PROBLEMS IN DISTRIBUTED SIGNAL PROCESSING IN WIRELESS SENSOR NETWORKS.

by

RAJET KRISHNAN

B.Tech., University of Kerala, India, 2004

A THESIS

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

Department of Electrical and Computer Engineering

College of Engineering

KANSAS STATE UNIVERSITY

Manhattan, Kansas

2009

Approved by:

Major Professor

Dr. Balasubramaniam Natarajan

ABSTRACT

In this thesis, we first consider the problem of distributed estimation in an energy and rate-constrained wireless sensor network. To this end, we study three estimators namely - (1) Best Linear Unbiased Estimator (BLUE-1) that accounts for the variance of noise in measurement, uniform quantization and channel, and derive its variance and its lower bound; (2) Best Linear Unbiased Estimator (BLUE-2) that accounts for the variance of noise in measurement and uniform quantization, and derive lower and upper bounds for its variance; (3) Best Linear Unbiased Estimator (BLUE-3) that incorporates the effects of probabilistic quantization noise and measurement noise, and derive an upper bound for its variance.

Then using BLUE-1, we analyze the tradeoff between estimation error (BLUE variance) at the fusion center and the total amount of resources utilized (power and rate) using three different system design approaches or optimization formulations. For all the formulations, we determine optimum quantization bits and transmission power per bit (or optimum actions) for all sensors jointly. Unlike prior efforts, we incorporate the operating state (characterized by the amount of residual battery power) of the sensors in the optimization framework. We study the effect of channel quality, local measurement noise, and operating states of the sensors on their optimum choice for quantization bits and transmit power per bit.

In the sequel, we consider a problem in distributed detection and signal processing in the context of biomedical wireless sensors and more specifically pulse-oximeter devices that record photoplethysmographic data. We propose an automated, two-stage PPG data processing method to minimize the effect of motion artifact. Regarding stage one, we present novel and consistent techniques to detect the presence of motion artifact in photoplethysmograms given higher order statistical information present in the data.For stage two, we propose an effective motion artifact reduction method that involves enhanced PPG data preprocessing followed by frequency domain

Independent Component Analysis (FD-ICA). Experimental results are presented to demonstrate the efficacy of the overall motion artifact reduction method.

Finally, we analyze a wireless ad hoc/sensor network where nodes are connected via random channels and information is transported in the network in a cooperative multihop fashion using amplify and forward relay strategy.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Acknowledgements

I would like to express my sincere gratitude to my advisor Dr. Bala Natarajan for his support, advice and friendship. He has not only given me wonderful insights into academics and research, but has also shared his invaluable experiences and lessons with me, which I am sure to cherish for the rest of my journey that I call life. I have thoroughly enjoyed various discussions, though impolite at times, that we have had and I thank him for patiently listening to me even when I would be wrong.

I would like to thank Dr. Caterina Scoglio for not only agreeing to serve on my committee but also introducing me to the complex and beautiful world of networking and control theory. Due thanks to Dr. Steve Warren for serving on my committee and also for the great help and support that he rendered to my work in Bio-medical signal processing.

It has been a pleasure to know and interact with Dr. Todd Easton from Industrial and Manufacturing Engineering Department - he has been a great teacher, and a source of several stimulating discussions that spanned across topics in graph theory, optimization, and complexity theory. His perspective on life has had a profound influence on me. I would like to thank Dr. Dave Auckly from Mathematics Department who helped me gain a strong and clear understanding of the mathematical theory of optimization.

I have had a great time being a part of the WiCom research group at K-State and I would like to thank my current and former colleagues - Krithika, Lutfa, Dalin,

Ahmad, Mark, Sohini, and Sunitha for their support, friendship and encouragement.

Over the past two and a half years in this journey of life, my path has criss-crossed with several wonderful people whose friendship I am sure to treasure lifelong - Anirudh, Arun, Ashish, Narayanan, Niranjan, Pritish, Sambit, Vishal and many more. I would like to thank them for making my life at K-state a memorable one.

This thesis would not have been possible without the unconditional love and care, and unwavering support of my entire family - in particular my love Mamta, my lovely siblings - Renju, Ichamma, my terrific *Aliyan*, my sweet nephew Gugu, and my great parents - Jaya and Radhakrishnan Nair. I feel fortunate and privileged for being loved and cared by them and for having them in my life.

Finally, I would like to dedicate my thesis to my late grandmother, *Ammachi*, whose love, courage, energy and care have influenced me, my life, my actions, my spirits and thoughts like no one else.

Chapter 1

Introduction

A wireless sensor network (WSN) consists of spatially distributed autonomous devices that are capable of communicating with each other or a fusion center wirelessly. Typically, they are employed for detecting or estimating an underlying physical phenomenon without human intervention. This requires the participating sensors to collect local information related to the physical process of interest, and then wirelessly communicate relevant information to other sensors or a fusion center.

These wireless sensor networks are well-suited for surveillance and monitoring applications(e.g., military surveillance, environment and habitat monitoring, traffic surveillance, smart homes and health care). For instance, a WSN can be used to detect/track targets, and coordinate actions among combat units based on data collected from the battlefield. In health care, the combination of body-area sensors and environmental sensors embedded in a home can be used for real-time health-monitoring and care. In short, these sensor networks are steadily leading the world towards an information technology revolution where networked autonomous devices are connecting the physical world, devices and humans like never before.

1.1 Distributed Detection

In a classical distributed detection setting, information regarding a state H is received by a set of distributed nodes. Based on observation, each node sends relevant data or decision to the fusion center that makes a final decision on the state by employing appropriate data/decision fusion. In the context of distributed detection in WSNs, factors such as spectral bandwidth constraints, energy constraints, imperfect communication between nodes and the fusion center are incorporated in the classical framework. However, if there are no constraints on the spectral bandwidth or imperfect communication, complete information can be transported to the fusion center for processing without distortion. The nodes may make a hard decision (binary decision) and transmit this information to the fusion center for decision fusion. The nodes may also choose to send multi-level decisions or soft decisions that indicate the level of confidence in their decisions. Then the fusion center is faced with the task of identifying the appropriate fusion technique. In some cases, the nodes may send raw measurement data to the fusion center that then applies appropriate data fusion techniques to formulate the final system decision. Data/decision fusion techniques for the classical distributed detection framework have been extensively covered in [1] and the references therein. Alternatively, techniques of data/decision fusion can be used in other contexts of signal processing that involve decision formulation; for e.g., let H be a state present in a noisy observation and $f_i, i \in \{1, \ldots, N\}$ be N features that ascertain its presence. Now the actual values of the N features or N decisions based on feature values can be fused by relevant data/decision fusion techniques to determine the presence of state H in the noisy observation.

In general, a large variety of distributed detection architecture can be conceived for different configurations and topologies in order to realize desired system-level performance objectives ([1] and the references therein).

1.2 Distributed Estimation

In many WSN applications, WSN nodes attempt to reconstruct a physical phenomenon or estimate a parameter based on their measurements. These measurements are locally processed and are communicated with other sensors or a fusion center before the final reconstruction or estimation. Design of such distributed estimation techniques pose greater challenges as compare to traditional centralized estimation for a multitude of reasons -

1. WSN nodes are distributed over a large geographical area and thus estimation using WSN require participating sensors to perform processing of local information and communicate this information with the other sensors or a fusion center. This introduces the complexity of wireless communication and networking to the problem of distributed estimation in wireless sensor networks, that is otherwise absent in traditional estimation problems.

2. WSN nodes are highly resource constrained -

   - They operate on limited battery power that needs to be optimally utilized in order to prolong the functional lifetime of the network. This strict power constraint stems from cost/size limitations and more often the fact that sensors are deployed in areas that are inaccessible, rendering them non-rechargeable.
   - Additionally, WSN nodes have only limited bandwidth available for communication with each other or with a remote fusion center. Subject to these constraints, each sensor may transmit only a quantized version of the actual measurements to each other or to the fusion center.

   There exists an ineludible trade-off exists between measurement/estimation accuracy attained and total amount of resource consumed by a WSN.

3. Sensor nodes are generally equipped with low-performance and low-memory processors and these factors result in corresponding constraints on computational speed, and memory. This not only makes low-complexity algorithms (with fast convergence rates) a desirable feature in WSNs, but makes tradeoff between resource efficiency, performance, and implementation complexity a critical aspect in WSN algorithm and protocol design.

4. Obtaining complete knowledge of signal (data and noise) models of all WSN nodes is impractical in many cases due to the dynamic nature of the sensing environment. In such cases estimation algorithms and techniques (regarded optimal in the context of centralized estimation) cannot be directly applied.

All these factors compel the need for a new paradigm in distributed and collaborative signal processing - Distributed estimation in wireless sensor networks that enables to achieve a desirable and efficient tradeoff between resource utilization, estimation accuracy or performance efficiency and implementation complexity.

1.3 Prior Work and Motivation

The problem of distributed estimation is a well investigated topic. [3, 4, 5] are some of the early works that studied distributed estimation in the context of spatially distributed observers, sensors or processors based on a linear measurement model and that the joint distribution of the measurements is known. [6] generalized distributed estimation to the case of nonlinear observation models under the assumption that joint distribution of the measurements is known. In all these works, the sensors communicate real values of their measurements to the central location (fusion center) with zero distortion where the final estimate is generated. [7, 8, 9] are some of the important works that first intertwined estimation and quantization by considering the problem of designing efficient distributed estimators where the information is first digitized

using joint distribution of measurement data and then communicated over noiseless channel links. Later, [10] considered the design of estimator for networks with communication constraints and unknown measurement statistics based on the use of a training sequence to realize optimal quantization. [11] investigates sequential signal encoding for distributed estimation for networks with power and delay constraints. Distributed estimator and quantizer design is studied in [12] that accounts for the spatial correlation among sensor measurements. [13] proposes a class of maximum likelihood estimators (MLEs) that achieves a variance close to the clairvoyant estimator when the observations are quantized to one bit. In [14], a universal decentralized estimator is proposed that is based on the rules of linearity and unbiasedness (BLUE) without the knowledge of measurement noise distribution. The premise adopted in all the above mentioned works assumes *distortion-less communication of sensor observations to the fusion center.*

In [15], the best linear unbiased estimator is used that considers the effect of channel noise on the variance of the estimator. Also in [15], an upper bound for the variance for the estimator is derived based on which an energy and rate-efficient estimation scheme is formulated. In [16], a rate-efficient distributed estimation scheme is proposed based on the upper bound of the BLUE variance from [15]. In [19], a tradeoff between the number of active sensors and bit-rate that minimizes estimation error is analyzed. Using the same formulation, [20] investigates the tradeoff between energy used by each sensor and number of active sensors. In [21], the concept of function based network lifetime is introduced and optimized for distributed estimation in order to achieve a particular estimation accuracy at the fusion center. The formulations in [19], [21] assume distortion-free communication links to the fusion center from the sensors. Optimum energy allocation and number of quantization bits in WSN to minimize estimation error in a binary symmetric channel with cross-over probabilities is analyzed in [23]. The work investigates optimal actions (power level

per bit and quantization bits in an information transmission) of the participating sensors that minimizes estimation error at the fusion center for various strategies namely - Optimal power allocation for fixed quantization bits, optimal quantization bits for fixed power per bit and the joint case of power per bit and quantization bits optimization. However, the effect of channel conditions on optimal sensor actions is not analyzed.

Distributed BLUEs that have been used in prior works in the context of wireless sensor networks incorporate either only measurement noise variance or measurement noise and quantization noise variance. To the best of our knowledge, [17] is one work that employs a generalized version of the best linear unbiased estimator in [14] and [15]; i.e., the variance of noise in observation, quantization and channel is incorporated into the design of the estimator. Here the BLUE is used locally at all sensors in order to determine optimal sensor locations and implement a decentralized motion-planning algorithm. The impact of imperfect channels incorporated in this estimator follows the model from [18] that investigates the average effect of channel fading on the performance of a mobile sensor node.

Moreover, the problem of distributed estimation has been addressed for a snapshot of the system without accounting for the history of utilization of each sensor node or the aspect of fairness in sensor scheduling that eventually affects its residual battery power. Note that intuitively the residual battery power can be used to characterize the operating state or the health of a sensor node. In this context, we are motivated by the notion that forcing optimum actions of the active sensors in the network to depend on their residual battery power, characterizing their operating state (or health), is essential for prolonging network lifetime and realizing a fair power, rate allocation.

Based on this concept, we investigate and develop distributed estimation techniques in wireless sensor networks with resource constraints. Unlike all prior

efforts in distributed estimation in resource constrained WSNs, we consider a distributed BLUE that captures the effects of noise in measurement, uniform quantization, and channel. It maybe highlighted that the distributed BLUE used in the case of imperfect channel links is optimal only when the estimator is weighted using weights that depend on the variance in measurement, quantization and channel. We analyze the tradeoff between estimation error at the fusion center, resource utilization of the sensors to achieve that accuracy and implementation complexity to achieve that accuracy. This forms the central theme of this thesis. Unlike all works in the past, we develop insights into different optimization formulations based on the tradeoff analysis between estimation error and resource utilization. Prior efforts have primarily focussed on the dependency of channel conditions, quantization noise and measurement noise on optimal sensors actions for distributed estimation. In this thesis, we account for the sensor operating states (characterized by residual battery power), which is also an important consideration in WSN design - Optimum actions of a sensor with low residual battery power are expected to be different from that with high residual battery power under similar conditions. This forms another novel part in our approach. Using these constructs, we seek to determine optimal sensor actions and their collaborative behavior, and understand their dependencies on various factors like measurement noise, channel conditions and operating state.

1.4 Contributions

The key contributions of this thesis are summarized in this section. From chapters 2 through to 6, we present our work in the area of distributed BLUE in WSN and our contributions can be summarized as follows -

- We study the use of three Best Linear Unbiased Estimators for distributed estimation in an energy and rate-constrained wireless sensor network namely -

- We consider the best linear unbiased estimator (BLUE-1) that accounts for the variance of noise in measurement, uniform quantization and channel. We derive the estimator (BLUE-1) variance and its lower bound [26] in section 2.3. We observe that the lower bound is tight when the participating sensors have comparable channel variances and depends on the sensor with the best channel conditions.
- We consider the best linear unbiased estimator (BLUE-2) that accounts for the variance of noise in measurement and uniform quantization. We derive lower and upper bounds for estimator (BLUE-2) variance [27] in section 2.4. We observe that both upper and lower bounds are tight as long as the channel noise variances of the participating sensors are comparable to each other and depend on the sensors with the worst and best channels respectively.
- We consider the best linear unbiased estimator (BLUE-3) in [15] that incorporates the effects of probabilistic quantization noise and measurement noise. We derive an upper bound for its variance [27] in section 2.5. We see that our bound is tighter than the bound in [15] for low measurement noise variances of the participating sensors.

For all the three estimators investigated, bounds are derived for any modulation scheme employed by the sensor nodes to communicate with each other or the fusion center in general. We present representative results based on BPSK and QAM modulation schemes.

- Unlike all prior efforts, we use BLUE-1 for distributed estimation in resource constrained WSN and study the optimal tradeoff between overall estimation accuracy at the fusion center and resource utilization of the sensors using three different system design approaches (or optimization formulations) [28] -

  - Formulation A - Minimize estimation error (BLUE variance) at the fusion center subject to a total system resource utilization constraint (rate and power) in section 3.2.
  - Formulation B - Minimize total system resource utilization subject to a constraint on the the estimation error in section 4.1.
  - Formulation C - Minimize the system resource utilization and estimation error jointly at the fusion center [25] as in section 5.1.

