in Chapter 2. The system parameters to be optimized are the α_i values and τ. The wireless sensor network is assumed to have five sensors with the noise variances given in Table 4.1. We can observe that the optimal value of α_i increases with the decrease in noise variance ${\sigma_i}^2$ (Table 4.2). Therefore the sensor with a high noise variance will have no importance in the decision making process. We plot the error probability for the optimal values of $\{\alpha_i\}$ (given in Table.4.2) as a function of τ in Figure 4.1. For comparison, we

TABLE 4.2. Optimum α values for Different Sensor Index.

Sensor	1	2	3	4	5
α	2.023	0.52	0.33	0.26	0.23

have also plotted the error probability for $\alpha_i = \frac{1}{{\sigma_i}^2}$. It can be seen that both cases result in similarly small error probabilities albeit for different values of τ. This indicates that if optimization over τ is performed then $\alpha_i = \frac{1}{{\sigma_i}^2}$ results in good performance. Since we obtained the optimal fusion rule for the error probability

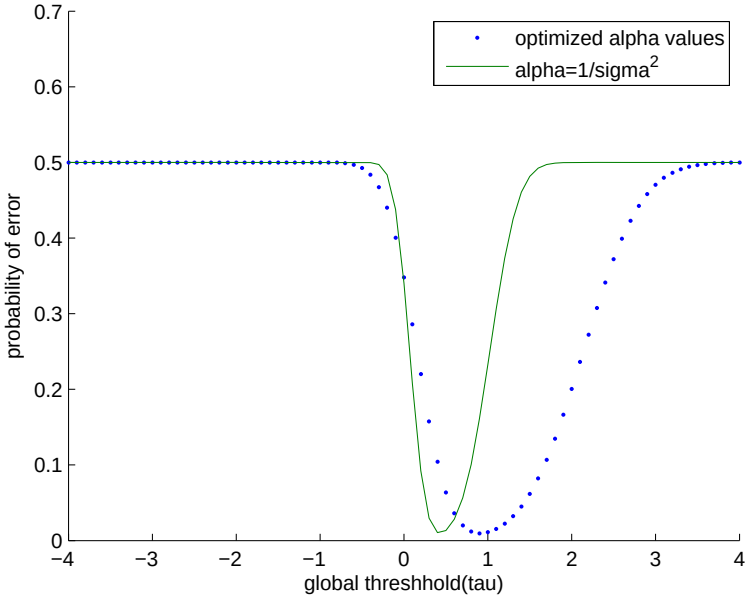

FIGURE 4.1. Probability of error vs global threshold.

criteria, we are now interested in the energy allocations for the nodes with a limit on the total energy for all the nodes. In this section, we will present the results for $N = 2$ and $N = 1$. For $N = 2$, we will be using QPSK modulation to transmit the quantizer outputs. The analytical formulation for this case is given in Chapter 2.

The WSN configuration for $L = 5$ and $L = 8$ are given in Table 4.1 and Table 4.6. The parameters to be optimized for this case are α_i, the quantizer thresholds t_{ij}, quantizer outputs q_{ij}, τ and θ_i. θ_i is the fraction of total energy E_T allocated to the sensor i. From Figure 4.1, we considered the optimal α_i to be $\frac{1}{{\sigma_i}^2}$. The optimal

quantizer thresholds obtained by solving the analytical formulations are given in Table 4.3. The corresponding quantizer outputs q_{ij} are presented in the Table 4.4. The spacing between the quantizer thresholds increases with the increase in the sensor noise variance. We also obtained the quantizer thresholds for Lloyd-Max quantizer (Table 4.5). The optimal quantizer thresholds follow the same trend as the Lloyd-Max quantizer. But the value of the thresholds are different to the Lloyd-Max quantizer.

TABLE 4.3. Optimum qunatizer thresholds for Four-level quantizer for QPSK Modulation for the Error Probability.

Sensor	t_1	t_2	t_3
1	-0.2443	0.0354	0.3151
2	-0.4694	-0.0671	0.3353
3	-1.0828	-0.0784	0.9261
4	-1.2740	0.0161	1.3602
5	-1.5888	0.0071	1.6301

TABLE 4.4. Optimum qunatizer output for Four-level quantizer for QPSK Modulation for the Error Probabiltiy Criteria.

Sensor	q_0	q_1	q_2	q_3
1	-2.5112	-0.1443	0.1360	2.6692
2	-2.8826	-0.1771	0.1653	2.7404
3	-2.9543	-0.6844	0.7009	2.9942
4	-3.1139	-0.9797	0.9294	3.1356
5	-3.2870	-0.9108	0.8802	3.2947

TABLE 4.5. Lloyd-Max qunatizer thresholds for the Four-level Quantizer.

Sensor	t_1	t_2	t_3
1	-0.8058	0	0.8058
2	-1.1491	0	1.1490
3	-1.470	0	1.471
4	-1.681	0	1.681
5	-1.8195	0	1.8195

The optimal energy allocation at the nodes for $L = 5$ is given by Figure 4.2. The corresponding values of τ and the error probability are presented in the Figure 4.2. The energy allocation for the nodes decreases with the increase in the noise variance

$(\sigma_i{}^2)$. This result sounds appropriate because the sensor with less noise variance provides more accurate information about the observed hypothesis at the fusion center for the decision making process. The sensor with a very high noise variance will be allocated no energy at all. This will censor the node from transmission. Thus saving the cost of energy and the bandwidth.

The optimal energy fractions (θ_i) at the nodes for $L = 8$ is presented in Figure 4.3. Similar to $L = 5$, the sensors with smaller noise variance are allocated higher energy for $L = 8$. The optimal τ and the resulting error probability are given in Figure 4.3. We can observe that the sensor with equal noise variance are allocated equal energy. The allocated energy is used by the sensor to modulate and transmit the symbols across the channel. The error probability for $L = 5$ and $L = 8$ is a quasi-convex function of τ for the given α_i and θ_i values. The convex function has its minimum at 0.0133 for $L = 5$. The minimum is at 0.062 for $L = 8$. It is noticeable that the error probability for $L = 8$ is better than that of $L = 5$. The error probability is expected to increase with the increase in the number of sensors (L). The error probability is also affected by the noise variances of the additional sensors. We also observed that the Lloyd-Max quantizer results in a low error probability. This is comparable to the error probability obtained by using the optimal theshhold. Thus we can reduce the complexity of the problem by using the Lloyd-Max quantizer.

TABLE 4.6. WSN Configuration for L=8.

Sensor node index(i)	1	2	3	4	5	6	7	8
σ^2	1	2	3	4	5	3	2	1

We now present the results for the binary level quantization. In this case the quantizer outputs are considered as the local decisions made by the sensors. Traditionally [34], identical local decisions are considered at the sensors. We consider

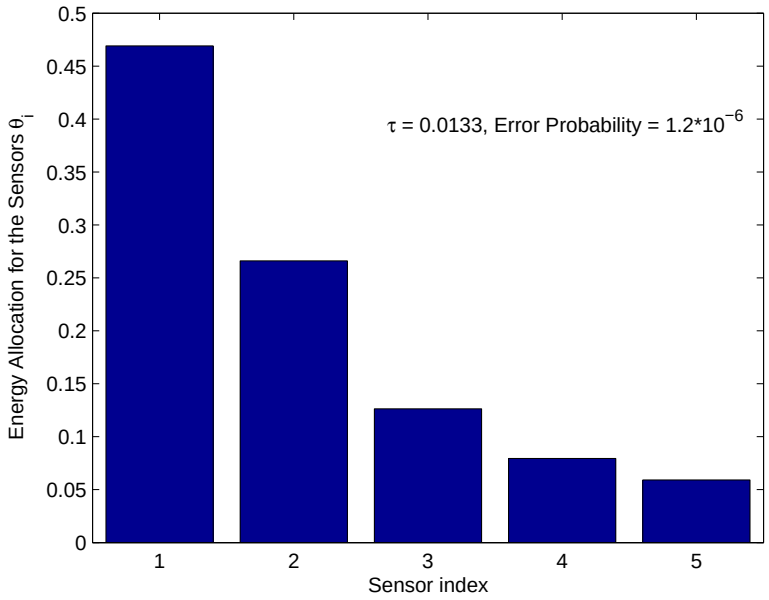

FIGURE 4.2. Energy Allocation for the Sensor Nodes for L=5 using Error Probability Criteria for Four-level Quantization