- We adopt a novel approach in optimization formulation by accounting for the operating state or the residual battery power of the sensors in the network. This is an important consideration in WSN design as optimum actions (power level per bit and quantization bits in an information transmission) of a sensor with low residual battery power are expected to be different from that with high residual battery power under similar conditions.

- We formulate energy and rate efficient schemes that enable optimal operation of the sensor nodes supposing that the sensors communicate with the fusion center over noisy channel links using two modulation schemes namely - BPSK and QAM.

- We counter non-convexity imbued in the three optimization problems framed, by applying techniques to transform it into a Difference of Convex Functions (D.C.) problem [22],[24]. Further, we approximate all the three D.C. formulations as convex optimization problems by applying first-order Taylor expansion. We observe that the solutions obtained from the D.C. problem version is the same as in the convex approximated version in all the three formulations. i.e., that the non-convex problem in all the three original formulations can be approximated as appropriate convex problems without affecting optimality.

- We study the relation between optimal actions of the sensors and channel conditions, operating states, and error in measurement. In effect, we see that the amount of error in estimation at the fusion center depends on the nodes' measurement quality, but also its operating state and local channel conditions.

- We develop insights into optimization formulations by performing a comparative analysis between the three optimization formulations in terms of the tradeoff between estimation error and resource used, optimal sensor actions, and collaborative behavior. We observe that Formulation B is the most economical approach in terms of resource consumed for a target BLUE variance and Formulation A enables achieve high quality estimator but at the cost of excess amount of resources.

In the ensuing chapters from 7 to 9, we present a problem in distributed detection and signal processing of PPG data obtained from pulse-oximeter sensors [29]. More specifically, the contributions are as follows -

- We formulate a motion artifact detection scheme in PPG data [30] using higher order statistics (HOS) properties of clean and motion-corrupted PPG data - In the time domain, we use skew and kurtosis measures associated with the data to aid detection. In the frequency domain, the presence of random components due to motion artifact is ascertained using a frequency-domain kurtosis measure as in [48]. Also, bispectral analyses of PPG data indicate the presence of strong quadratic phase coupling (QPC) and more specifically self coupling in the case of clean PPG data. In motion-artifact-corrupted data, QPC between random frequency components is observed, but the self coupling feature is absent. To the best of our knowledge, this is the first effort to employ HOS analysis for motion detection in PPG data.

- We formulate Neyman Pearson (NP) tests based on these time-domain and

frequency-domain metrics. Using practical test data, we characterize the performance (probability of false alarm - P_F, probability of detection - P_D, and probability of error - P_{error}) of the artifact detection tests. The performance results illustrate the potency of the proposed method for consistent and robust detection of PPG motion artifact.

- Treating each of the measures as observations from independent sensors, we perform soft decision fusion from [33] and hard-fusion (Varshney-Chair rule) from [32] to fuse individual decisions to form a global system decision.

- In chapter 9, we present a new motion artifact reduction method [31] that combines an enhanced signal preprocessing unit and a frequency-domain ICA unit.

- We propose a new enhanced preprocessing unit incorporates a Fourier series reconstruction of the PPG data that utilizes the spectrum variability and quasi-periodicity of the pulse waveform.

- We develop a novel frequency-domain ICA routine (FD-ICA) that considers only magnitude information is presented. This technique assumes instantaneous mixing of statistically independent sources in the time domain and a constant mixing matrix for the time frame considered. A comparison of the technique used in this thesis with the time-domain ICA and complex FD-ICA techniques in the literature implies that the new magnitude-based frequency domain ICA approach more effectively reduces motion artifact.

Finally in Appendix A, we present an analysis of a wireless ad hoc/sensor network where nodes are connected via random channels and information is transported in the network in a cooperative multihop fashion using amplify and forward relay strategy [55]. The contributions are summarized as follows -

- To the best of our knowledge, this is the first work that attempts to analyze and characterize such a random network using important parameters like: (1) SNR degradation with hop, (2) outage probability, (3) maximum permissible number of hops and (4) source-destination node pairs that communicate with each other simultaneously.

- Using constructs from graph theory, we formulate the basic operation of the network by demonstrating a scheme for choosing appropriate nodes for relaying information over disjoint routes between all source-destination node pairs. We establish the condition of existence of such disjoint paths between all source and destination nodes in the network and their characteristics.

- We evaluate the achievable throughput of the network and its asymptotic scaling for channel strengths drawn from an exponential density and observe that the throughput scales asymptotically as $O(\log n)$, where n is the number of nodes in the network.

1.5 Organization

The thesis is organized as follows - In Chapter 2, the concept of BLUE for centralized and distributed estimation are introduced. Three BLUE designs and bounds on their variances are investigated. From Chapter 3 through to 5, the three optimization formulations along with their analysis and the results are presented. In Chapter 6, we first present a comparison between the three formulations in terms of estimation error achieved and the resources utilized to achieve that error. In the same chapter, we discuss the optimal actions of sensors and their collaborative behavior observed in each of the formulations. In Chapter 7, we propose a model for the two-stage approach in detection and reduction of motion artifacts in photoplethysmographic data. Chapters 8 and 9 details the PPG data analysis, motion artifact detection

and its reduction. In Chapter 10, we present our key contributions, and plausible future directions and extensions to our work in distributed estimation in WSNs. Finally in Appendix A, we present a throughput analysis of a wireless ad hoc/sensor network. We consider a network where nodes are connected via random channels and information is transported in the network in a cooperative multihop fashion using amplify and forward relay strategy.

Chapter 2

Best Linear Unbiased Estimator (BLUE)

In this chapter, we study the use of the best linear unbiased estimator (BLUE) for distributed parameter estimation in wireless sensor networks. We investigate three types of BLUE - (1) Best Linear Unbiased Estimator (BLUE-1) that accounts for the variance of noise in measurement, uniform quantization and channel; (2) Best Linear Unbiased Estimator (BLUE-2) that accounts for the variance of noise in measurement and uniform quantization; (3) Best Linear Unbiased Estimator (BLUE-3) that incorporates the effects of probabilistic quantization noise and measurement noise. For all three estimators, we derive bounds for their variance considering any modulation scheme in general, and specifically for BPSK and QAM modulation schemes employed by the sensor nodes to communicate with each other or the fusion center.

2.1 Definition - BLUE

In many practical scenarios related to parameter estimation, the design of an optimal minimum variance estimator may not be possible more often, though not always,

due to the lack of knowledge of the probability density function (pdf) of the data set. Such cases san the use of CRLB or sufficient statistics making it reasonable to consider suboptimal estimators [2]. One such consideration is treating the estimator to be linear in data. i.e., if $y_1, y_2, \ldots, y_N$ is the data set under consideration that has a pdf denoted by $p(y_1, y_2, \ldots, y_N, \Theta)$ where, Θ is an unknown parameter, then the estimator that is linear in the data is given as -

$$\hat{\Theta} = \sum_{i=1}^{N} a_i y_i, \tag{2.1}$$

where, $a_i, \forall i \in \{1, \ldots, N\}$ are determined such that the variance of $\hat{\Theta}$ is minimized. It may be noted that the BLUE requires the knowledge of the first and second moments of the pdf of the data set $y_i, \forall i \in \{1, \ldots, N\}$. If the data set is uncorrelated and has zero mean with variance $\sigma_i^2, \forall i \in \{1, \ldots, N\}$, then the BLUE $\hat{\Theta}$ is as follows -

$$\hat{\Theta} = \left(\sum_{i=1}^{N} \frac{1}{\sigma_i^2} \right)^{-1} \sum_{i=1}^{N} \frac{y_i}{\sigma_i^2} \tag{2.2}$$

Apparently in (2.1), the samples with the smallest variances are weighted most heavily. The variance of $\hat{\Theta}$ is given as -

$$Var(\hat{\Theta}) = \left(\sum_{i=1}^{N} \frac{1}{\sigma_i^2} \right)^{-1} \tag{2.3}$$

Additionally, it may be highlighted that if the pdf of the data is Gaussian, then the BLUE is also the minimum variance unbiased estimator (MVUE).

2.2 Distributed Best Linear Unbiased Estimator

Consider a WSN consisting of N distributed sensors (observers) that measure a source signal θ and report the observations to a fusion center. The observation of the $i-$th sensor, $\forall i \in \{1, \ldots, N\}$ is the following linear model -

$$x_i \;\; = \;\; \theta + n_i, \tag{2.4}$$

where, n_i is the $i-$th sensor's measurement noise, with zero mean, and spatially uncorrelated with variance σ_i^2. The sensor measurement noise distribution is otherwise considered to be unknown. After some local processing of these measurements at the sensor nodes, they are transmitted to a central location or a fusion center without distortion where they are fused to produce a final estimate of θ using a fusion function f. If the fusion center has complete knowledge of the sensor measurement noise variances, it suffices to linearly combine all sensor observations $x_i \; \forall i \in \{1, \ldots, N\}$ to form an unbiased estimate of θ with the minimum variance as follows -

$$\begin{aligned} f(x_1, \ldots, x_N) = \hat{\theta} &= \left(\sum_{i=1}^{N} \frac{1}{\mathbb{E}(x_i - \theta)^2} \right)^{-1} \sum_{i=1}^{N} \frac{x_i}{\mathbb{E}(x_i - \theta)^2} \\ &= \left(\sum_{i=1}^{N} \frac{1}{\sigma_i^2} \right)^{-1} \sum_{i=1}^{N} \frac{x_i}{\sigma_i^2}, \end{aligned} \tag{2.5}$$

where, $\mathbb{E}(.)$ denotes the expectation operator. The estimate $\hat{\theta}$ has a mean-squared error (MSE) denoted by D, that also gives a measure of the quality of the final estimate generated, given as -

$$\begin{aligned} D &= \mathbb{E} \left[\left(\sum_{i=1}^{N} \frac{1}{\mathbb{E}(x_i - \theta)^2} \right)^{-1} \sum_{i=1}^{N} \frac{x_i - \theta}{\mathbb{E}(x_i - \theta)^2} \right]^2 \\ &= \left(\sum_{i=1}^{N} \frac{1}{\sigma_i^2} \right)^{-1} \end{aligned} \tag{2.6}$$

This is the notion of distributed best linear unbiased estimator or distributed BLUE. It may be noted that if the sensor nodes communicate their real-valued measurements to the fusion center without performing any local processing, then the model under consideration is called centralized BLUE. In a realistic wireless sensor network, transmission of real-valued measurements incur high communication cost in terms of bandwidth and power expended for transmission. Moreover, the channel links between the fusion center and the sensor nodes are noisy and subject to loss due to fade. Thus, it requires that the sensors locally process their measurements by means of quantization and transmit digitized information over these noisy, fading channel

links. This aspect essentially interlaces quantization and estimation in the context of wireless sensor networks.

2.3 BLUE-1

In this section, we study a BLUE for distributed estimation that considers the effect of observation noise, uniform quantization noise, and channel noise. By accounting for instantaneous channel knowledge, we derive the estimator variance and its lower bound for any modulation scheme, and more specifically for BPSK and QAM modulation schemes. We analyze the performance of the lower bound by drawing a comparison with the actual variance of the estimator.

2.3.1 Problem Formulation

For the WSN model described earlier in the chapter, we assume that each sensor locally performs uniform quantization of its observation x_i, $\forall i \in \{1, \ldots, N\}$ as follows -

$$x_{i,q} \quad = \quad x_i + n_{i,q}, \tag{2.7}$$

where, $n_{i,q}$ is the quantization noise of the $i-$th sensor, $\forall i \in \{1, \ldots, N\}$. The quantized information is transmitted by all N sensors to the fusion center over independent AWGN channels (realized by means of orthogonal signaling). Information received at the fusion center from the $i-$th sensor is given as -

$$x_{i,c} \quad = \quad x_{i,q} + n_{i,c}, \tag{2.8}$$

where, $n_{i,c}$ is the noise due to imperfect channel experienced by the $i-$th sensor, $\forall i \in \{1, \ldots, N\}$. A detailed system model is illustrated in figure 2.1.

The fusion center linearly combines all sensor observations $x_{i,c}$ $\forall i \in \{1, \ldots, N\}$ (laden with measurement, quantization, and channel noises) using the best unbiased

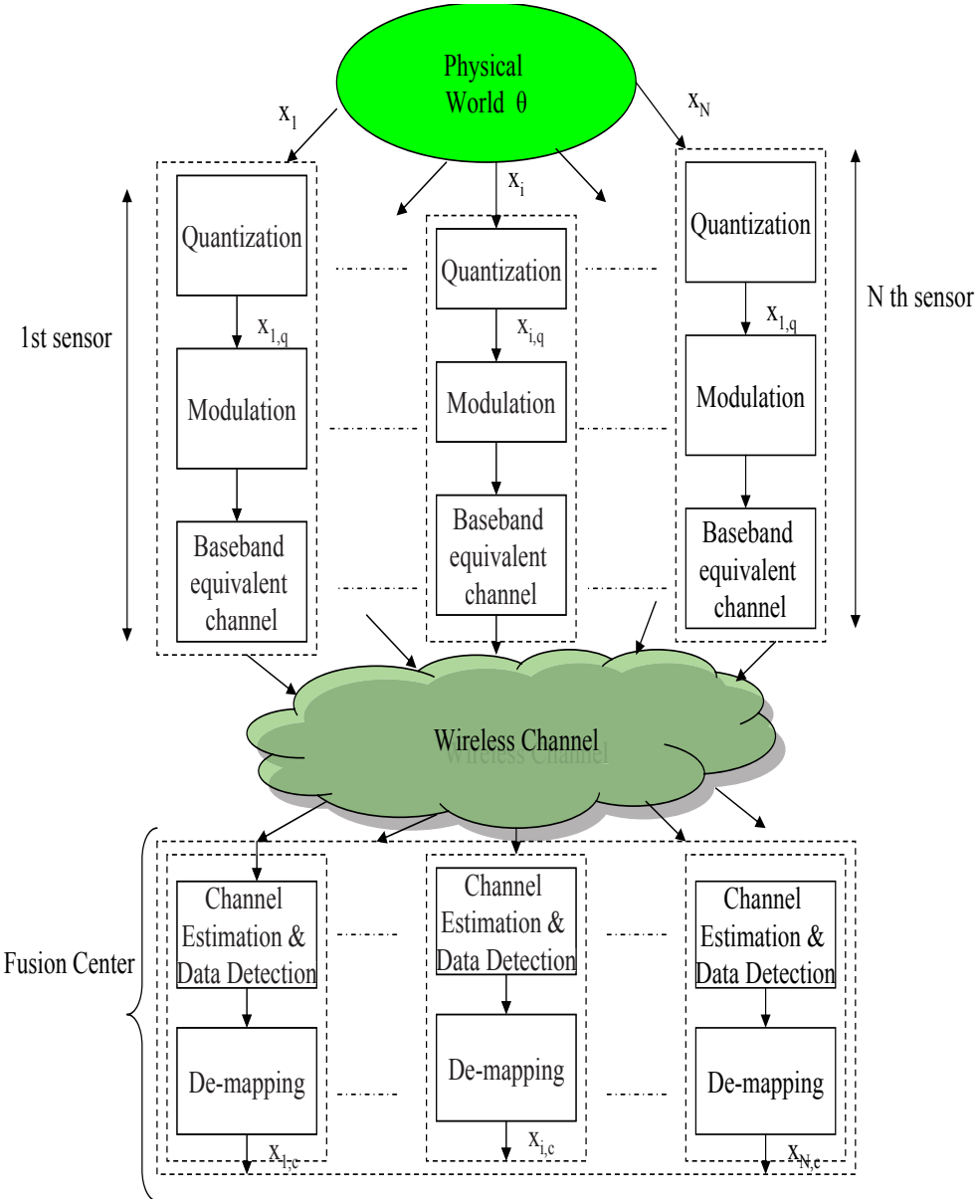


Figure 2.1: Detailed System Model

linear estimator to form an estimate of θ. More specifically, the estimator at the fusion center weighs the information from each sensor linearly with its variance that depends on its measurement, quantization, and channel noises and is given as -

$$\hat{\theta} = \left(\sum_{i=1}^{N} \frac{1}{\mathbb{E}(x_{i,c} - \theta)^2} \right)^{-1} \sum_{i=1}^{N} \frac{x_{i,c}}{\mathbb{E}(x_{i,c} - \theta)^2}, \tag{2.9}$$