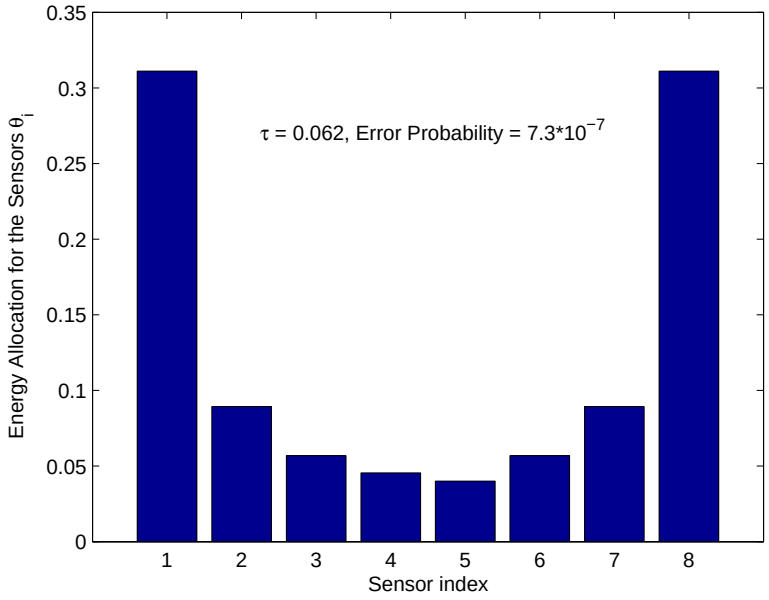

FIGURE 4.3. Energy Allocation for the Sensor Nodes for L=8 using Error Probability Criteria for Four-level Quantization

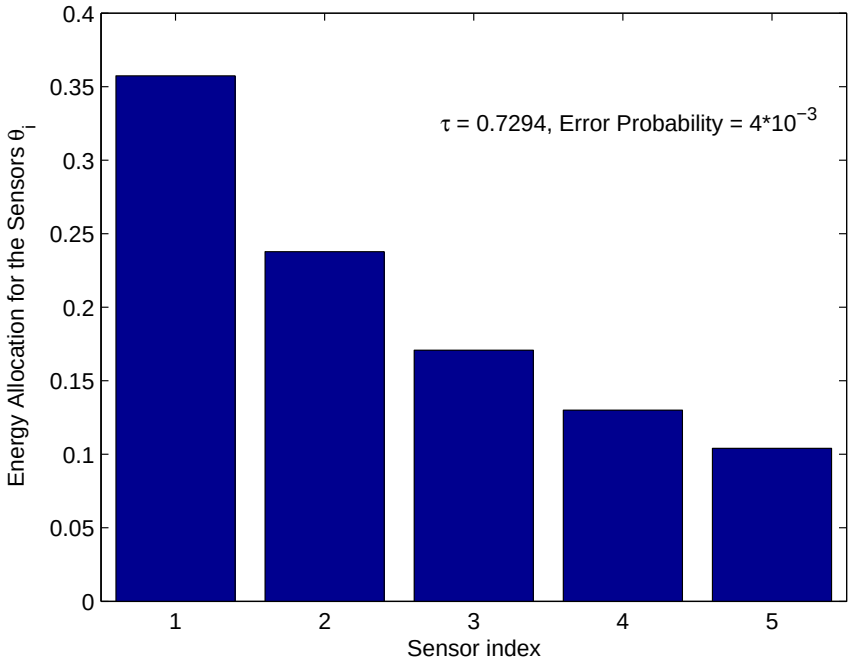

FIGURE 4.4. Energy Allocation for the Sensor Nodes for L=5 using Error Probability Criteria for Binary-level Quantization

non-identical local decisions because the observations made by the sensors are not i.i.d. The local decisions for the individual sensors is given by (3.21). The parameters to be optimized are τ and the energy fractions (θ_i). The quantizer outputs are modulated using binary modulation. Thus the channel is modeled by binary symmetric channel with cross over probability ($Q(\frac{E_T\theta_i}{N_0})$). The optimal energy allocations, τ and the resulting error probabilities for $L = 5$ are presented in Figures 4.4. As expected the sensors with smaller noise variance are allocated higher energy. The energy fractions for $L = 8$ (Figure 4.5) follow the same trend as the $L = 5$. As expected the error probability for $L = 8$ is better than that of $L = 5$ for the given noise variances.

Comparing the results for $N = 2$ with $N = 1$, we can observe that the error probability for $N = 2$ is better than that of $N = 1$. But the case $N = 1$ is easy to

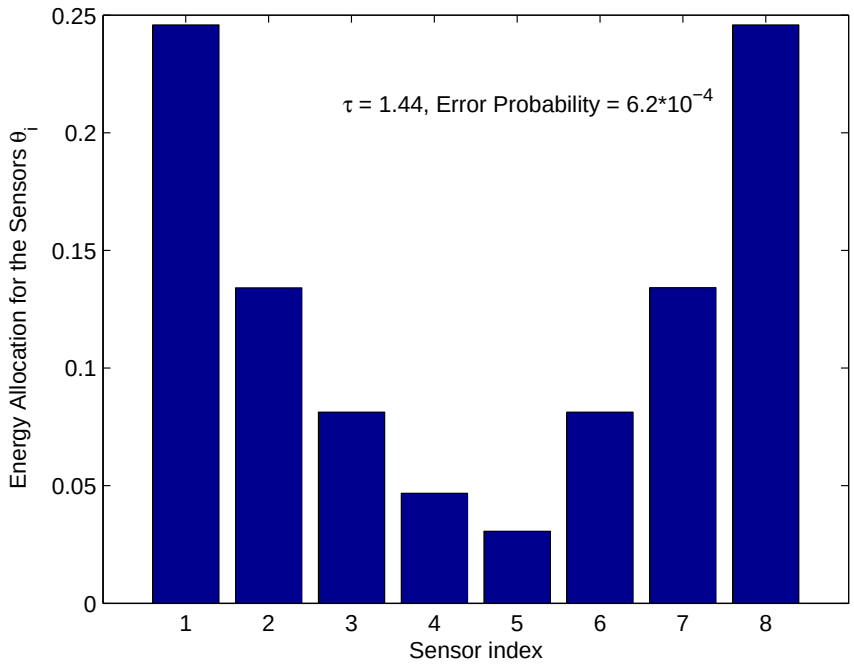

FIGURE 4.5. Energy Allocation for the Sensor Nodes for L=8 using Error Probability Criteria for Binary-level Quantization

design and has lot of work done with respect to the identical sensors ([31], [32], [34], [36] and [33]). The number of parameters to be optimized are naturally reduced with the quantizer levels. This will result in a trade off between the performance, complexity and the bandwidth requirement.

4.2.2 Distance Measure

In the previous section, we presented the results for the case where error probability is used as the cost function. This requires a closed form solution for the error probability. For the fusion rule (3.6), it is difficult to obtain a closed form solution for the error probability. Therefore we opt for an alternative cost function termed as the J-Divergence distance measure (3.10). Our goal is to design an optimal fusion rule and an energy allocation for the nodes so as to maximize the J-divergence cost function subject to a limit on the total energy of all the nodes. The analytical

formulation for the J-Divergence cost function is given in Chapter 3. In this section we will present the results that will corroborate the analytical formulation.

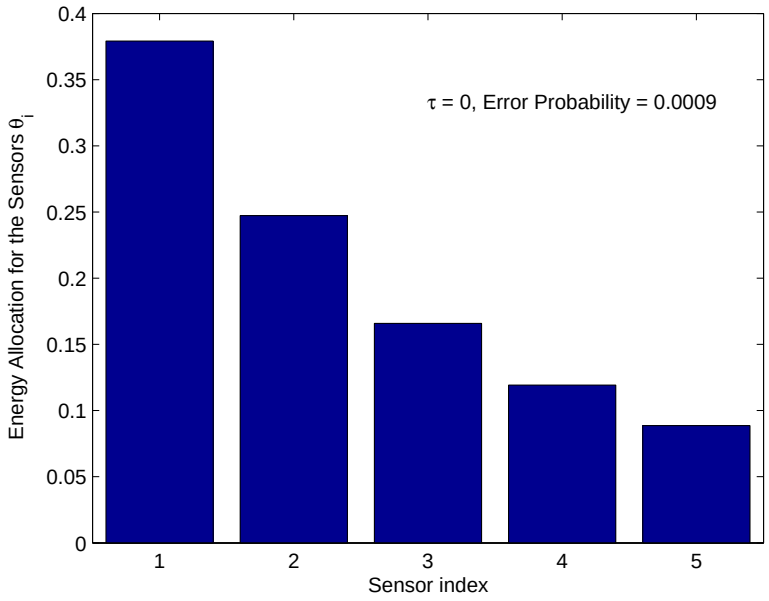

FIGURE 4.6. Energy Allocation for the Sensor Nodes for L=5 using Distance measure criteria for Four-level Quantization

TABLE 4.7. Optimum qunatizer thresholds for Four-level quantizer for QPSK Modulation for the Distance measure.

Sensor	t_1	t_2	t_3
1	-0.4305	0.02	0.4280
2	-0.6002	0.52	0.6001
3	-0.9179	0.0124	0.9175
4	-1.5896	0	1.5895
5	-1.8423	0	1.8423

For the distance measure criteria, we use the log-likelihood ratio test as the optimal fusion rule (3.6). According to [34], log-likelihood ratio test is an optimal fusion rule if you are willing to minimize the error probability of the system. The optimal energy calculations for the distance measure criteria are presented in the Chapter 3. The error probability for the distance measure is calculated by

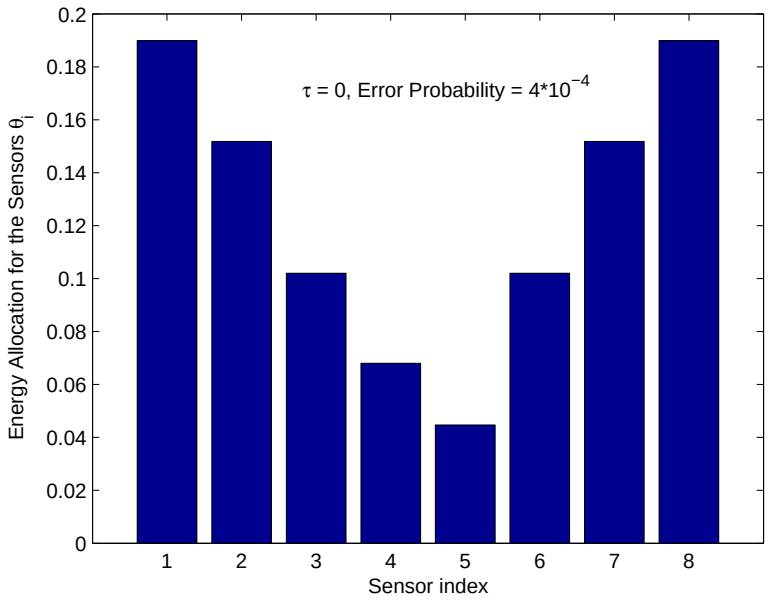

FIGURE 4.7. Energy Allocation for the Sensor Nodes for L=8 using Distance Measure criteria for Four-level Quantization

simulating the system model (Figure 2.3). The simulation is done at an optimal value of $\tau = 0$.

The optimal quantizer thresholds for the distance measure criteria are presented in Table 4.7. The quantizer thresholds follow the same trend as the error probability criteria. The spacing between the thresholds increases with the increase in the noise variances. The optimal energy allocations and the resulting error probabilities for $L = 5$ are given in Figure 4.6. The M-ary modulation provides energy allocation only among the sensors. As expected the sensor with less noise variance is allocated more energy. The optimal energy allocation for $L = 8$ is given by Figure 4.7. The energy fractions for the sensors increase with the decrease in the sensor noise variance and the sensors with equal noise variance are allocated equal energy. The error probability for $L = 8$ is presented in Figure 4.7. As discussed, the error probability is obtained by simulating system model (Figure 2.3) at $\tau = 0$. The error probability for $L = 8$ is better than that of $L = 5$ for the distance measure

criteria. The results for the distance measure follow the same trend as the error probability criteria.

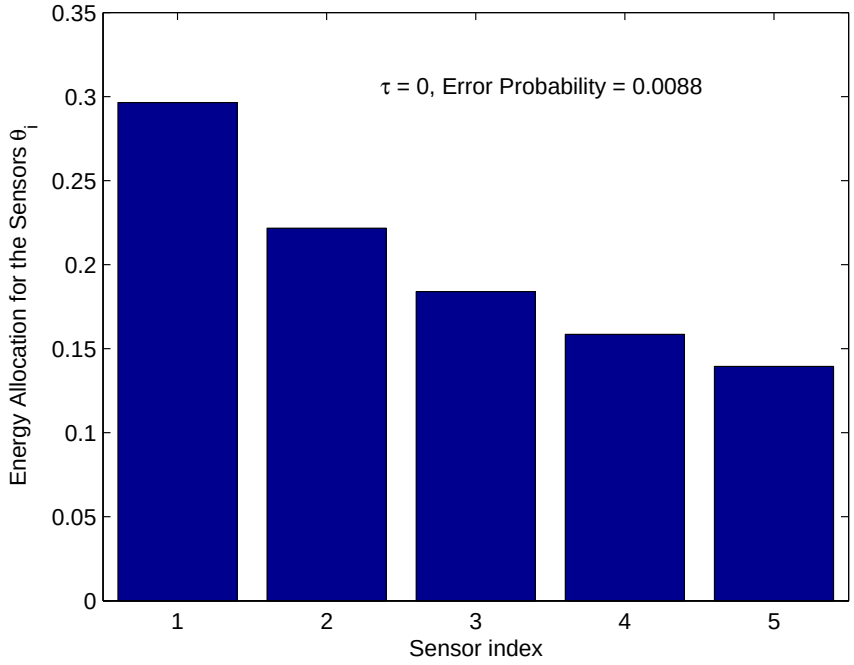

FIGURE 4.8. Energy Allocation for the Sensor Nodes for L=5 using Distance Measure for binary quantization

TABLE 4.8. WSN Configuration for L=5.

Sensor index (i)	1	2	3	4	5
${\sigma_i}^2$	1	4	9	16	25

We now present the results for the binary-level quantization. The local thresholds for the quantizer are given by (3.21). The results (energy fractions and the error probabilities) for the binary-level quantization are presented in Figures 4.8 for $L = 5$. The sensor with less noise variance gets more energy as expected. Figure 4.10 shows that the sensor with a very high noise variance (${\sigma_i}^2 = 25$) will be allocated no energy. The WSN configuration for the Figure 4.10 is given by Table 4.8. The modulation and transmission costs are reduced for the given WSN configuration (Table 4.8) because the sensor with a very high noise variance is censored.

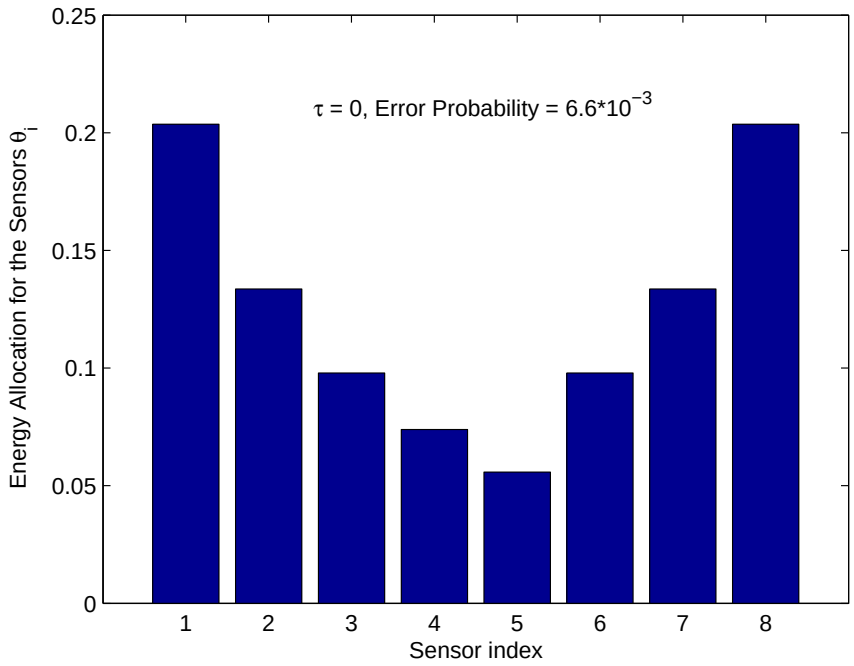

FIGURE 4.9. Energy Allocation for the Sensor Nodes for L=8 using Distance Measure for binary quantization

The energy fractions for $L = 8$ is presented in Figure 4.9. The energy fractions for $L = 8$ follows the same trend as $L = 5$. The sensor with smaller noise variance gets higher energy. The error probability for $L = 8$ is better than that of $L = 5$. We can notice a decrease in the error probability with an increase in the quantization levels. We can observe that the error probability for the error probability criteria is better than that of the distance measure criteria for the M-ary modulation case. This trend follows for both the four-level quantization and binary-level quantization. However, the distance measure criteria is easy to design and provides a more tractable solution. The distance measure requires less amount of complexity compared to the error probability.