We assume that the fusion center has complete knowledge of the variance associated

with the information received from each sensor. The unbiased estimate $\hat{\theta}$ has a mean-squared error (MSE) given by -

$$\begin{aligned} D &= \mathbb{E}\left[\left(\sum_{i=1}^{N}\frac{1}{\mathbb{E}(x_{i,c}-\theta)^2}\right)^{-1}\sum_{i=1}^{N}\frac{x_{i,c}-\theta}{\mathbb{E}(x_{i,c}-\theta)^2}\right]^2 \\ &= \left(\sum_{i=1}^{N}\frac{1}{\mathbb{E}(x_{i,c}-\theta)^2}\right)^{-2}\sum_{i=1}^{N}\frac{\mathbb{E}(x_{i,c}-\theta)^2}{(\mathbb{E}(x_{i,c}-\theta)^2)^2} \\ &= \left(\sum_{i=1}^{N}\frac{1}{\mathbb{E}(x_{i,c}-\theta)^2}\right)^{-1} \end{aligned} \tag{2.10}$$

Let $R_i = \mathbb{E}(n_i^2)$, $R_{i,q} = \mathbb{E}(n_{i,q}^2)$, $R_{i,c} = \mathbb{E}(n_{i,c}^2)$, $\forall i \in \{1,\ldots,N\}$. It can be seen that $\mathbb{E}(x_{i,c}-\theta)^2 = R_i + R_{i,q} + R_{i,c}$, $\forall i \in \{1,\ldots,N\}$, since measurement, quantization, and channel noises can be considered to be statistically independent of each other. If $[-W,W]$ represents the dynamic range of the signal source, then $R_{i,q} = \frac{W^2}{3(2^{l_i}-1)^2}$, where $l_i \in [1, BW]$ is the number of quantization bits used in a transmission by the i-th sensor and BW denotes the total rate constraint of the system.

If the $i-$th sensor communicates with the fusion center using a particular modulation scheme that results in a bit error probability of $P_{i,k}^{\{b\}}$ for the $k-$th bit in the transmitted information, then the variance due to imperfect channel can be derived as follows -

$$n_{i,c} = \begin{cases} \pm 2^k \times \Delta_i & P_{i,k}^{\{b\}} \\ 0 & 1 - \sum_{k=0}^{l_i-1} P_{i,k}^{\{b\}}, \end{cases} \tag{2.11}$$

where, $\Delta_i = \frac{2W}{2^{l_i}-1}$ is the quantizer step size. (2.11) assumes that there is at most one bit in error in each information transmission consisting of l_i bits. Also, we assume that all the bits in the transmitted information have the same bit error probability associated with it (i.e. $P_{i,k}^{\{b\}} = P_i^{\{b\}}$). Under the assumption that channel noise variance remains unchanged during a complete information transmission, the variance

contribution from the channel is -

$$\begin{aligned} R_{i,c} &= \sum_{k=0}^{l_i-1} (\pm 2^k \times \Delta_i)^2 P_{i,k}^{\{b\}}, \\ &= P_i^{\{b\}} \times \Delta_i^2 \sum_{k=0}^{l_i-1} 4^k \\ &\approx \frac{4W^2}{3} P_i^{\{b\}} \end{aligned} \tag{2.12}$$

If the $i-$th sensor uses BPSK modulation scheme for transmission, then equation (2.12) becomes -

$$\frac{4W^2}{3} P_i^{\{b\}} = \frac{4W^2}{3} \mathcal{Q}\sqrt{SNR_i}, \tag{2.13}$$

where, SNR_i is the signal to noise ratio associated with each bit transmission. Suppose that the $i-$th sensor chooses uncoded QAM modulation for transmission of l_i bits such that b_i is the size of each symbol transmitted. Let $c_i \in \mathbb{Z}^+, \forall i \in \{1, \ldots, N\}$ denote the number of symbols transmitted. Then the variance due to imperfect channel associated with the complete information is given as -

$$\frac{4W^2}{3} P_i^{\{b\}} = \frac{4W^2 P_i^{\{s\}} l_i}{3b_i^2}, \tag{2.14}$$

where, $P_i^{\{s\}}$ is the symbol error probability associated with the $i-$th sensor's information transmission.

2.3.2 Analytical Results

The BLUE variance in equation (3.5) corresponds to -

$$D = \left(\sum_{i=1}^{N} \frac{1}{(R_i + R_{i,q} + R_{i,c})} \right)^{-1} \tag{2.15}$$

Lemma 1: The variance of the best linear unbiased estimator in equation (2.15), where the information is weighed by its measurement noise, quantization noise, and channel noise variances, is lower bounded as follows -

$$D \geq \frac{R_c^{\{min\}}}{N} + \left(\sum_{i=1}^{N} \frac{1}{(R_i + R_{i,q})} \right)^{-1}, \tag{2.16}$$

where, $R_c^{\{min\}}$ corresponds to the minimum channel noise variance such that $R_c^{\{min\}} = \min(R_{1,c}, R_{2,c}, \ldots, R_{N,c})$.

Proof: Define function $H(.)$ such that -

$$\begin{aligned} & H(R_1 + R_{1,q} + R_{1,c}, \ldots, R_N + R_{N,q} + R_{N,c}) \\ & = \left(\frac{1}{N} \sum_{i=1}^{N} \frac{1}{(R_i + R_{i,q} + R_{i,c})} \right)^{-1} \end{aligned} \tag{2.17}$$

Then we have -

$$\begin{aligned} & H(R_1 + R_{1,q} + R_{1,c}, \ldots, R_N + R_{N,q} + R_{N,c}) \\ & \geq H(R_1 + R_{1,q} + R_c^{\{min\}}, \ldots, R_N + R_{N,q} + R_c^{\{min\}}) \end{aligned}$$

For convenience we set $R_c^{\{min\}} = K$. Now we have -

$$\begin{aligned} NH^{-1} &= \sum_{i=1}^{N} \frac{1}{R_i + R_{i,q} + K} \\ \frac{d(NH^{-1})}{dK} &= -NH^{-2} \frac{dH}{dK} \\ &= -\sum_{i=1}^{N} \frac{1}{(R_i + R_{i,q} + K)^2} \\ \Rightarrow \frac{dH}{dK} &= \frac{d(NH^{-1})}{dK} / \left(-NH^{-2}\right) \end{aligned} \tag{2.18}$$

NH^{-2} can be written from (2.17) as -

$$NH^{-2} = \frac{1}{N} \left(\sum_{i=1}^{N} \frac{1}{R_i + R_{i,q} + K} \right)^2 \tag{2.19}$$

We replace all the cross-product terms in the right hand side of equation (2.19) by applying the Cauchy-Schwarz inequality $2.\frac{1}{R_i+R_{i,q}+K} \cdot \frac{1}{R_j+R_{j,q}+K} \leq \frac{1}{(R_i+R_{i,q}+K)^2} + \frac{1}{(R_j+R_{j,q}+K)^2}$. Therefore, we have-

$$\sum_{i=1}^{N} \frac{1}{(R_i + R_{i,q} + K)^2} \geq \frac{1}{N} \left(\sum_{i=1}^{N} \frac{1}{R_i + R_{i,q} + K} \right)^2 \tag{2.20}$$

Upon substituting (2.20) in (2.18), we can see that $\frac{dH}{dK}$ is always greater or equal to unity. Now applying the mean value theorem in the interval $[0, K]$ on the function

$H(.)$, we have -

$$\frac{H(R_1 + R_{1,q} + K, \ldots) - H(R_1 + R_{1,q}, \ldots)}{K} = \frac{dH}{dK} \tag{2.21}$$

From (2.21) and (2.18) we have $H(R_1 + R_{1,q} + K, \ldots) \geq H(R_1 + R_{1,q}, \ldots) + K$. And substituting from (2.18) we have -

$$\begin{aligned} D &\geq \frac{H(R_1 + R_{1,q}, \ldots, R_N + R_{N,q})}{N} + \frac{K}{N} \\ \Rightarrow D &\geq \left(\sum_{i=1}^{N} \frac{1}{(R_i + R_{i,q})} \right)^{-1} + \frac{K}{N} = \\ & \left(\sum_{i=1}^{N} \frac{1}{(R_i + R_{i,q})} \right)^{-1} + \frac{R_c^{\{min\}}}{N} \end{aligned} \tag{2.22}$$

Lemma 1.1: When the participating sensors use BPSK modulation scheme for transmission, then estimation variance is lower bounded as follows -

$$D \geq \left(\sum_{i=1}^{N} \frac{1}{(R_i + R_{i,q})} \right)^{-1} + \frac{4W^2}{3N} \mathcal{Q} \sqrt{SNR_{max}}, \tag{2.23}$$

where, $SNR_{max} = \max\{SNR_i\}, \forall i \in \{1, \ldots, N\}$.

Lemma 1.2: When the participating sensors choose uncoded QAM for transmission of information, the estimation variance is lower bounded as follows -

$$D \geq \left(\sum_{i=1}^{N} \frac{1}{(R_i + R_{i,q})} \right)^{-1} + (\frac{4W^2 P^{\{s\}} l}{3Nb^2})_{min}, \tag{2.24}$$

where, $(\frac{4W^2 P^{\{s\}} l}{3Nb^2})_{min} = \min(\frac{4W^2 P_1^{\{s\}} l_1}{3Nb_1^2}, \ldots, \frac{4W^2 P_N^{\{s\}} l_N}{3Nb_N^2})$. It maybe noted that the lower bound for the variance of this estimator differs from what is achievable with perfect sensor channels by an additive factor.

2.3.3 Results and Discussion

We perform Monte-Carlo simulations for evaluating the performance of the best linear unbiased estimator discussed in the previous section. We consider a wireless sensor

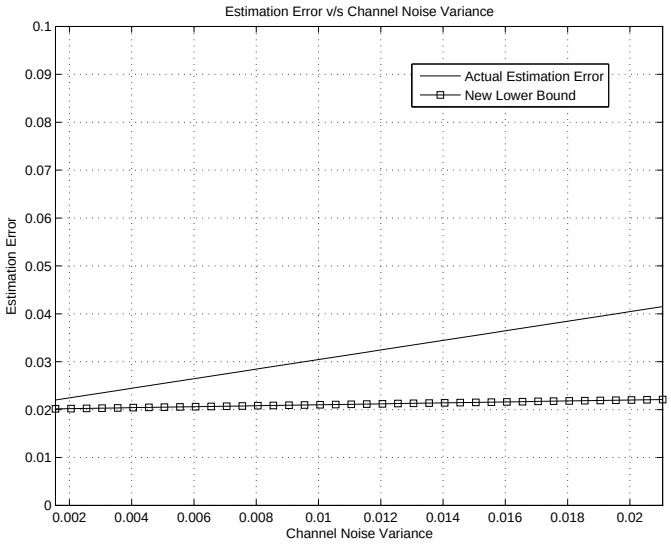


Figure 2.2: Variation of Estimator Variance with Channel Noise Variance

network with $N = 20$ sensors. We set the dynamic range of the source signal as $W \in [-1, 1]$ and consider BPSK modulation scheme. By simulations, we draw a comparison between the actual estimation error and the bounds derived. We plot the estimation variance against the channel noise variance (with the channel noise variance of only one of the sensors being varied). As expected from equation (2.18), we observe that the lower bound depends on the sensor with the best channel conditions. In figure 2.2, we see that the lower bound is tight when the channel noise variances of the sensors are comparable, and the deviation becomes prominent as the noise variances greatly vary from each other.

2.4 BLUE-2

Next, we consider a BLUE that accounts for variance of the noise in observation and quantization for the design of the estimator. We assume that each of the participating sensors uniformly quantizes its measurement and transmits over AWGN channels to a fusion center where the final estimate is generated. We derive lower and upper

bounds for the variance of this estimator for any modulation scheme in general. We then analyze the performance of the bounds by drawing a comparison with the actual estimation error.

2.4.1 Problem Formulation

Using the WSN model as before, we suppose that each sensor locally quantizes its observation x_i, $\forall i \in \{1, \ldots, N\}$ uniformly as follows -

$$x_{i,q} \;=\; x_i + n_{i,q}, \tag{2.25}$$

where, $n_{i,q}$ is the quantization noise of the $i-$th sensor, $\forall i \in \{1, \ldots, N\}$. The quantized information is transmitted by all N sensors to a fusion center over independent AWGN channels that are realized by means of orthogonal signaling schemes like TDMA, FDMA or CDMA. Information received at the fusion center from the $i-$th sensor is given as -

$$x_{i,c} \;=\; x_{i,q} + n_{i,c}, \tag{2.26}$$

where, $n_{i,c}$ is the noise due to imperfect channel experienced by the $i-$th sensor, $\forall i \in \{1, \ldots, N\}$. By weighing the information $x_{i,c}$ $\forall i \in \{1, \ldots, N\}$ from each sensor with its variance that depends on its measurement, and quantization noises, the fusion center forms the best unbiased linear estimate of θ. This is given as follows -

$$\hat{\theta} \;=\; \left(\sum_{i=1}^{N} \frac{1}{\mathbb{E}(x_{i,q} - \theta)^2} \right)^{-1} \sum_{i=1}^{N} \frac{x_{i,c}}{\mathbb{E}(x_{i,q} - \theta)^2}, \tag{2.27}$$

The unbiased estimate $\hat{\theta}$ has a mean-squared error (MSE) given by -

$$\begin{aligned} D \;&=\; \mathbb{E}\left[\left(\sum_{i=1}^{N} \frac{1}{\mathbb{E}(x_{i,q} - \theta)^2} \right)^{-1} \sum_{i=1}^{N} \frac{x_{i,c} - \theta}{\mathbb{E}(x_{i,q} - \theta)^2} \right]^2 \\ &=\; \left(\sum_{i=1}^{N} \frac{1}{\mathbb{E}(x_{i,q} - \theta)^2} \right)^{-2} \sum_{i=1}^{N} \frac{\mathbb{E}(x_{i,c} - \theta)^2}{(\mathbb{E}(x_{i,q} - \theta)^2)^2} \end{aligned}$$

Let $R_i = \mathbb{E}(n_i)^2) = \sigma_i^2$, $R_{i,q} = \mathbb{E}(n_{i,q}^2)$, $R_{i,c} = \mathbb{E}(n_{i,c}^2)$, $\forall i \in \{1, \ldots, N\}$ and we consider measurement, quantization, and channel noises to be statistically independent of each other. Hence the variance of this estimator is given as -

$$\begin{aligned} D &= \mathbb{E}\left[\left(\sum_{i=1}^{N} \frac{1}{R_i + R_{i,q}}\right)^{-1} \sum_{i=1}^{N} \frac{x_{i,c} - \theta}{R_i + R_{i,q}}\right]^2 \\ &= \left(\sum_{i=1}^{N} \frac{1}{(R_i + R_{i,q})}\right)^{-2} \sum_{i=1}^{N} \frac{\mathbb{E}(x_{i,c} - \theta)^2}{(R_i + R_{i,q})^2} \end{aligned}$$

In this premise, we assume that the fusion center has complete knowledge of the variance associated with the information received from each sensor. Let $[-W, W]$ denote the dynamic range of the signal source. Then, the uniform quantization noise variance is $R_{i,q} = \frac{W^2}{3(2^{l_i}-1)^2}$, where $l_i \in [1, BW]$ represents the number of quantization bits used by the i-th sensor in a transmission and BW denotes the bandwidth of the system.