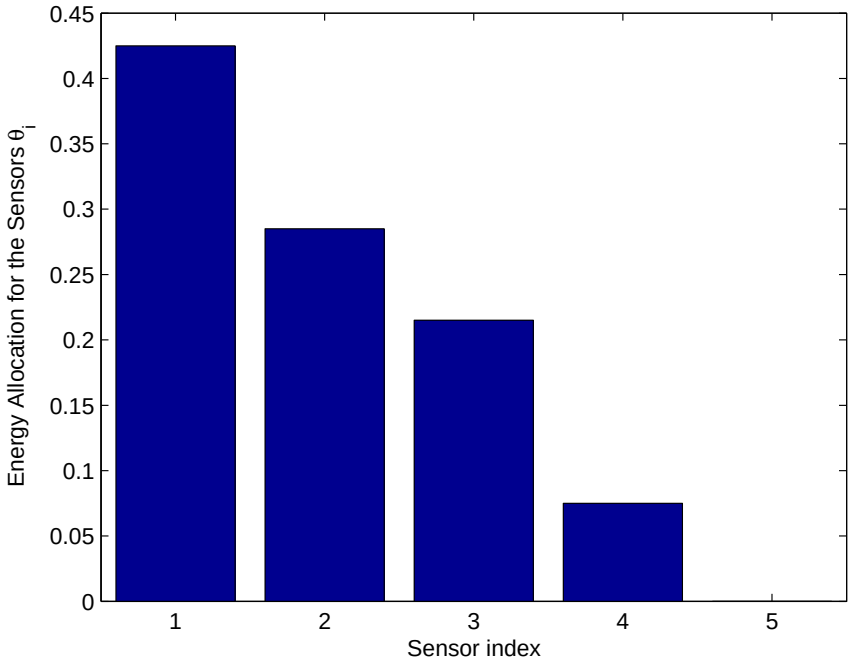

FIGURE 4.10. Energy Allocation for the Sensor Nodes for L=5 using Distance Measure for binary quantization

4.3 Binary Modulation

In the M-ary modulation case, we observed that there is energy allocation only among the sensors. We could not protect the individual bits because all the bits are allocated equal energy. In order to protect the individual bits, we transmit them individually across the channel using the binary modulation. This will provide energy allocation among the sensors and the individual bits. For the sake of concreteness the sensors use BPSK modulation. Thus the channel is modeled as a binary symmetric channel with cross over probability $\epsilon_i = Q(\sqrt{\frac{2\beta_{ij}\theta_i E_T}{N_0}})$. The channel output is demodulated into bits which are used to reconstruct the observation signal X_i, denoted by $\hat{X}_i$. The fusion center makes a decision on the value of H using the reconstructed signal sequence. The goal is to design an optimum fusion rule and energy allocations so as to optimize the cost functions. The ana-

lytical formulation for both the cost functions is given in Chapter 2 and Chapter 3. The results obtained by solving the analytical formulations are presented in this section.

4.3.1 Error Probability Criteria

As discussed previously, the fusion rule (2.3) is considered optimal for $\alpha_i = \frac{1}{{\sigma_i}^2}$ given the fusion center has access to the complete observations. From the Figure 4.1, we can observe that the fusion rule (2.3) is optimal for $\alpha_i = \frac{1}{{\sigma_i}^2}$ even if the fusion center does not have access to the complete observations.

TABLE 4.9. Optimum qunatizer thresholds for Four-level quantizer for Binary Modulation for the Error Probability criteria.

Sensor	t_1	t_2	t_3
1	-1.0176	0	1.0176
2	-1.2733	0	1.2733
3	-1.9987	0.001	1.9985
4	-2.1261	0	2.1261
5	-2.6334	0.0002	2.6335

TABLE 4.10. Optimum qunatizer output for Four-level quantizer for Binary Modulation for the Error Probability criteria.

Sensor	q_0	q_1	q_2	q_3
1	-2.4833	-0.593	0.593	2.4833
2	-2.9025	-0.766	0.766	2.9025
3	-2.9663	-0.952	0.952	2.9663
4	-2.9820	-0.978	0.978	2.9820
5	–2.9841	-1.0581	1.0581	2.9841

The optimal quantizer thresholds and outputs for this case are presented in Table 4.9 and Table 4.10. The spacing between the thresholds increases with the increase in the noise variances. They follow the same trend as the Lloyd-Max quantizer. The optimal energy allocation for the binary modulation includes the energy allocation at the nodes and the individual bits. The optimal energy fractions (θ) at the nodes are depicted in Figure 4.11. The optimal value of τ and the resulting error probability are presented in the Figure 4.11. As expected the sensors with smaller

noise variance are allocated a higher fraction of the energy. Figure 4.11 depicts the fraction of the total energy E_T allocated to the nodes. The total energy allocated to the nodes for processing its observations is given by $E_T\theta_i$. This energy ($E_T\theta_i$) is used by the sensor i to modulate and transmit the quantized bits across the channel to the fusion center for the final decision on the value of H.

The energy allocations β_{ij} for the individual bits is given by the Figure 4.12. From the Figure 4.12, we can observe that most of the energy is allocated to the most significant bit for a SNR of 10 dB. This will reduce the bandwidth requirement because the second bit is censored from transmission. The bitwise energy allocation for a SNR of 20 dB is given in Figure 4.13. This shows the priority of the the most significant bit over the remaining bits. This is an important result because an error in the most significant bit causes the reconstructed signal $\hat{X}_i$ to favor a hypothesis which is different to the observed hypothesis.

The error probability for the binary modulation is similar to the M-ary modulation. But the binary modulation reduces the bandwidth requirement because some of the bits are censored from transmission. Thus saving the bandwidth without any degradation in the performance.

The energy fractions for $N = 3$ at SNR of 10 dB is presented in Figure 4.14. Similar to $N = 2$, all the energy is allocated to the first bit for $N = 3$ at SNR of 10 dB. Thus saving the bandwidth by censoring the remaining bits. Therefore it is advisable to use binary-level quantization at a low SNR. The energy fractions at SNR of 20 dB are given in the Figure 4.15. The Figure 4.15 shows the importance of the first bit over the remaining bits.

We can observe from the results that the amount of energy allocated to the most significant bit decreases with the SNR. In order to fortify this observation, we determined the energy allocations for the bits at different SNR (Figure 4.16).

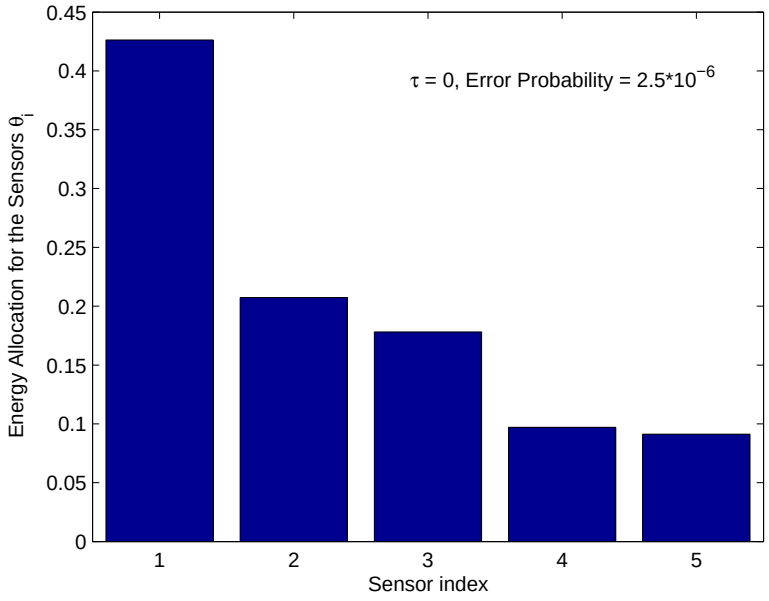

FIGURE 4.11. Energy Allocation for nodes using Error Probability Criteria for Four-level Quantization

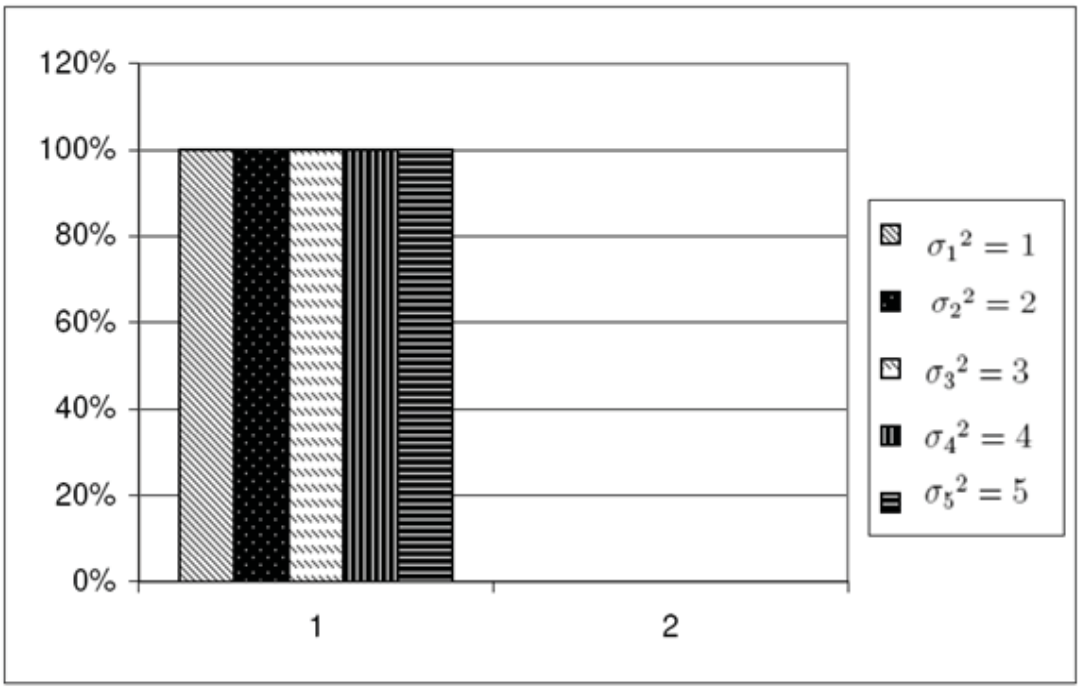

FIGURE 4.12. Bitwise Energy Allocation for a Four-level Quantization at SNR of 10 dB for Error Probability Criteria

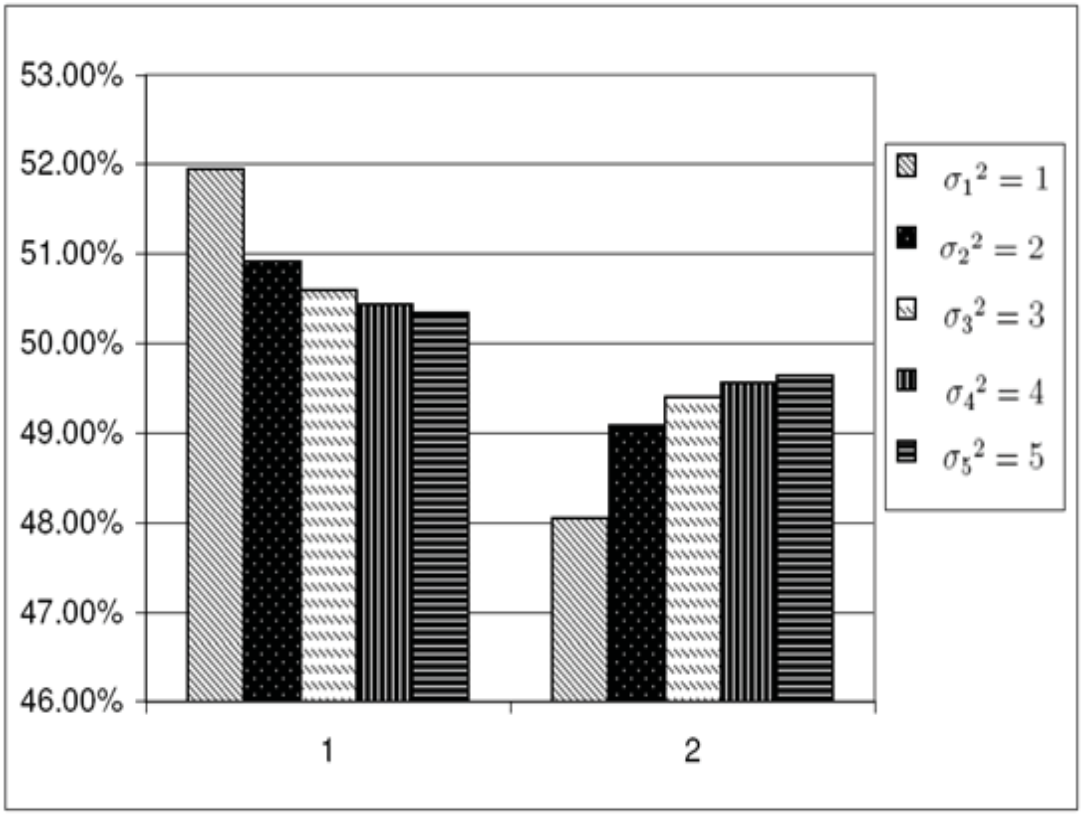

FIGURE 4.13. Bitwise Energy Allocation for a Four-level Quantization at SNR of 20 dB for Error Probability Criteria

We plotted the bitwise energy allocations for different SNR in Figure 4.16. The energy fraction for the first bit decreases with the SNR and remains constant at 0.333 for higher SNR. The energy fraction for the remaining bits increases with the SNR and remains constant at higher SNR. Thus it is advisable to use low-level quantization for a low SNR. From Figure 4.15 and Figure 4.13, we can notice that the energy allocation to the first bit decreases with the increase in the noise variances (σ_i^2) for the error probability criteria.

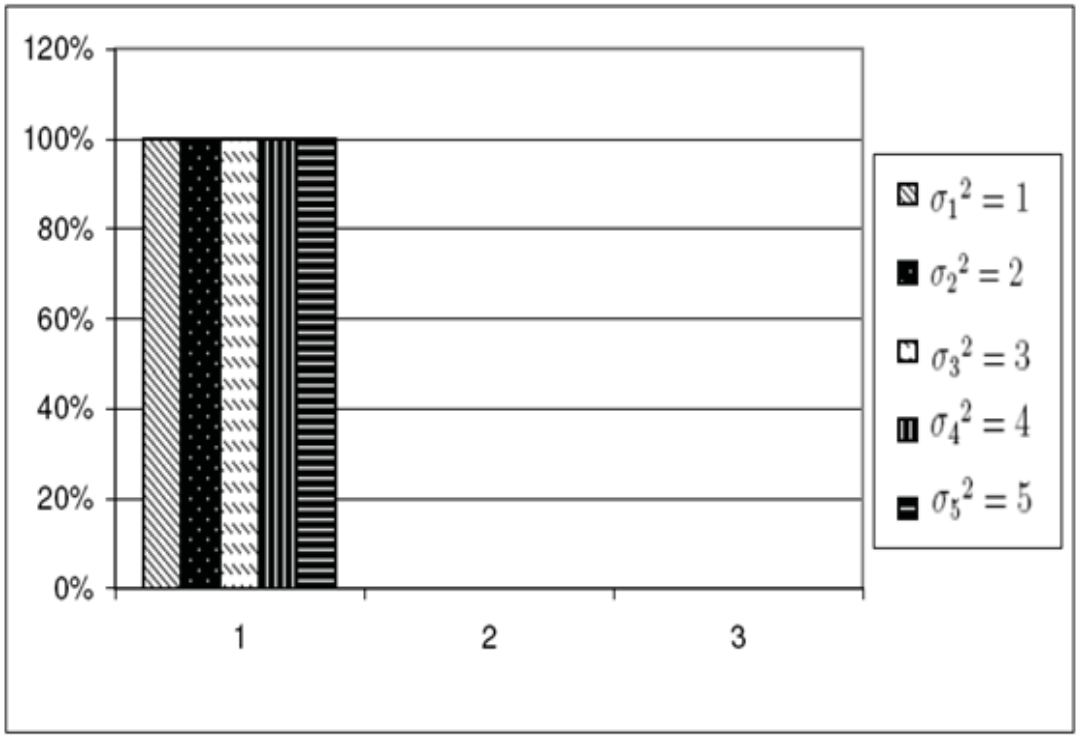

FIGURE 4.14. Bitwise Energy Allocation for a Eight-level Quantization at SNR = 10 dB for Error Probability Criteria

4.3.2 Distance Measure

In this section we will present the optimal fusion rule and the energy allocations that maximizes the J-divergence distance measure with a limit on the total energy of the network.

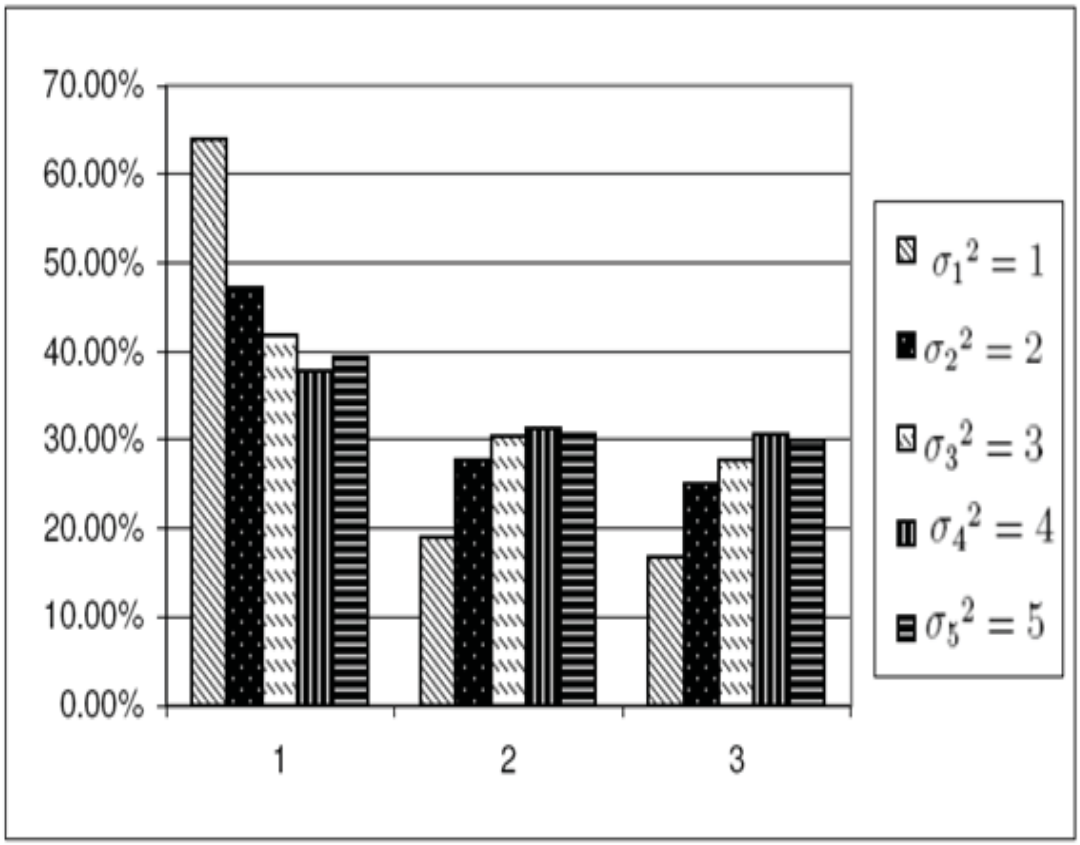

FIGURE 4.15. Bitwise Energy Allocation for a Eight-level Quantization at SNR = 20 dB for Error Probability Criteria