From [26], we know that when the $i-$th sensor communicates with the fusion center using a particular modulation scheme that results in a bit error probability of $P_{i,k}^{\{b\}}$ for the $k-$th bit in the transmitted information, the variance due to imperfect channel is -

$$\mathbb{E}(n_{i,c}^2) \approx \frac{4W^2}{3} P_i^{\{b\}} \tag{2.28}$$

In deriving (2.28), we assume that in each information transmission consisting of l_i bits, there is at most one bit in error. Also, we assume that all the bits in the transmitted information have the same bit error probability (i.e. $P_{i,k}^{\{b\}} = P_i^{\{b\}}$) and that during a complete information transmission, the channel condition experienced by a sensor remains unchanged.

Thus, if BPSK modulation scheme for transmission is used by the $i-$th sensor, then (2.28) becomes -

$$\mathbb{E}(n_{i,c}^2) \approx \frac{4W^2}{3} P_i^{\{b\}} = \frac{4W^2}{3} \mathcal{Q}\sqrt{SNR_i}, \tag{2.29}$$

where, SNR_i is the signal to noise ratio associated with each bit transmission.

Now, consider uncoded QAM scheme to be used for the transmission of l_i bits by the $i-$th sensor such that b_i is the size of each symbol transmitted. Let $c_i = \frac{l_i}{b_i}, c_i \in \mathbb{Z}^+, \forall i \in \{1, \ldots, N\}$ denote the number of symbols transmitted. Then the variance due to imperfect channel associated with complete information is given as -

$$\mathbb{E}((n_{i,c})^2) \approx \frac{4W^2}{3} P_i^{\{b\}} = \frac{4W^2 P_i^{\{s\}} l_i}{3b_i^2}, \tag{2.30}$$

where $P_i^{\{s\}}$ is the symbol error probability associated with the $i-$th sensor's information transmission.

2.4.2 Analytical Results

We first derive a lower and upper bound for the variance of the estimator that uses measurement and uniform quantization noise variance for weighing the information from sensors that is transmitted over independent noisy AWGN channels.

Lemma 2: Let D denote the variance associated with the best linear unbiased estimator where the information from the sensors is weighed by weights that depends on its measurement noise and uniform quantization noise variances at the fusion center. Then D is bounded as follows -

$$\begin{aligned} &\left(\sum_{i=1}^{N} \frac{1}{(R_i + R_{i,q})} \right)^{-1} + R_c^{\{min\}} \leq D \\ &\leq \left(\sum_{i=1}^{N} \frac{1}{(R_i + R_{i,q})} \right)^{-1} + R_c^{\{max\}}, \end{aligned} \tag{2.31}$$

where, $R_c^{\{max\}}$ corresponds to the maximum channel noise variance, such that $R_c^{\{max\}} = \max(R_{1,c}, R_{2,c}, \ldots, R_{N,c})$, and $R_c^{\{min\}}$ denotes the minimum channel noise variance, such that $R_c^{\{min\}} = \min(R_{1,c}, R_{2,c}, \ldots, R_{N,c})$.

Proof: The variance of this estimator corresponds to -

$$D = \mathbb{E}\left[\left(\sum_{i=1}^{N}\frac{1}{R_i+R_{i,q}}\right)^{-1}\sum_{i=1}^{N}\frac{x_{i,c}-\theta}{R_i+R_{i,q}}\right]^2$$
$$= \left(\sum_{i=1}^{N}\frac{1}{(R_i+R_{i,q})}\right)^{-2}\sum_{i=1}^{N}\frac{\mathbb{E}(x_{i,c}-\theta)^2}{(R_i+R_{i,q})^2}$$

For convenience, we let $D' = \left(\sum_{i=1}^{N}\frac{1}{(R_i+R_{i,q})}\right)^{-1}$, where D' is the MSE of the estimator at the fusion center in case of perfect sensor channels. After some algebraic manipulations we have -

$$\begin{aligned} D &= D'^2\left(\sum_{i=1}^{N}\frac{R_i+R_{i,q}+R_{i,c}}{(R_i+R_{i,q})^2}\right) \\ &= D'^2\left(\sum_{i=1}^{N}\frac{R_{i,c}}{(R_i+R_{i,q})^2}+\frac{R_i+R_{i,q}}{(R_i+R_{i,q})^2}\right) \\ &= D'^2\left(\sum_{i=1}^{N}\frac{R_{i,c}}{(R_i+R_{i,q})^2}+\frac{1}{(R_i+R_{i,q})}\right) \end{aligned}$$

From the previous section, we have that $R_{i,c} = \frac{4W^2P_{bi}}{3}$. Let $P_{max}^{\{b\}}$ =max$\{P_i^{\{b\}}\}, \forall i \in \{1,\ldots,N\}$ denote the maximum probability of bit error among the N collaborating sensors and $R_c^{\{max\}}$ be the corresponding variance due to the channel. Then we have -

$$\begin{aligned} D &\leq D'^2\left(\sum_{i=1}^{N}\frac{R_c^{\{max\}}}{(R_i+R_{i,q})^2}+\frac{1}{(R_i+R_{i,q})}\right) \\ &= D'+R_c^{\{max\}} = D'+\frac{4W^2P_{max}^{\{b\}}}{3} \end{aligned}$$

In the above treatment, $R_{i,c}$ can be replaced by $R_c^{\{min\}}$ that corresponds to the minimum channel variance (minimum bit error probability) among all sensors in order to obtain a lower bound for the estimator variance. When the sensors use BPSK modulation scheme for transmission of information, we have -

$$D \leq D'+\frac{4W^2}{3}\mathcal{Q}\sqrt{SNR_{min}},$$

where, SNR_{min} =min$\{SNR_i\}, \forall i \in \{1, \ldots, N\}$. For uncoded QAM, we have -

$$D \leq D' + (\frac{4W^2 P^{\{s\}} l}{3b^2})_{max},$$

where, $(\frac{4W^2 P^{\{s\}} l}{3Nb^2})_{max} = \max(\frac{4W^2 P_1^{\{s\}} l_1}{3Nb_1^2}, \ldots, \frac{4W^2 P_N^{\{s\}} l_N}{3Nb_N^2})$ It can be seen that both the lower and upper bound for the variance of this estimator differs from what is achievable with perfect sensor channels by an additive factor.

2.4.3 Results and Discussion

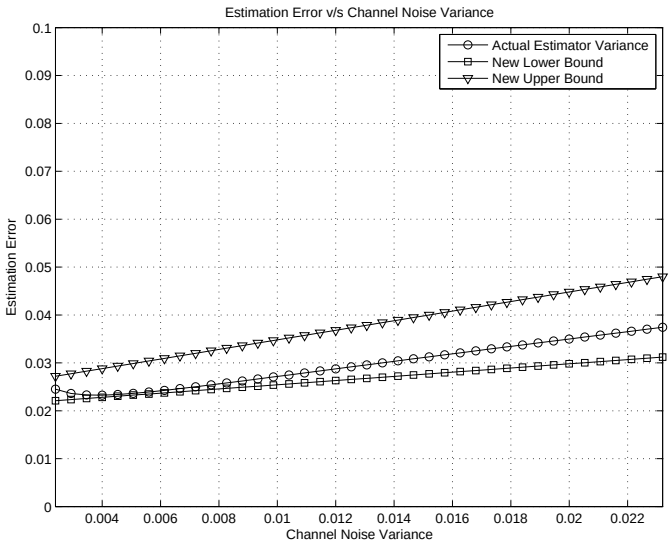


Figure 2.3: Estimator Variance versus Channel Noise Variance for Uniform Quantization.

In this section, we present Monte-Carlo simulation results for the estimation error associated with the best linear unbiased estimators discussed in the previous sections. We draw a comparison between the actual estimation error and the bounds derived as in figure 2.3. We consider a wireless sensor network with $N = 20$ sensors; fix the dynamic range of the source signal as $W \in [-1, 1]$, and consider BPSK modulation scheme. We plot the estimation error against the channel noise variance (and the channel noise variance of only one of the sensors is varied). From the figure, it can

be seen that the lower bound depends on the sensor with the best channel condition and the upper bound depends on the sensor with the worst channel condition.

2.5 BLUE-3

Finally, we investigate a BLUE that incorporates variance of the noise in observation and quantization for the design of the estimator. We assume that the sensors perform uniform probabilistic quantization of its information that is then transmitted over noisy wireless AWGN connections. This is similar to the BLUE in [15] that incorporates the effects of noise in measurement and uniform random quantization. We derive a new upper bound for the variance of this estimator and compare it with the upper bound derived in [15].

2.5.1 Problem Formulation

Based on the same system model as detailed earlier in the chapter, we consider the case where uniform probabilistic quantization as in [15] is performed locally at all sensors. From [15] we have -

$$\begin{aligned} \mathbb{E}(x_{i,q} - \theta)^2 &\leq \sigma_i^2 + \frac{W^2}{(2^{l_i} - 1)^2} \\ &= \sigma_i^2 + \delta_i^2, \end{aligned} \tag{2.32}$$

where, $\delta_i^2 = \frac{W^2}{(2^{l_i}-1)^2}$. The BLUE at the fusion center in the case of perfect sensor channels is given as -

$$\hat{\theta}_q = \left(\sum_{i=1}^{N} \frac{1}{\sigma_i^2 + \delta_i^2} \right)^{-1} \sum_{i=1}^{N} \frac{x_{i,q}}{\sigma_i^2 + \delta_i^2} \tag{2.33}$$

From [15], the MSE of this estimator in the case of perfect sensor channels is upper bounded by $D_1 = \left(\sum_{i=1}^{N} \frac{1}{\sigma_i^2 + \delta_i^2} \right)^{-1}$. However if the sensor channels are noisy, then the BLUE is as follows -

$$\hat{\theta}_c = \left(\sum_{i=1}^{N} \frac{1}{\sigma_i^2 + \delta_i^2} \right)^{-1} \sum_{i=1}^{N} \frac{x_{i,c}}{\sigma_i^2 + \delta_i^2}, \tag{2.34}$$

and the estimation error D_2 at the fusion center for this estimator is given as -

$$D_2 = \mathbb{E}\left[\left(\sum_{i=1}^{N}\frac{1}{R_i+R_{i,q}}\right)^{-1}\sum_{i=1}^{N}\frac{x_{i,c}-\theta}{R_i+R_{i,q}}\right]^2$$
$$\leq \left(\sum_{i=1}^{N}\frac{1}{\sigma_i^2+\delta_i^2}\right)^{-2}\sum_{i=1}^{N}\frac{\mathbb{E}(x_{i,c}-\theta)^2}{(\sigma_i^2+\delta_i^2)^2}$$

2.5.2 Analytical Results

We derive an upper bound for the variance of the estimator used in [15] with the underlying assumption that the sensor observation noise distribution is unknown.

Lemma 3: Let D_2 denote the variance associated with the best linear unbiased estimator where the information from the sensors is weighed by weights that depends on its measurement noise and random quantization noise variances at the fusion center. Then D_2 is upper bounded as follows -

$$D_2 \quad \leq (p_0+1)^2 D_1, \tag{2.35}$$

where, $p_0 = \sqrt{4W^2 P_{max}^{\{b\}}}$, and $P_{max}^{\{b\}}$ =max$\{P_i^{\{b\}}\}, \forall i \in \{1,\ldots,N\}$. D_1 is the upper bound on the MSE of the BLUE at the fusion center incorporating effects of observation and random quantization noises when the sensor channels are perfect.

Proof: The quantized information can be written as -

$$x_{i,q} = \left(\sum_{k=1}^{l_i} b_{i,k}2^{l_i-k} - 2^{l_i-1}\right)\Delta_i, \tag{2.36}$$

where, $b_{i,1}$ to b_{i,l_i} represents the MSB of the information to its LSB and $\Delta_i = \frac{2W}{2^{l_i}-1}$. Let $\{\hat{b}_{i,1},\ldots,\hat{b}_{i,c_i}\}$ be the information bits received at the fusion center in each transmission. Then the information received at the fusion center is as follows -

$$x_{i,c} = \left(\sum_{k=1}^{l_i} \hat{b}_{i,k}2^{l_i-k} - 2^{l_i-1}\right)\Delta_i \tag{2.37}$$

Now let $\hat{\theta}_{ch}$ be the final estimate made at the fusion center after information is trans-

mitted across noisy wireless channels. We have -

$$\mathbb{E}|\hat{\theta}_q - \hat{\theta}_c|^2 = \mathbb{E}\left|\left(\sum_{i=1}^{N}\frac{1}{\sigma_i^2+\delta_i^2}\right)^{-1}\sum_{i=1}^{N}\frac{x_{i,q}-x_{i,c}}{\sigma_i^2+\delta_i^2}\right|^2$$
$$= \left(\sum_{i=1}^{N}\frac{1}{\sigma_i^2+\delta_i^2}\right)^{-2}\mathbb{E}\left(\sum_{i=1}^{N}\frac{\left|\sum_{k=1}^{l_i}\left(b_{i,k}-\hat{b}_{i,k}\right)2^{l_i-k}\Delta_i\right|^2}{(\sigma_i^2+\delta_i^2)^2}\right)$$

Now for any random variable Z bounded in $[-U, U]$ we have, $\mathbb{E}(|Z|^2) = \int_{-u}^{u}|Z|^2 p(z)dz \leq U\mathbb{E}(|Z|)$. Hence we have -

$$D_1^2\mathbb{E}\left(\sum_{i=1}^{N}\frac{\left|\sum_{k=1}^{l_i}\left(b_{i,k}-\hat{b}_{i,k}\right)2^{l_i-k}\Delta_i\right|^2}{(\sigma_i^2+\delta_i^2)^2}\right)$$
$$\leq \quad 2WD_1^2\mathbb{E}\left(\sum_{i=1}^{N}\frac{\left|\sum_{k=1}^{l_i}\left(b_{i,k}-\hat{b}_{i,k}\right)2^{l_i-k}\Delta_i\right|}{(\sigma_i^2+\delta_i^2)}\right)$$
$$\leq \quad 2WD_1^2\left(\sum_{i=1}^{N}\frac{\mathbb{E}\left(\sum_{k=1}^{l_i}\left|b_{i,k}-\hat{b}_{i,k}\right|2^{l_i-k}\Delta_i\right)}{(\sigma_i^2+\delta_i^2)}\right)$$