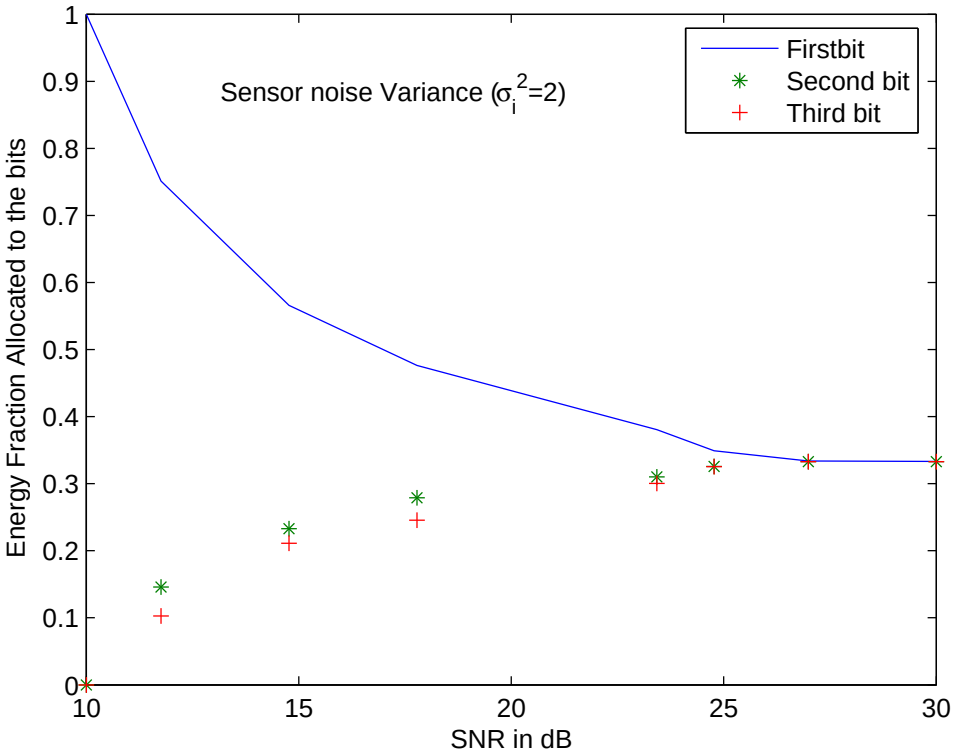

FIGURE 4.16. Energy Allocation for the bits at different SNR

As discussed previously, the log-likelihood ratio test is considered as an optimal fusion rule for the distance measure criteria [34]. The optimal quantizer thresholds are given by Table 4.11. We now determine the optimal energy allocation for the wireless sensor network. The optimal energy fractions (θ_i) at the nodes for $N = 2$ is given by Figure 4.17. The resulting error probability is presented in the Figure 4.17. The energy allocation at the nodes increases with the decrease in the sensor noise variances. The total energy allocated to the sensor i is given by $E_T\theta_i$. This energy is used by the sensor to modulate and transmit its observations. The amount of energy required to transmit the individual bits is given in Figure 4.18. We can observe that all the sensor energy ($E_T\theta$) is allocated to the most significant bit for a small SNR ($SNR = 10$ dB). This reduces the bandwidth requirement because the second bit is censored from transmission. We also obtained the bitwise energy allocation for $SNR = 20$ dB (Figure 4.19). Similar to the error probability criteria, the most significant bit is given more priority compared to the second bit for the distance measure criteria. The error probability for the error probability criteria is better than that of Distance measure.

TABLE 4.11. Optimum qunatizer thresholds for Four-level quantizer for Binary Modulation for the Distance measure.

Sensor	t_1	t_2	t_3
1	-1.44	0	1.44
2	-1.6456	0	1.6476
3	-1.767	0.0176	1.7720
4	-1.9749	0	1.9751
5	-2.101	0	2.102

The bitwise energy allocation for $N = 3$ at SNR of 10 dB is given by Figure 4.20. The second and the third bit are censored from transmission because all the energy is allocated to the first bit. This will reduce the cost of transmission and the bandwidth requirement. The bitwise energy allocation at SNR of 20 dB

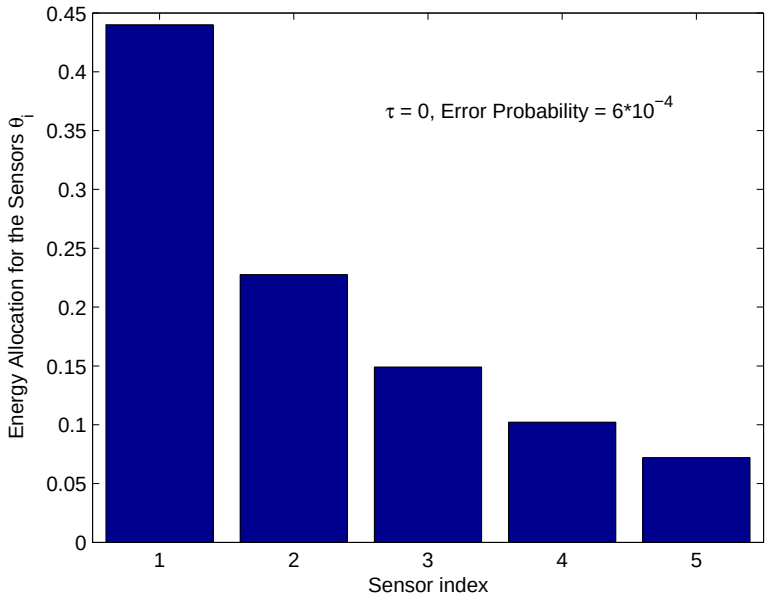

FIGURE 4.17. Energy Allocation for nodes using Distance measure for Four-level Quantization

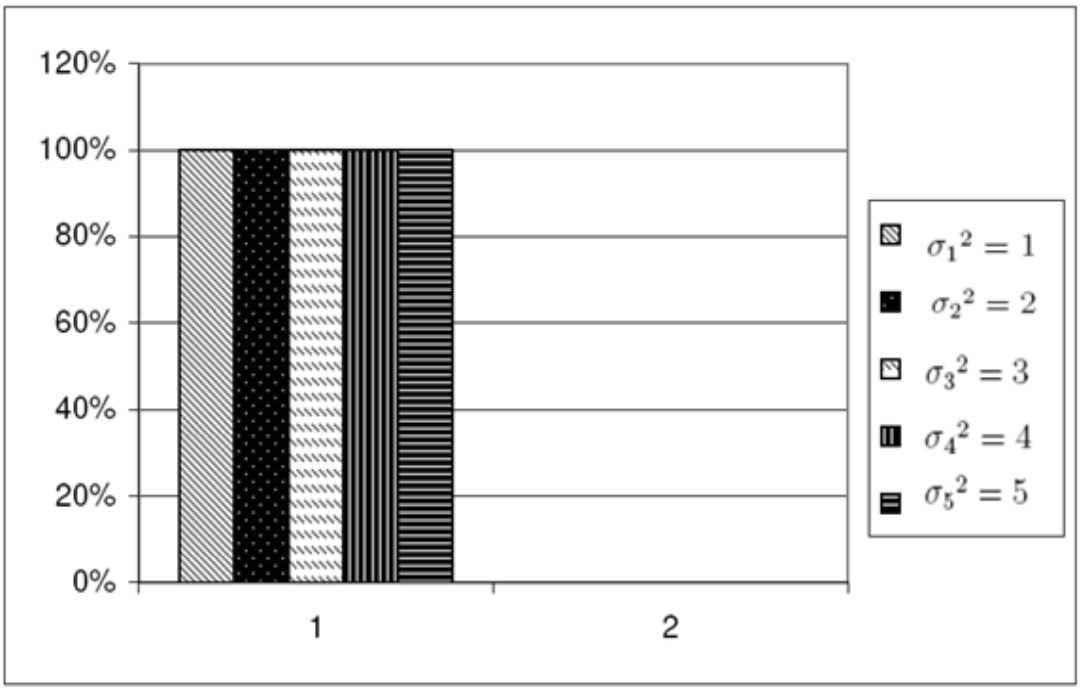

FIGURE 4.18. Bitwise Energy Allocation for a Four-level Quantization at SNR of 10 dB for Distance Measure Criteria