It can be seen that $\left|b_{i,k}-\hat{b}_{i,k}\right|$ is a bernoulli random variable that takes value 1 with probability $P_i^{\{b\}}$ and 0 with probability $1-P_i^{\{b\}}$.Thus we have $\mathbb{E}\left|b_{i,k}-\hat{b}_{i,k}\right| = P_i^{\{b\}}$ and -

$$2WD_1^2\left(\sum_{i=1}^{N}\frac{\mathbb{E}\left(\sum_{k=1}^{l_i}\left|b_{i,k}-\hat{b}_{i,k}\right|2^{l_i-k}\Delta_i\right)}{(\sigma_i^2+\delta_i^2)}\right)$$
$$= \quad 2WD_1^2\sum_{i=1}^{N}\frac{P_i^{\{b\}}\Delta_i 2^{l_i}\sum_{k=1}^{l_i}2^{-k}}{(\sigma_i^2+\delta_i^2)}$$
$$= \quad 2WD_1^2\sum_{i=1}^{N}\frac{P_i^{\{b\}}\Delta_i 2^{l_i}(2^{l_i}-1)}{2^{l_i}(\sigma_i^2+\delta_i^2)}$$
$$= \quad D_1^2\left(\sum_{i=1}^{N}\frac{4W^2P_i^{\{b\}}}{(\sigma_i^2+\delta_i^2)}\right) \leq 4W^2P_{max}^{\{b\}}D_1^2\left(\sum_{i=1}^{N}\frac{1}{(\sigma_i^2+\delta_i^2)}\right),$$

where, $P_{max}^{\{b\}} = \max\{P_i^{\{b\}}\}, \forall i \in \{1,\ldots,N\}$. We set $p_0 = \sqrt{4W^2P_{max}^{\{b\}}}$. Hence we have -

$$\mathbb{E}|\hat{\theta}_q - \hat{\theta}_c|^2 \leq p_0^2 D_1$$

Now the overall estimation error upper bound can be evaluated as follows -

$$\begin{aligned}\mathbb{E}\left(|\theta-\hat{\theta}_c|^2\right) &= \mathbb{E}\left(|\theta-\hat{\theta}_q+\hat{\theta}_q-\hat{\theta}_c|^2\right)\\ &\leq \mathbb{E}\left(|\theta-\hat{\theta}_q|^2\right)+\mathbb{E}\left(|\hat{\theta}_q-\hat{\theta}_c|^2\right)\\ &+2\sqrt{\mathbb{E}\left(|\hat{\theta}_q-\hat{\theta}_c|^2\right).\mathbb{E}\left(|\theta-\hat{\theta}_q|^2\right)}\\ &\leq (p_0^2+2p_0+1)D_1\\ &\leq (p_0+1)^2D_1, \end{aligned} \tag{2.38}$$

where the term $2\sqrt{\mathbb{E}\left(|\hat{\theta_q}-\hat{\theta_c}|^2\right).\mathbb{E}\left(|\theta-\hat{\theta_q}|^2\right)}$ is bounded by the Cauchy-Schwarz inequality.

2.5.3 Results and Discussion

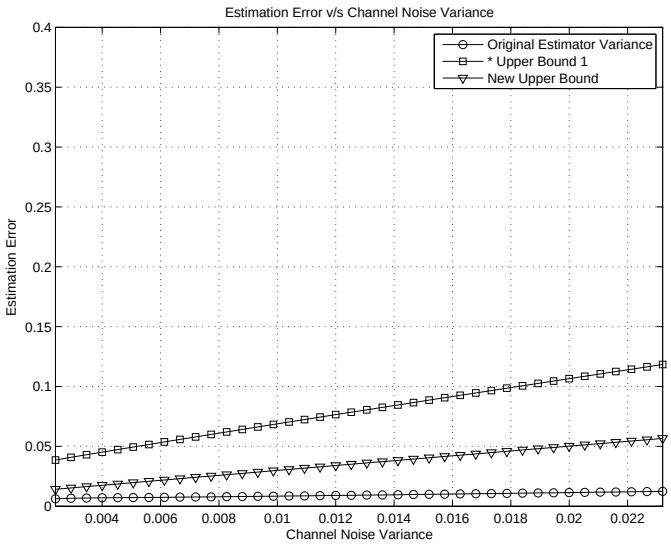


Figure 2.4: Estimator Variance versus Channel Noise Variance for Random Quantization (*Upper Bound 1 refers to the upper bound derived in [15]).

We perform Monte-Carlo simulations for evaluating the performance of the best linear unbiased estimator discussed in the previous section. We consider a wireless sensor network with $N = 20$ sensors. We set the dynamic range of the source signal as $W \in [-1, 1]$ and consider BPSK modulation scheme. Figure 2.4 compares

the performance of the new bound derived in this paper with the result in [15], and actual variance of the estimator that accounts for noise in measurement and uniform random quantization in noisy AWGN channels. We can see that the new bound is tighter than the one in [15] for low values of measurement noise variance of the participating sensors. For both the estimators, it can be seen that the upper and lower bounds derived are tight when the channel noise variance of the participating sensors are similar, and the deviation becomes more pronounced as the channel noise variances vary greatly from each other.

2.6 Summary

We consider three estimators - BLUE-1, BLUE-2, and BLUE-3 for distributed estimation in wireless sensor networks and derive bounds on their variance. We observe for BLUE-1 that the lower bound is an additive factor away from the estimator variance in the case of perfect sensor channels and depends on the sensor with the best channel condition. For BLUE-2, the upper and lower bound for the variance are an additive factor away from the BLUE variance evaluated for perfect sensor channels and can be seen to depend on the sensors with the worst and best channels respectively. For BLUE-3, the upper bound is a multiplicative factor away from the BLUE variance in the case of ideal sensor channels and depends on the sensor with the worst channel condition. We observe that the new upper bound is tighter than the bound in [15] for lower values of measurement noise. Finally, for all the estimators considered, deviation of the bound is observed to be more pronounced when the channel noise variances of the participating sensors vary from each other.

Chapter 3

Estimation Error Minimization

In this chapter, we consider the problem of distributed estimation and resource optimization in an energy and rate-constrained wireless sensor network. To this end, we consider the best linear unbiased estimator (BLUE-1) from chapter 2 that accounts for variance of noise in measurement, uniform quantization and channel.

We analyze the tradeoff between estimation error (BLUE variance) at the fusion center and the total amount of resources utilized (power and rate), and determine optimal sensor actions (power and rate) using three different system design approaches or optimization formulations.In all three formulations, the original optimization problem is observed to be intricately non-convex that is transformed to a Difference of Convex functions (D.C.) problem. Further, using Taylor expansion we present the convex approximated version of all three problems whose solutions (the global minimizers) are verified to be the same as that of the original optimization problem and its D.C. version. For all the formulations, we determine optimum quantization bits and transmission power per bit (or optimum actions) for all sensors jointly. Unlike prior efforts, we incorporate the operating state (characterized by the amount of residual battery power) of the sensors in the optimization framework. We study the effect of channel quality, local measurement noise, and operating states of the sensors on their optimum choice for quantization bits and transmit power per bit.

First, we present the system model in section 3.1. Then we present the first optimization formulation for Scheme 1 (BPSK) and Scheme 2 (QAM) - Formulation A in section 3.2, where, we seek to minimize BLUE variance at the fusion center subject to total resource constraint in the network. This is then ensued by KKT analysis, results and discussion.

3.1 System Model

Consider a scenario where a wireless sensor network consisting of N spatially distributed sensors are actively estimating a deterministic source signal θ related to an underlying physical phenomenon of interest. Considering a linear observation model, measurement made by the $i-$th sensor, $\forall i \in \{1, \ldots, N\}$ is -

$$x_i \quad = \quad \theta + n_i, \tag{3.1}$$

where, n_i is the $i-$th sensor's measurement noise, with zero mean, and spatially uncorrelated with variance σ_i^2. Each sensor uniformly quantizes its observation x_i, $\forall i \in \{1, \ldots, N\}$ locally as follows -

$$x_{i,q} \quad = \quad x_i + n_{i,q}, \tag{3.2}$$

where, $n_{i,q}$ is the quantization noise of the $i-$th sensor, $\forall i \in \{1, \ldots, N\}$. The quantized information is then mapped to b_i bits and transmitted over independent (realized by orthogonal signalling schemes like TDMA, FDMA, or CDMA) noisy wireless channels to a data fusion center. We consider two cases of modulation schemes employed by the active sensor nodes to transmit the quantized information to the fusion center namely: BPSK modulation and uncoded QAM modulation. We refer to the cases as Scheme 1 and Scheme 2 respectively. It is quite straight-forward to extend this to other modulation schemes as well. At the fusion center, a linear combination of the information received from the active sensors is performed thereby generating the final estimate of θ.

For both schemes considered, information received from the $i-$th sensor at the fusion center is given as -

$$x_{i,c} \;=\; x_{i,q} + n_{i,c}, \tag{3.3}$$

where, $n_{i,c}$ is the additive white gaussian noise due to imperfect channel with channel noise power spectral density of $N_0/2$, $\forall i \in \{1, \ldots, N\}$.

At the fusion center, observations received from the sensors $x_{i,c}$ $\forall i \in \{1, \ldots, N\}$, that are laden with measurement, quantization, and channel noises, are combined linearly to form the best linear unbiased estimate or the centralized BLUE of θ. More specifically, we consider a generalized version of the BLUE at the fusion center that weighs each information or observation received with its total variance that depends on its measurement, quantization, and channel noises. This is given as -

$$\hat{\theta} \;=\; \left(\sum_{i=1}^{N} \frac{1}{\mathbb{E}(x_{i,c} - \theta)^2} \right)^{-1} \sum_{i=1}^{N} \frac{x_{i,c}}{\mathbb{E}(x_{i,c} - \theta)^2}, \tag{3.4}$$

where, $\mathbb{E}(.)$ denotes the expectation operator. We assume that the data fusion center has complete knowledge of the variance associated with the observations received from each sensor. We use the mean squared error associated with this estimator, denoted by D, as a measure of the quality of the final estimate generated and is given as [**?**] -

$$\begin{aligned} D \;&=\; \mathbb{E}\left[\left(\sum_{i=1}^{N} \frac{1}{\mathbb{E}(x_{i,c} - \theta)^2} \right)^{-1} \sum_{i=1}^{N} \frac{x_{i,c} - \theta}{\mathbb{E}(x_{i,c} - \theta)^2} \right]^2 \\ &=\; \left(\sum_{i=1}^{N} \frac{1}{\mathbb{E}(x_{i,c} - \theta)^2} \right)^{-1} = \left(\sum_{i=1}^{N} \frac{1}{(R_i + R_{i,q} + R_{i,c})} \right)^{-1} \end{aligned} \tag{3.5}$$

Let $R_i = \mathbb{E}(n_i^2)$, $R_{i,q} = \mathbb{E}(n_{i,q}^2)$, $R_{i,c} = \mathbb{E}(n_{i,c}^2)$, $\forall i \in \{1, \ldots, N\}$. Considering measurement, quantization, and channel noises to be statistically independent of each other, it can be seen that $\mathbb{E}(x_{i,c} - \theta)^2 = R_i + R_{i,q} + R_{i,c}$, $\forall i \in \{1, \ldots, N\}$. Let $\theta \in [-W, W]$, where $[-W, W]$ represents the dynamic range of the signal source. Then $R_{i,q} = \frac{W^2}{3(2^{l_i}-1)^2}$ is the uniform quantization noise variance. Here $l_i \in [1, BW]$ is

the number of quantization bits used in a transmission by the i-th sensor and BW denotes the total rate constraint of the system.

We know from the previous chapter that when a particular modulation scheme that results in a bit error probability of $P_{i,k}^{\{b\}}$ for the $k-$th bit in the transmitted information is used by the $i-$th sensor for reporting to the fusion center, the variance due to imperfect channel is -

$$R_{i,c} = \mathbb{E}(n_{i,c}^2) \approx \frac{4W^2}{3} P_i^{\{b\}} \tag{3.6}$$

We assume that each information transmission consists of l_i bits, and that there is at most only one bit in error in the transmitted information. Also, we assume that all the bits in the transmitted information have the same bit error probability (i.e. $P_{i,k}^{\{b\}} = P_i^{\{b\}}$) and that the channel condition experienced by a sensor remains unchanged during a complete information transmission.

Thus, in Scheme 1 when the $i-$th sensor uses BPSK modulation scheme for transmission in noisy Rayleigh fading channels, equation (3.6) becomes -

$$R_{i,c} = \frac{4W^2}{3} P_i^{\{b\}} = \frac{4W^2}{3} \left(1 - \sqrt{\frac{0.5\Gamma_i}{1 + 0.5\Gamma_i}}\right), \tag{3.7}$$

where, $\Gamma_i = \frac{2p_i\overline{|h_i|^2}}{n_i}$ represents the average received signal to noise ratio, $p_i \in (p_{min}^{(i)}, p_{max}^{(i)})$ is the power level per bit used in a transmission. $p_{min}^{(i)}$ is the minimum power level per bit in a transmission based on channel conditions so as to achieve a minimum SNR requirement of the system. $p_{max}^{(i)}$ is the maximum power per bit in an RF transmission of information. p_i does not account for the electronics circuit power as we assume that the RF transmission power for all sensors is significantly larger than circuit power consumption. $\overline{|h_i|^2}$ is the average power of the Rayleigh fading channel coefficient that is assumed to be constant during a single information transmission, but varies across multiple transmissions. n_i is the channel noise power experienced by the i-th sensor. Suppose the $i-$th sensor chooses Scheme 2 where uncoded QAM modulation

is used for transmission of l_i bits such that b_i is the size of each symbol transmitted. Let $c_i \in \mathbb{Z}^+, \forall i \in \{1, \ldots, N\}$ denote the number of symbols transmitted. Then the variance due to imperfect channel associated with the complete information is given as -

$$R_{i,c} = \frac{4W^2}{3} P_i^{\{b\}} = \frac{4W^2 P_i^{\{s\}} l_i}{3b_i^2}, \tag{3.8}$$

where, $P_i^{\{s\}}$ is the symbol error probability associated with the $i-$th sensor's information transmission.

3.2 Formulation A - Minimize D Subject to Total Resource Constraint.

We first consider the problem of minimizing variance D of the estimator described in the previous section subject to a total system resource constraint. We address the following question -

(Q1) What is the optimal power level per bit and the optimal number of quantization bits in an information transmission for every active sensor such that the variance of the best linear unbiased estimator D at the fusion center is minimized subject to a total resource constraint?

In other words, we seek to determine the optimal sensor actions (in terms of power level per bit and number of quantization bits in an information transmission) that renders the best linear estimator given the maximum amount of resources permitted to be expended by the active sensors in the system. Formally, this problem is

expressed in the standard form as -

Minimize D

Subject to

$J - E_{sys}^{\{max\}} \leq 0;$ - Resource constraint

$\sum_{i=1}^{N} b_i - BW \leq 0;$ - Rate constraint

$-b_i + b_{low} \leq 0;$

$p_i - p_{max}^{(i)} \leq 0, -p_i + p_{min}^{(i)} \leq 0;$

$\forall i \in \{1, \ldots, N\}$

or equivalently -

Maximize D^{-1}

Subject to

$\sum_{i=1}^{N} \Lambda_i p_i b_i - E_{sys}^{\{max\}} \leq 0;$

$\sum_{i=1}^{N} b_i - BW \leq 0, -b_i + b_{low} \leq 0;$

$p_i - p_{max}^{(i)} \leq 0, -p_i + p_{min}^{(i)} \leq 0;$

$\forall i \in \{1, \ldots, N\},$

where, the BLUE variance $D = \left(\sum_{i=1}^{N} \frac{1}{(R_i + R_{i,q} + R_{i,c})}\right)^{-1}$ is as from previous section. For Scheme 1 (BPSK), variance due to imperfect channel links $R_{i,c}$ is given as follows -

$$R_{i,c} = \frac{4W^2}{3} P_i^{\{b\}} = \frac{4W^2}{3} \left(1 - \sqrt{\frac{0.5\Gamma_i}{1 + 0.5\Gamma_i}}\right),$$

In this scheme, the total resource expended by the active sensors in the system is defined as -

$$J = \sum_{i=1}^{N} \Lambda_i p_i b_i, \tag{3.9}$$

where, p_i denotes the power associated with each bit in the information, b_i refers to the total number of bits constituting the information that is transmitted, and $\Lambda_i \in [0, 1]$ is a weighing parameter. It may be noted that in Scheme 1 (BPSK), the number of bits used for quantization is equal to the number of bits used for transmission (i.e. one information transmission), i.e. $l_i = b_i$ and $b_{low} = 1$. In equation (3.9), the total amount of resources expended by all sensors in the system is expressed as a weighted sum of the total power $p_i b_i$ used by each sensor in an information transmission. Λ_i can be treated as a parameter that reflects the resource policy of each sensor based on its operating state: If a sensor is operating on low residual battery power, it would choose a high Λ_i value (close to unity) thereby showing greater affinity for resource economical actions that help in conserving its battery power. A smaller value for Λ_i (close to zero) indicating higher residual battery power would imply that the sensor is ready to expend as much as resource possible in order to minimize the estimator variance at the fusion center. $E_{sys}^{\{max\}}$ is the maximum amount of resources that could be used by all the active sensors in the system for an information transmission.