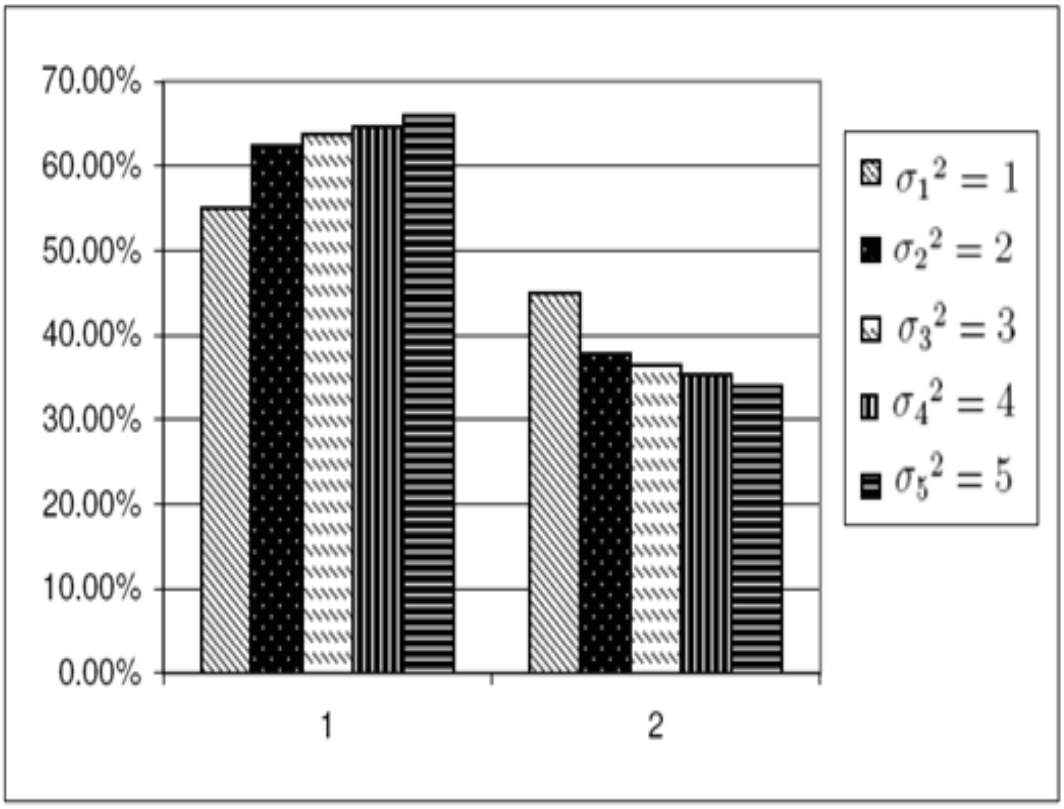

FIGURE 4.19. Bitwise Energy Allocation for a Four-level Quantization at SNR of 20 dB for Distance Measure Criteria

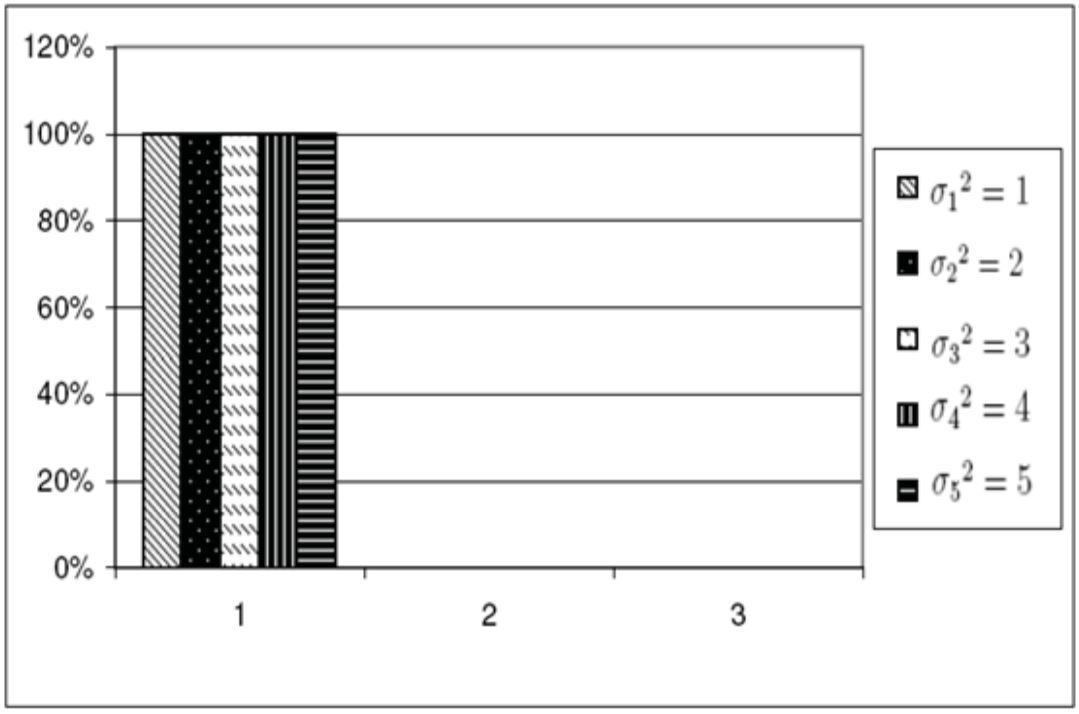

FIGURE 4.20. Bitwise Energy Allocation for a Eight-level Quantization at SNR = 10 dB for Distance Measure Criteria

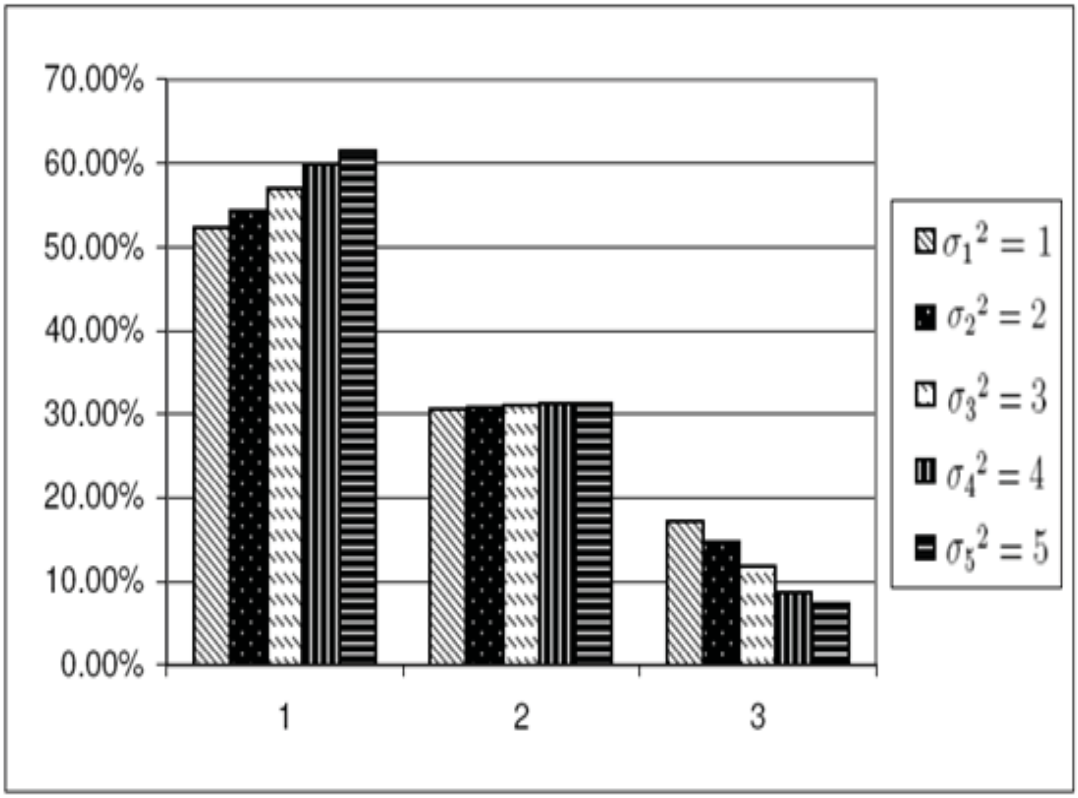

FIGURE 4.21. Bitwise Energy Allocation for a Eight-level Quantization at SNR = 20 dB for Distance Measure Criteria

is presented in Figure 4.21. Figure 4.21 demonstrates the importance of the first bit over the remaining bits. This result is important because an error in the first bit will cause the reconstructed variable $\hat{X}_i$ to favor a hypothesis different to the observed hypothesis. Similar to the error probability criteria, the energy allocated to the first bit decreases with the SNR for the distance measure. The amount of energy allocated to the first bit increases with the noise variances of the sensor for the distance measure (Figure 4.19 and Figure 4.21). This trend follows for all values of SNR.

4.4 Comparisons

The efficiency of the detection process is determined by the error probability. The error probabilities for different methods used in this thesis are presented in the Table 4.12. The error probability for the distance measure is calculated by simulating the system model for a global threshold of $\tau = 0$ because it is difficult to obtain a closed form for the error probability.

TABLE 4.12. Error Probabilities for Different methods for the two Cost functions

Cost functions	Error Probability	Distance Measure
Bitwise Allocation	$2.5 * 10^{-6}$	$6 * 10^{-4}$
Multiple Bits with Equal Energy	$1.2 * 10^{-6}$	$9 * 10^{-4}$
Single bit case	$4 * 10^{-3}$	$8.8 * 10^{-3}$

Although the error probabilities are similar for the cases where binary modulation and M-ary modulation is used to transmit the quantizer outputs, the amount of bandwidth required is reduced for binary modulation at small SNR. This is because most of the energy is allocated to the first bit for a small SNR. This trend follows for both the cost functions, error probability and distance measure.

The error probability for the error probability cost function is better than that of distance measure cost function. Although the distance measure has less error probability, it results in a much easier and tractable solution for the system design.

The amount of complexity required is less for the distance measure. Thus creating a trade off between performance, complexity and the bandwidth requirement. We can also observe an increase in the performance with the quantization levels. The results for the distance measure follow the same trend as the error probability criteria. Therefore distance measure is an appropriate alternative for the error probability.

Chapter 5
Conclusions

We have studied the problem of binary hypothesis testing in a wireless sensor network in the presence of noisy channels and for non-identical sensors. We have designed a mathematically tractable fusion rule for which optimal energy allocation for individual sensors can be achieved. We also designed an optimal energy allocation for the bits when multi-level quantization is used by the sensors to quantize its observation. The objective is to optimize a cost function with a constraint on the total network energy. Two cost functions were considered; the probability of error and the J-divergence distance measure.

We have presented the results of optimal energy allocation for sensors and bits for the case where binary modulation is used to transmit the quantized observations. For M-ary modulation, we presented the optimal energy allocation at the sensors. The optimal fusion rule, energy allocation and the resulting error probability are presented for the two cost functions. Comparisons are drawn between the performance of the two cost functions. The error probability cost function has better error probability compared to the J-divergence cost function. However, the J-divergence cost function yields a much easier and tractable solution.

This work can be extended in many ways. We could incorporate a fading channel between the sensors and the fusion center. Diversity techniques could be used at the fusion center to process the signal received. This thesis assumes a constraint on the total energy of the network. We can extend it by assuming a limit on the individual sensor energy.

References

[1] S. M. Ali and S. D. Silvey. A general class of coefficients of divergence of one distribution from another. *J.Royal Stat. Soc. Series B*, 28:131–142, 1966.

[2] S. Appadawedula, V. V. Veerwalli, and D. L. jones. Energy-efficient detection in sensor networks. *IEEE J. Select . Areas. commun*, 23(4):693–702, June 2005.

[3] K.S.J. Pister B. Warneke. Mems for distributed wireless sensor networks. In *9th Int'l Conf on Electronics, Circuits and Systems*, Dubrovnik,Croatia, Sep. 2002.

[4] B.chen and P.k Willet. On the optimality of the likelihood ratio test for local sensor decision rules in the presence of non-ideal channels. *IEEE trans on Signal processing*, 51(2):407–416, Feburary 2003.

[5] V. Borkar and P. Varaiya. Asymptotic agreement in distributed estimation. *IEEE trans. Automat. Contr.*, 27(3):873–876, June 1982.

[6] Stephen Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.

[7] J. F. chamberland and V. V. Veeravalli. Wireless sensors in distributed detection applications. *IEEE Signal Processing Magazine*, pages 16–24, May 2007.

[8] Biao Chen, Ruixiang Jiang, Teerasit Kasetkasem, and Pramod K .Varshney. Channel aware decision fusion in wireless sensor networks. *IEEE trans on Information Theory*, 52(12):3454–3458, December 2004.

[9] Shuguang Cui, Andrea J goldsmith, and Ahmad Bahai. Energy-constrained modulation optimization. *IEEE trans. Automat. Contr.*, 4(5):2349– 2360, September 2005.

[10] E. V. Eremin. A central limit theorem for linear estimators of measurement results. *Measurement Techniques*, 21(11):1009–1016, October 2006.

[11] A. J. goldsmith and S. B. Wicker. Design challenges for energy constrained ad hoc wireless sensor netwroks. *IEEE Wireless Communications Magazine*, 9(4):8–27, August 2002.

[12] Harry L. Van TreesMichel Goossens, Frank Mittlebach, and Alexander Samarin. *Detection , Estimation and Modulation theory*. Wiley, New york, 1968.

[13] T. L. Grettenberg. Signal selection in communication and radar systems. *IEEE trans. Information Theory*, IT-9:265–275, October 1963.

[14] Simon Haykin. *Communication Systems*. Wiley, New york, 2001.

[15] I. Duncan I. Khemapech and A. Miller. *A Survey of Wireless Sensor Networks Technology*. University of St Andrews, North Haugh, St Andrews.

[16] W. Ye. J, Heidemann, and D. Estrin. An energy efficient mac protocol for wireless sensor networks. In *21st Annual Joint Conference of the IEEE Computer and Communications socities*, volume 3, pages 1567–1576, New york, June 2002.

[17] Holger Karl. *A Short Survey of Wireless Sensor Networks*. Telecommunication Networks Group,Technical university at Berlin.

[18] W. Karush. Minima of functions of several variables with inequalities as side constraints. In *M.Sc. Dissertation.*, Dept. of Mathematics, Univ. of Chicago, Chicago, Illinois., 1939.

[19] Alexey kranospeev, jin-jun Xiao, and Zhi-Quan Luo. Minimum energy decentralized estimation in a wireless sensor network with correlated sensor noises. *EURASIP Journal on Wireless Communications and Networking*, 2005(4):473–482, 2005.

[20] H. W. Kuhn and Tucker. Nonlinear programming. In *2nd Berkeley Symposium*, pages 481–492, University of california, Berkeley, 1951.

[21] S. Kullback. *Information Theory and Statistics*. Wiley, New york, 1959.

[22] X. Luo and G. B Giannakis. Energy-constrained optimal quantization for wireless sensor networks. In *1st IEEE Annual Communications Society Conference on Sensor and Ad Hoc Communications and Networks*, 2004.

[23] K. Matsusita, Y. Suzuki, and H. Hudimoto. On testing statistical hypothesis. *Ann. Inst. Stat. Math*, 6:133–141, 1954.

[24] A. Michail and A. Ephermides. Energy-efficient routing for connection oriented traffic in ad-hoc wireless networks. In *11th IEEE International Sysposium on Personal Indoor and Mobile Radio Communications*, pages 762–766, London, September 2002.

[25] Chongzhao Han Ming Xiang. Optimization of distributed networks with tree structures. In *Fifth International Conference on Information Fusion*, volume 1, pages 164–169, 2002.

[26] H. Vincent Poor and John B.Thomas. Applications of ali-silvey distance measures in the design of the generalized quantizers for binary decision systems. *IEEE Trans on Communications*, 25(9):893–900, September 1977.

[27] P.Swaszek and P.Willet. Parley as an approach to distributed detection. *IEEE trans on aerospace.Electron.Syst*, 31(1):447–457, January 1995.

[28] J. M. Rabaey, M. J. Ammer, J. L. da. Silva, and S. Roundy. Picoradio supports ad hoc ultra-low power wireless networking. *IEEE computer*, 33:42–48, July 2000.

[29] C. Rago, P. Willet, and Y. Bar-shalom. Censoring sensors: a low communication-rate scheme for distributed detection. *IEEE trans. Aerosp. Electron. Syst*, 32(2):554–568, April 1996.

[30] R.Nui, B. chen, and P. K. Varshney. Fusion of decisions transmitted over rayleigh fading channels in wireless sensor networks. *IEEE trans. Signal processing*, 54(3):1018–1027, March 2006.

[31] R.R. Tenney and N.R. Sandell. Detection with distributed sensors. *IEEE Trans on Aerospace and Electronice Systems.*, 17(4):501–509, July 1981.

[32] John Tsitiklis. Decentralized detection. *Advances in Signal Processing*, 2:297–304, 1993.

[33] John Tsitsiklis. Decentralized detection by a large number of sesnors. *Math. Control, Signals, System*, 1(2):167–182, 1988.

[34] Pramod K. Varshney. *Distributed Detection and Data Fusion*. Springer, New york, 1997.

[35] P. Willet, P. F. Swaszek, and R. S. Blum. The good, bad and ugly: Distributed detection of a known signal in dependent gaussian noise. *IEEE trans on Signal Processing*, 48(12):3266–3279, December 2000.

[36] Qain Zhang, Pramod K. Varshney, and Richard D. Wesel. optimum bi-level quantization of i.i.d sensor observations for binary hypothesis testing. *IEEE trans on Information theory*, 48(7):2105–2111, 1993.

Vita

Krishna K. Gunturu was born in Guntur, India. He did his schooling at Loyola Public School and Gowthan Junior college. He later got the degree of Bachelor of Engineering in Electrical and Communications Engineering from Andhra University in 2005. He joined the Department of Electrical Engineering at Louisiana State University to do his master's in 2005. He will be graduating with a master's degree in electrical engineering in December 2007.