For Scheme 2 (QAM), $R_{i,c}$ is as follows -

$$R_{i,c} = \frac{4W^2 P_i^{\{s\}} l_i}{3b_i^2} \leq \frac{16W^2}{3b_i} \exp\left(\sqrt{\frac{3\gamma_i^{\{s\}}}{2(2^{b_i}-1)}}\right), \tag{3.10}$$

where, $\gamma_i^{\{s\}}$ is the average SNR per symbol. The total resource consumption by the active sensors in the system is defined as -

$$J = \sum_{i=1}^{N} \Lambda_i p_i c_i, \tag{3.11}$$

where, c_i is the number of symbols transmitted as a part of one information transmission, $p_i = B_s K_i a_i \log\left(\frac{2}{p_b}\right)(2^{b_i} - 1)$ is the average power associated with each symbol, $K_i = 2N_f N_0 G_d$, B_s is the sampling rate of the observed/measured signal by the $i-$th sensor, N_f is the receiver noise figure. G_d is a power gain factor defined as $G_d = G_1 d^{\{\kappa\}} M_l$, , $a_i = d^{\kappa}$,where, $d^{\{\kappa\}}$ is the κ-power path-loss model at

distance d for the $i-$th sensor. G_1 is the gain factor at $d = 1$ (depends on antenna gain, carrier frequency and other system parameters). M_l is the link margin that compensates for variations in hardware processes and other background noise or interference. We set all these values as in [15]. We also assume that all the sensors have the same target symbol error probability. $\Lambda_i \in [0, 1]$ is a weighing parameter and has the same implications as in Scheme 1 (BPSK). We consider the case of $l_i = b_i \Rightarrow c_i = 1; \forall i \in \{1, \ldots, N\}$ and $b_{low} = 2$, implying that the only variable to optimize for Scheme 2 is b_i that in turn determines the power per transmission.

All the problem variables for both schemes are constrained to be non-negative by the constraints $b_i \geq 1$ and $p_{min}^{(i)} \leq p_i \leq p_{max}^{(i)}$. We shall henceforth refer to these constraints as box constraints.

In general, call the formulation as *OR-A* referring it as the original formulation of the problem. The above formulation is intricately non-convex in its variables p_i and $b_i, \forall i \in \{1, \ldots, N\}$. More specifically, it is essentially a mixed-integer nonlinear programming (MINLP) optimization problem as bits are discrete valued while power levels are continuous valued. However relaxing the bits to be continuous valued, the formulation transforms to a nonlinear optimization problem. We set $x_i = \frac{1}{(R_i + \frac{W^2}{3(2^{b_i}-1)^2} + R_{i,c})}$ and then reformulate the above optimization problem as follows -

$$\begin{aligned}
&\text{Minimize } \sum_{i=1}^{N} -x_i \\
&\text{Subject to} \\
&x_i - \frac{1}{(R_i + \frac{W^2}{3(2^{b_i}-1)^2} + R_{i,c})} = 0 \\
&J - E_{sys}^{\{max\}} \leq 0; \\
&\sum_{i=1}^{N} b_i - BW \leq 0,
\end{aligned}$$

where, the box constraints are implicit. Call the above reformulation as *MOD-A* referring it as the modified version of the original problem. For both schemes, it can be

seen that the objective function is decreasing in x_i, the equality constraint is increasing in x_i but decreasing in p_i, b_i, and the resource constraint is increasing in $p_i, b_i, \forall i \in \{1, \ldots, N\}$. Thus the equality in the above formulation can be transformed to an inequality to form an equivalent problem. This is because the inequality introduced is strictly active at the minima. Hence the reformulated problem is of the form -

$$\begin{aligned}
&\text{Minimize } \sum_{i=1}^{N} -x_i \\
&\text{Subject to} \\
&R_i + \frac{W^2}{3(2^{b_i}-1)^2} + R_{i,c} - \frac{1}{x_i} \leq 0 \\
&J - E_{sys}^{\{max\}} \leq 0; \\
&\sum_{i=1}^{N} b_i - BW \leq 0,
\end{aligned}$$

with the box constraints being implicit. In the above reformulation for both schemes, it can be seen that $R_i + \frac{W^2}{3(2^{b_i}-1)^2} + R_{i,c}$ is a sum of convex functions and hence convex in its variables for both the schemes under consideration, while $-\frac{1}{x_i}$ is concave in x_i. Hence the expression in the inequality is a difference of convex functions and the above problem assumes a typical form of difference of convex programming (D.C. programming). We refer to the above formulation as *DC-A* referring it as the D.C formulation of the modified problem. Global optimizers of such problems can be determined using branch and bound techniques and outer approximation methods. However the complexity of D.C. programming is NP hard and the convergence time of any algorithm to determine global minimizer is long - a feature that is undesirable in the context of wireless sensor networks. It may be noted that by making a convex approximation of the difference of convex functions constraint in the above reformulation, we can pose the optimization problem as a convex optimization problem. Thus

the approximated problem is as follows -

$$\begin{aligned}
&\text{Minimize } \sum_{i=1}^{N} -x_i \\
&\text{Subject to} \\
&R_i + \frac{W^2}{3(2^{b_i}-1)^2} + R_{i,c} - \left(\frac{1}{x_i^{\{m\}}} + \frac{x_i - x_i^{\{m\}}}{(x_i^{\{m\}})^2} \right) \leq 0 \\
&J - E_{sys}^{\{max\}} \leq 0; \\
&\sum_{i=1}^{N} b_i - BW \leq 0,
\end{aligned}$$

where, the box constraints are implicit and the $\frac{1}{x_i}, \forall i \in \{1, \ldots, N\}$ term in the D.C. constraint has been expressed via first-order Taylor expansion about $x_i^{\{m\}}$. We refer to this convex approximated form as *CA-A*.

3.2.1 Analysis

By applying KKT conditions, we determine the conditions satisfied by the minima for the formulation in *DC-A* for Scheme 1 (BPSK). The Lagrangian $G_{1.a}$ is given as -

$$\begin{aligned}
G_{1.a} = &\sum_{i=1}^{N} -x_i + \sum_{i=1}^{N} \eta_i \left(R_i + \frac{W^2}{3(2^{b_i}-1)^2} + \right. \\
&\frac{4W^2}{3} \left(1 - \sqrt{\frac{p_i \overline{|h_i|^2}}{p_i \overline{|h_i|^2} + n_i}} \right) - \frac{1}{x_i} \Bigg) + \eta_{N+1} \left(\sum_{i=1}^{N} \Delta_i p_i b_i \right. \\
&\left. - E_{sys}^{(max)} \right) + \eta_{N+2} \left(\sum_{i=1}^{N} b_i - BW \right) + \sum_{i=1}^{N} \eta_{N+2+i}(-b_i + 1) + \\
&\sum_{i=1}^{N} \eta_{2N+2+i}(p_i - p_i^{(max)}) + \sum_{i=1}^{N} \eta_{3N+2+i}(-p_i + p_i^{(min)}),
\end{aligned}$$

where, $\eta = [\eta_1, \ldots, \eta_{4N+2}]^T$ are the Lagrange multipliers that necessarily exist such that -

$$\begin{aligned}
&\frac{\partial G_{1.a}}{\partial p_k} = -\eta_k \left(\frac{4W^2}{3} \frac{n_k \sqrt{\overline{|h_k|^2}}}{(p_k^* \overline{|h_k|^2} + n_k)^{3/2} p_k^{*1/2}} \right) \\
&+\eta_{N+1} \Lambda_k b_k^* + \eta_{2N+2+k} - \eta_{3N+2+k} = 0 \\
&\frac{\partial G_{1.a}}{\partial b_k} = -\eta_k \frac{W^2 \ln 2}{3(2^{b_k^*} - 1)^2} + \eta_{N+1} \Lambda_k p_k^* \\
&+\eta_{N+2} - \eta_{N+2+k} = 0 \\
&\frac{\partial G_{1.a}}{\partial x_k} = -1 + \frac{\eta_k}{x_k^{*2}} = 0
\end{aligned} \tag{3.12}$$

Similarly for Scheme 2 (QAM) in *DC-A*, the Lagrangian $G_{1.b}$ is as follows -

$$\begin{aligned}
G_{2.a} = &\sum_{i=1}^{N} -x_i + \sum_{i=1}^{N} \eta_i \left(R_i + \frac{W^2}{3(2^{b_i} - 1)^2} + \right. \\
&\left. \frac{16W^2}{3} \exp \sqrt{\frac{3B_s K_i a_i \log\left(\frac{2}{p_b}\right)}{2N_0}} - \frac{1}{x_i} \right) \\
&+\eta_{N+1} \left(\sum_{i=1}^{N} \Delta_i B_s K_i a_i \log\left(\frac{2}{p_b}\right) (2^{b_i} - 1) \right. \\
&\left. - E_{sys}^{(max)} \right) + \eta_{N+2} \left(\sum_{i=1}^{N} b_i - BW \right) \\
&+ \sum_{i=1}^{N} \eta_{N+2+i}(-b_i + 2),
\end{aligned} \tag{3.13}$$

where, $\eta = [\eta_1, \ldots, \eta_{3N+2}]^T$ are the Lagrange multipliers that necessarily exist such that -

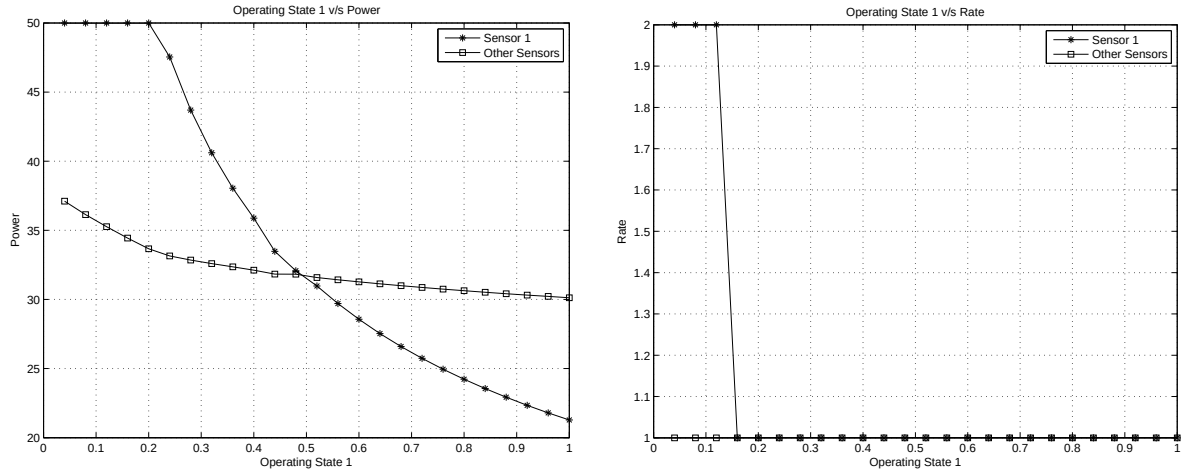
$$\begin{aligned}
&\frac{\partial G_{2.a}}{\partial b_k} = -\eta_k \left(\frac{W^2 \ln 2}{3(2^{b_k^*} - 1)^2} + \frac{16W^2}{3 b_k^{*2}} \exp \sqrt{\frac{3B_s K_i a_i \log\left(\frac{2}{p_b}\right)}{2N_0}} \right) \\
&+\eta_{N+1} \Delta_k B_s K_k a_k \log\left(\frac{2}{p_b}\right) 2^{b_k^*} \ln 2 + \eta_{N+2} - \eta_{N+2+k} = 0 \\
&\frac{\partial G_{1.a}}{\partial x_k} = -1 + \frac{\eta_k}{x_k^{*2}} = 0
\end{aligned} \tag{3.14}$$

3.2.2 Results

It may be noted that the discussion in the sequel for Formulation A applies to both schemes unless otherwise explicitly specified. For both Scheme 1 (BPSK) and Scheme 2 (QAM), we perform numerical simulations for the formulation in *DC-A*, *CA-A* and an exhaustive search for the global minimizers $x_i^*, p_i^*, b_i^*, \forall i \in \{1, \ldots, N\}$ of the original problem *OR-A*. We set $N = 5$, $BW = 20$, minimum system SNR requirement as 5dB, $p_{max}^{(i)} = 50$ units, $\forall i \in \{1, \ldots, N\}$ and $E_{sys}^{\{max\}} = 100$. First, we observe that the points of minima obtained from simulations for *DC-A*, *CA-A* and the exhaustive search are the same. This implies that the intricately non-convex problem in *OR-A* and *DC-A* can be approximated as a convex problem *CA-A* whose minima is guaranteed to be global minimizer of the problem. This is further bolstered by results from the exhaustive search for global minimizers of the problem.

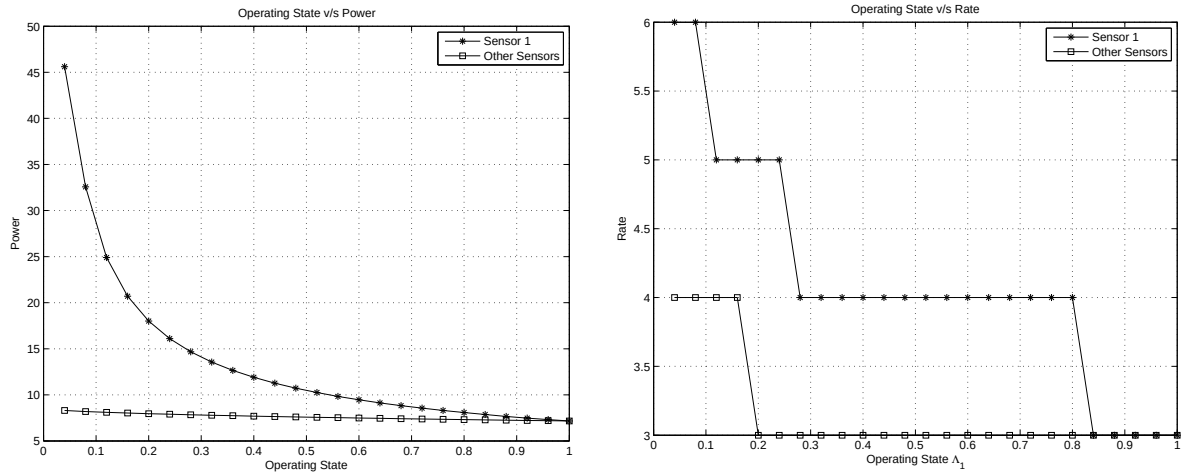
Optimal Actions and Operating State - We first analyze the variation of optimal sensor actions with operating states (for different values of $\Lambda_i, \forall i \in \{1, \ldots, N\}$). Λ_i values close to zero indicate that the active sensors are healthy; Λ_i values around 0.5 indicate mid-range healthy operating state and unhealthy operating state is indicated by Λ_i value close to unity. The Λ_i value is varied for $i = 1$ and the dependency of optimal actions of all the sensors is observed. In all the three cases of operating states of the active sensors, we observe that the optimal actions p_i^*, b_i^* relate to the operating states $\Lambda_i, \forall i \in \{1, \ldots, N\}$ inversely. This is illustrated in figures 3.1 for Scheme 1 (BPSK) and figures in 3.2 for Scheme 2 (QAM), where, $\Lambda_i, i = 1$ is varied from zero to one and the collaborating sensors are in mid-range healthy operating state. In general $\forall i \in \{1, \ldots, N\}$, we observe that for larger values of Λ_i, the corresponding sensor show greater tendency to conserve its battery power resulting in lower values of p_i^*, b_i^*. This comes at the cost of higher estimator variance at the fusion center. Conversely for the same channel conditions and measurement noise variances, smaller Λ_i values result in higher values of p_i^*, b_i^*, consequently reducing the amount

of variance due to imperfect communication. Hence, the overall amount of error at the fusion center does not depend on the quality of observation alone, but also the operating state or the residual battery power of the active sensors in the system.



(a) (b)

Figure 3.1: Scheme 1 (BPSK) - (a) Variation of Λ_1 with the Power used by Sensor (i=1) and Collaborating Mid-Range Healthy Sensors. (b) Variation of Λ_1 with the Bits used by Sensor (i=1) and Collaborating Mid-Range Healthy Sensors.



(a) (b)

Figure 3.2: Scheme 2 (QAM) - (a) Variation of Λ_1 with the Power used by Sensor (i=1) and Collaborating Mid-Range Healthy Sensors. (b) Variation of Λ_1 with the Bits used by Sensor (i=1) and Collaborating Mid-Range Healthy Sensors.

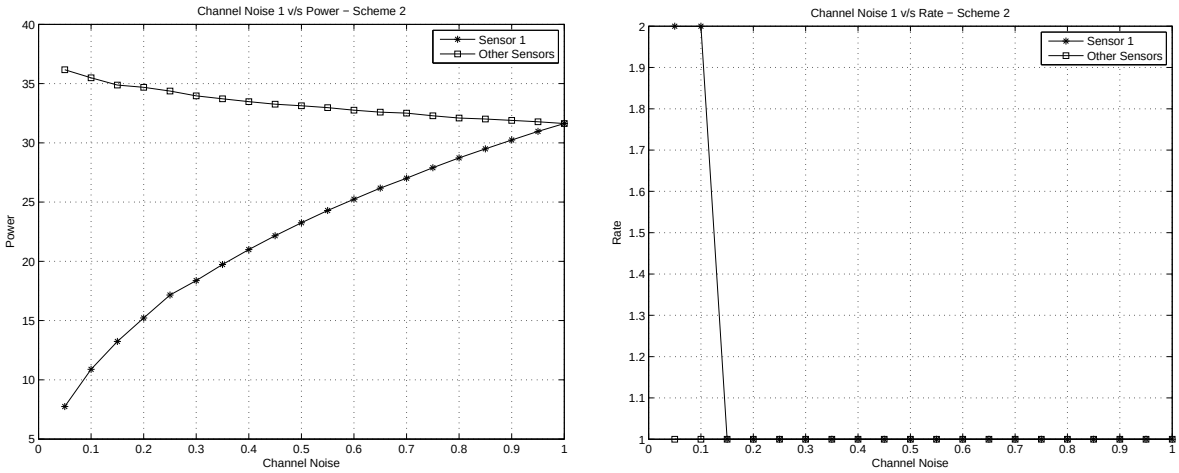
We also make other interesting observations as follows -

1. The value of η_{N+1}, Lagrange multiplier for the system resource constraint, increases with increase in $\Lambda_i \forall i \in \{1, \ldots, N\}$ (or deteriorating operating states of the collaborating sensors). Intuitively, this means that the cost of resource utilization in the system increases in with deteriorating operating states of the active sensors.

2. We observe in figure 3.1 that the optimal p_i, b_i values of the collaborating sensors decrease with increase in Λ_i value of any one of the active sensors. i.e., the collaborating sensors do not compensate for the deterioration of operating state of any of the active sensors by using more resources to minimize system variance. This is because of the system resource expenditure constraint that is modeled as $J \leq E_{sys}^{(max)}$ where any deteriorating operating states cause the corresponding Λ_i to increase.

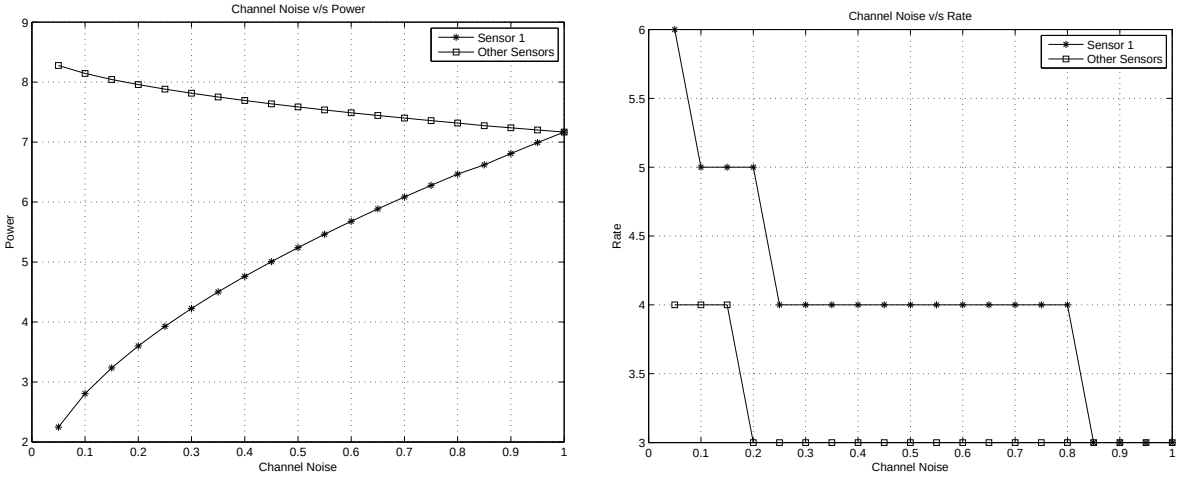
3. The Lagrange multiplier associated with the total rate constraint for Scheme 1 (BPSK), η_{2N+2}, is significantly low in magnitude compared to the Lagrange associated with resource constraint. This implies that changing the total rate constraint of the system does not alter the optimal actions of the active sensors significantly. However for Scheme 2 (QAM), the lagrange associated with the total rate constraint is comparable in magnitude with that of the resource constraint.

Optimal Actions and Channel Conditions - We vary channel condition metricized as $\frac{n_i}{|h_i|^2}$ ($\overline{|h_i|^2} = 1$) for Scheme 1 (BPSK) and as n_i for Scheme 2 (QAM) for $i = 1$, and observe the variation of optimal actions for different operating states of the active sensors. Figures 3.3 and 3.4 illustrate the dependency of $p_i^*, b_i^*, \forall i \in \{1, \ldots, N\}$ on channel condition for $\Lambda_i = 0.5, \forall i \in \{1, \ldots, N\}$ (implying that all the active sensors are in mid-range health condition) for Scheme 1 and Scheme 2 respectively.



(a) (b)

Figure 3.3: Scheme 1 (BPSK) - (a) Variation of $\frac{n_i}{|h_i|^2}$ with the Power used by Sensor (i=1) and Collaborating Mid-Range Healthy Sensors. (b) Variation of $\frac{n_i}{|h_i|^2}$ with the Bits used by Sensor (i=1) and Collaborating Mid-Range Healthy Sensors.



(a) (b)

Figure 3.4: Scheme 2 (QAM) - (a) Variation of n_i with the Power used by Sensor (i=1) and Collaborating Mid-Range Healthy Sensors. (b) Variation of n_i with the Bits used by Sensor (i=1) and Collaborating Mid-Range Healthy Sensors.

We observe that $p_i^*, \forall i \in \{1, \ldots, N\}$ relates to n_i directly. If the channel quality (low n_i) of a sensor is high, then it uses lesser amount of power for transmission of its information. The converse is observed for degraded channel quality or high n_i

values. This is similar to a reverse water-filling process where the total amount of power used for transmission in a channel is inversely related to the channel gain and directly related to channel noise. However it can be seen that $b_i^*, \forall i \in \{1, \ldots, N\}$ or the optimal information length relates inversely to the channel condition metric. For high n_i, lower values of b_i^* ensues and the converse is observed for low n_i values. Similar variations in p_i^*, b_i^* are observed for all the three cases of operating states of the active sensors. Other interesting observations are summarized as follows -

1. The system resource cost η_{N+1} increases with channel noise. This causes the sensors collaborating with the active sensor with bad channel quality to decrease their resources expenditure.

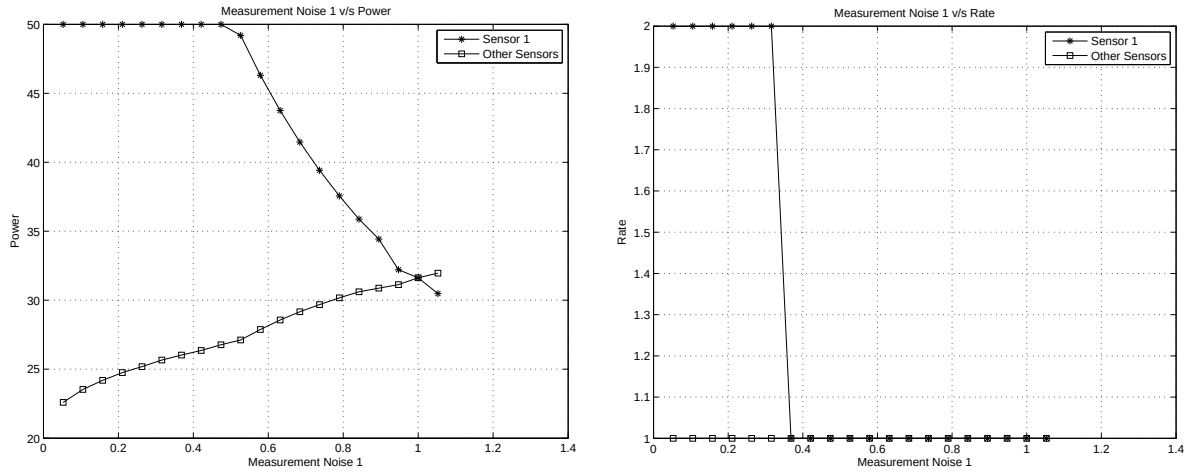
2. It can be deduced that sensors with poor channel quality are severely disadvantaged due to increased utilization of power for transmission of information in order to minimize the estimator variance at the fusion center.

Optimal Actions and Measurement Noise Variance - The measurement noise variance R_i for $i = 1$ is varied for different operating states of the active sensors. Figures 3.5 and 3.6 show variation of the optimal actions p_i, b_i with measurement noise R_1 for sensors in mid-range health $\Lambda_i = 0.5, \forall i \in \{1, \ldots, N\}$ for Scheme 1 (BPSK) and Scheme 2 (QAM) respectively. We observe that for all operating states of the active sensors considered, p_1^*, b_1^* decreases with increase in R_1, while $p_i^*, b_i^*, \forall i \in \{2, \ldots, N\}$, i.e., the total power expended by all the other sensors increases correspondingly increases.

Some of the other inferences are -

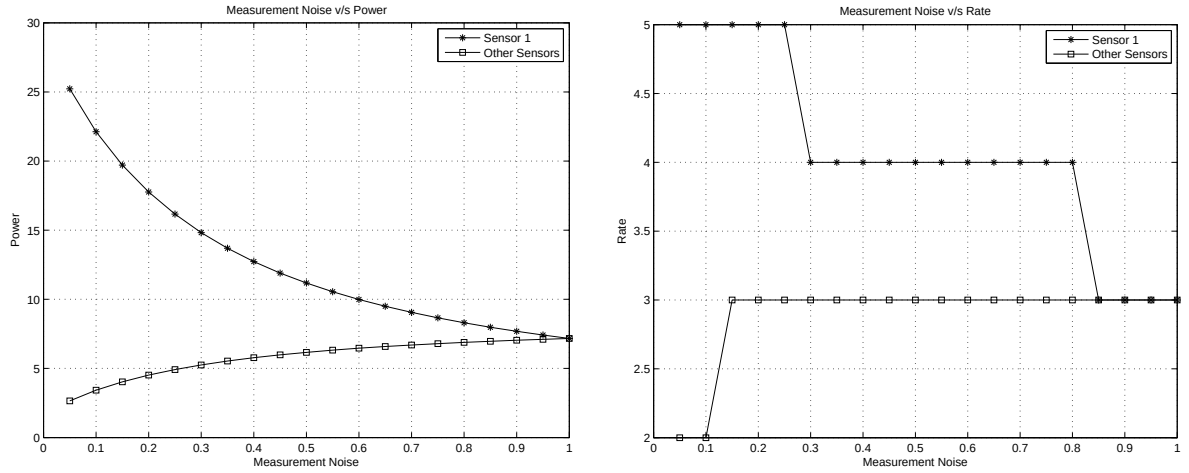
1. In this formulation, active sensors with relatively low measurement noise variance are disadvantaged from a resource utilization perspective when they collaborate with sensors with relatively higher measurement noise variance.

2. The Lagrange associated with the resource constraint η_{N+1} significantly decreases with increase in measurement noise variance of any of the active sensors.



(a) (b)

Figure 3.5: Scheme 1 (BPSK) - (a)Variation of R_1 with the Power used by Sensor (i=1) and Collaborating Mid-Range Healthy Sensors (b) Variation of R_1 with the Bits used by Sensor (i=1) and Collaborating Mid-Range Healthy Sensors.



(a) (b)

Figure 3.6: Scheme 2 (QAM) - (a) Variation of R_1 with the Power used by Sensor (i=1) and Collaborating Mid-Range Healthy Sensors (b) Variation of R_1 with the Bits used by Sensor (i=1) and Collaborating Mid-Range Healthy Sensors.

A Note on the Integer Relaxation of Bits : All the formulations are solved by first treating $b_i, \forall i \in \{1, \ldots, N\}$ as continuous. Upon obtaining optimal b_is, we search exhaustively for the values in $[b_i^*]$ and $[b_i^*] + 1$ $\forall i \in \{1, \ldots, N\}$ subject to the

total rate constraint BW. It is straight-forward to see that the optimal values would lie in the immediate integer neighborhood of b_i for the convex approximated version of all formulations. Since the solutions of the convex approximated version of all formulations match that of the original formulations and the D.C. versions, we are assured that a similar exhaustive search yields optimal solutions for b_i in all cases. This is verified by the exhaustive search performed to find the global minimizers for all the formulations. Note that determining the optimal b_is is NP hard (can be reduced to a knapsack problem) and complexity of the search increases exponentially with the number of active sensor nodes.

3.3 Summary

In this chapter, we presented an optimization formulation that minimizes BLUE variance at the fusion center subject to a total resource constraint in the network for Scheme 1 (BPSK) and Scheme 2 (QAM). We analyzed the tradeoff between the estimation error and the total resource utilized and determined optimal sensor actions. We also studied the effect of operating state, channel conditions, and the measurement noise variance on the optimum choice of power and rate per transmission. Observing that the minimizer for DC-A, CA-A, and the original problem are the same, it appears the original non-convex problem in OR-A can be approximated as a convex optimization problem CA-A for both Schemes without compromising on the optimality of the solutions attained.

Chapter 4

Resource Utilization Minimization

In this chapter, we consider the second optimization formulation - Formulation B, for Scheme 1 (BPSK) and Scheme 2 (QAM), using the same system model presented in chapter 3. Here we seek to minimize the total resource utilization of the active sensors in the network subject to a BLUE variance constraint at the fusion center. In section 4.1, we present the optimization formulation and its modified versions (D.C. and convex approximation), followed by KKT analysis, results and discussion.

4.1 Formulation B - Minimize J Subject to BLUE Variance Constraint.

We seek to minimize the total resource expenditure in the system subject to a constraint on the BLUE variance at the fusion center. We ask -
(Q2) What is the minimum resource (optimal p_i and $b_i, \forall i \in \{1, \ldots, N\}$) that needs to be expended in an information transmission by all active sensors given the maximum tolerable variance of the best linear unbiased estimator at the fusion center?

In order to address this question, we consider both Scheme 1 (BPSK) and Scheme 2 (QAM), and use the resource expenditure and the BLUE variance model

from Formulation A. The problem is modeled as follows -

$$\begin{aligned}
&\text{Minimize } J \\
&\text{Subject to} \\
&D \leq D_{sys}^{\{max\}}; \qquad \text{- BLUE Variance Constraint} \\
&\sum_{i=1}^{N} b_i - BW \leq 0; \qquad \text{- Rate Constraint}
\end{aligned}$$

where the box constraints from Formulation A follow implicitly, and $D_{sys}^{\{max\}}$ is the maximum tolerable variance of the BLUE at the fusion center. $\Lambda_i \in [0,1], \forall i \in \{1,\ldots,N\}$ is a weighing parameter that reflects the relative operating states (as in residual battery powers) of the sensors. When any of the Λ_i values is much higher compared to the rest, it implies that the corresponding sensor has comparatively much lesser residual battery power and hence would be conservative in resource expenditure. The converse applies to any of the sensors whose Λ_i is much lesser than the rest. We refer to the formulation in (4.1) as *OR-B* that has a convex objective function in $p_i, b_i, \forall i \in \{1,\ldots,N\}$ and the BLUE variance constraint being non-convex. By setting $x_i = \frac{1}{(R_i + \frac{W^2}{3(2^{b_i}-1)^2} + R_{i,c})}$, we reformulate the above optimization problem as follows -

$$\begin{aligned}
&\text{Minimize } J \\
&\text{Subject to} \\
&x_i - \frac{1}{(R_i + \frac{W^2}{3(2^{b_i}-1)^2} + R_{i,c})} = 0 \\
&D_{sys}^{\{max\}^{-1}} - \sum_{i=1}^{N} x_i \leq 0; \\
&\sum_{i=1}^{N} b_i - BW \leq 0,
\end{aligned} \tag{4.1}$$

We refer to this problem as *MOD-B*. It can be seen that the objective function is increasing in $p_i, b_i, \forall i \in \{1,\ldots,N\}$, the equality constraint is increasing in x_i but

decreasing in ,p_i, b_i, and the BLUE variance constraint is decreasing in $x_i, \forall i \in \{1, \ldots, N\}$. Thus the equality in (4.1) can be treated as an inequality to formulate an equivalent problem since it is guaranteed to be always active at all points of minima. We reformulate *MOD-B* as a difference of convex functions problem *DC-B* as below -

$$\begin{aligned}
&\text{Minimize } \sum_{i=1}^{N} \Lambda_i p_i b_i \\
&\text{Subject to} \\
&R_i + \frac{W^2}{3(2^{b_i}-1)^2} + R_{i,c} - \frac{1}{x_i} \leq 0 \\
&D_{sys}^{\{max\}^{-1}} - \sum_{i=1}^{N} x_i \leq 0; \\
&\sum_{i=1}^{N} b_i - BW \leq 0,
\end{aligned}$$

Motivated by reasons mentioned earlier, we perform a convex approximation of the D.C. constraint in *DC-B* by using its first-order Taylor expansion. We denote this formulation as *CA-B* -

$$\begin{aligned}
&\text{Minimize } \sum_{i=1}^{N} \Lambda_i p_i b_i \\
&\text{Subject to} \\
&R_i + \frac{W^2}{3(2^{b_i}-1)^2} + R_{i,c} - \\
&\left(\frac{1}{x_i^{\{m\}}} + \frac{x_i - x_i^{\{m\}}}{(x_i^{\{m\}})^2} \right) \leq 0 \\
&D_{sys}^{\{max\}^{-1}} - \sum_{i=1}^{N} x_i \leq 0; \\
&\sum_{i=1}^{N} b_i - BW \leq 0,
\end{aligned}$$

where, the first-order Taylor approximation of $-\frac{1}{x_i}, \forall i \in \{1, \ldots, N\}$, about $x_i^{\{m\}}$ is used.

4.1.1 Analysis

We determine the conditions satisfied by the minima for the formulation in *DC-B* by applying KKT conditions. The Lagrangian $G_{1.b}$ for Scheme 1 (BPSK) in the formulation *DC-B* is as follows -

$$
\begin{aligned}
G_{1.b} &= \sum_{i=1}^{N} \Delta_i p_i b_i + \eta_1 (D_{sys}^{\{max\}^{-1}} - \sum_{i=1}^{N} x_i) + \sum_{i=1}^{N} \eta_{i+1} \left(R_i \right. \\
&+ \frac{W^2}{3(2^{b_i}-1)^2} + \frac{4W^2}{3} \left(1 - \sqrt{\frac{p_i \overline{|h_i|^2}}{p_i \overline{|h_i|^2} + n_i}}\right) - \frac{1}{x_i} \Bigg) + \\
&\eta_{N+2} \left(\sum_{i=1}^{N} b_i - BW \right) + \sum_{i=1}^{N} \eta_{N+2+i}(-b_i + 1) + \\
&\sum_{i=1}^{N} \eta_{2N+2+i}(p_i - p_i^{(max)}) + \sum_{i=1}^{N} \eta_{3N+2+i}(-p_i + p_i^{(min)}),
\end{aligned}
$$

where, $\eta = [\eta_1, \ldots, \eta_{4N+2}]^T$ are the Lagrange multipliers that necessarily exist such that -

$$
\begin{aligned}
&\frac{\partial G_{1.b}}{\partial p_k} = \Lambda_k b_k^* - \eta_{k+1} \left(\frac{4W^2}{3} \frac{n_k \sqrt{\overline{|h_k|^2}}}{(p_k^* \overline{|h_k|^2} + n_k)^{3/2} p_k^{*1/2}} \right) \\
&+\eta_{2N+2+k} - \eta_{3N+2+k} = 0 \\
&\frac{\partial G_{1.b}}{\partial b_k} = \Lambda_k p_k^* - \eta_{k+1} \frac{W^2 \ln 2}{3(2^{b_k^*}-1)^2} \\
&+\eta_{N+2} + \eta_{N+2+k} = 0 \\
&\frac{\partial G_{1.b}}{\partial x_k} = -\eta_1 + \frac{\eta_{k+1}}{x_k^{*2}} = 0
\end{aligned} \tag{4.2}
$$

Similarly for Scheme 2 (QAM) in *DC-A*, the Lagrangian $G_{2.b}$ is as follows -

$$
\begin{aligned}
G_{2.b} &= \sum_{i=1}^{N} \Delta_i B_s K_i a_i \log\left(\frac{2}{p_b}\right)(2^{b_i}-1) + \eta_1 (D_{sys}^{\{max\}^{-1}} - \sum_{i=1}^{N} x_i) \\
+ &\quad \sum_{i=1}^{N} \eta_{i+1} (R_i + \frac{W^2}{3(2^{b_i}-1)^2} + \frac{16W^2}{3} \exp \sqrt{\frac{3B_s K_i a_i \log\left(\frac{2}{p_b}\right)}{2N_0}} - \frac{1}{x_i}) \\
+ &\quad \eta_{N+2} \left(\sum_{i=1}^{N} b_i - BW \right) + \sum_{i=1}^{N} \eta_{N+2+i}(-b_i + 2),
\end{aligned} \tag{4.3}
$$

where, $\eta = [\eta_1, \ldots, \eta_{3N+2}]^T$ are the Lagrange multipliers that necessarily exist such that -

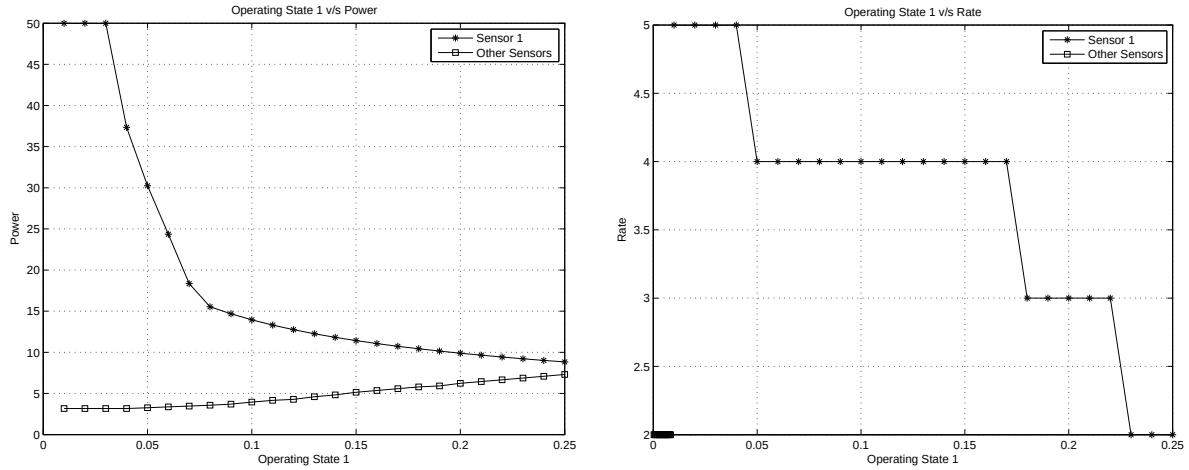
$$\frac{\partial G_{2.a}}{\partial b_k} = \Delta_k B_s K_k a_k \log\left(\frac{2}{p_b}\right) 2^{b_k^*} \ln 2$$
$$-\eta_{k+1}\left(\frac{W^2 \ln 2}{3(2^{b_k^*}-1)^2} + \frac{16W^2}{3b_k^2}\exp\sqrt{\frac{3B_s K_i a_i \log\left(\frac{2}{p_b}\right)}{2N_0}}\right) + \eta_{N+2} - \eta_{N+2+k} = 0$$
$$\frac{\partial G_{1.a}}{\partial x_k} = -1 + \frac{\eta_k}{x_k^{*2}} = 0 \tag{4.4}$$

4.1.2 Results

We consider a network with system parameters similar to that described in the previous chapter and set $D_{sys}^{\{max\}} = 0.35$. Numerical simulations are performed for both the schemes using formulations *DC-B* and *CA-B*, followed by an exhaustive search for the global minimizers $p_i^*, b_i^*, x_i^*, \forall i \in \{1, \ldots, N\}$ of the original problem *OR-B*. We observe that the points of minima are the same in all the cases of simulations implying that the non-convex problem in *OR-B* and *DC-B* can be reduced to the approximated formulation in *CA-B* without compromising on the quality of the solutions.

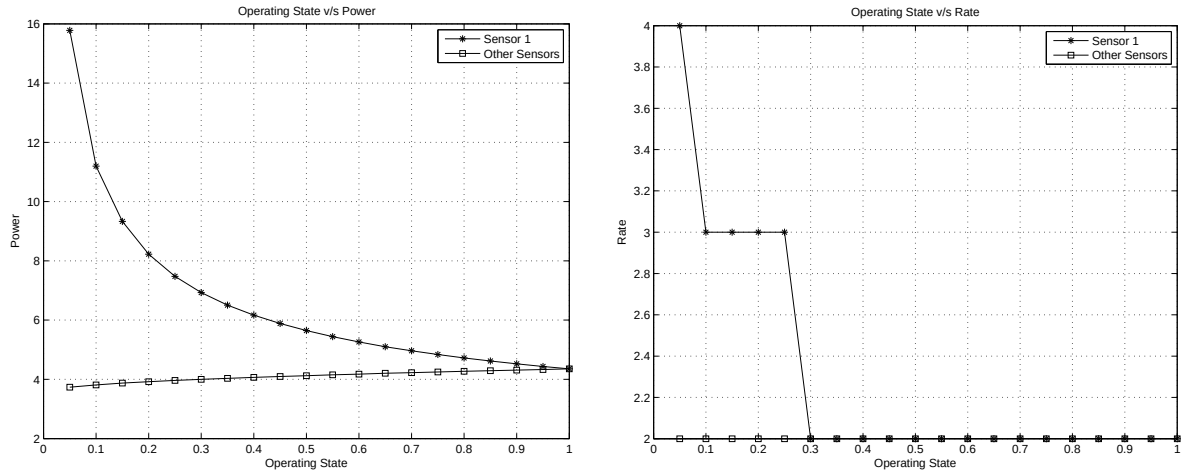
Optimal Actions and Operating States - We first set the value of Λ_i for $i = 1$ significantly low compared to the other Λ_i values in order to study the case of a healthy active sensor amongst unhealthy active sensors. The variation of $p_i^*, b_i^*, \forall i \in \{1, \ldots, N\}$ with $\Lambda_i, i = 1$, are presented in figures 4.1 and 4.2 for Scheme 1 (BPSK) and Scheme 2 (QAM) respectively. As before, we observe that $p_i^*, b_i^*, i = 1$, depends on Λ_1 inversely. In general, we note that as Λ_1 increases, there is an increase in the resource spending of the collaborating sensors $i = \{2, \ldots, N\}$. This is because as the healthy sensor's ($i = 1$) state deteriorates (increasing Λ_1), there is a decrease in its power spending that results in a corresponding increase in the total power spending of the collaborating sensors in order to satisfy the BLUE variance constraint at the fusion center. A similar trend is observed when all $\Lambda_i, \forall i \in \{1, \ldots, N\}$ are comparable (i.e.

comparable operating states) and when $\Lambda_1 > \Lambda_i, \forall i \in \{2, \ldots, N\}$ (i.e. when sensor $i = 1$ is much unhealthier compared to the other collaborating sensors).



(a) (b)

Figure 4.1: Scheme 1 (BPSK) - (a) Variation of Λ_1 with the Power used by Sensor (i=1) and Collaborating Mid-Range Healthy Sensors. (b) Variation of Λ_1 with the Bits used by Sensor (i=1) and Collaborating Mid-Range Healthy Sensors.



(a) (b)

Figure 4.2: Scheme 2 (QAM) - - (a) Variation of Λ_1 with the Power used by Sensor (i=1) and Collaborating Mid-Range Healthy Sensors. (b) Variation of Λ_1 with the Bits used by Sensor (i=1) and Collaborating Mid-Range Healthy Sensors.

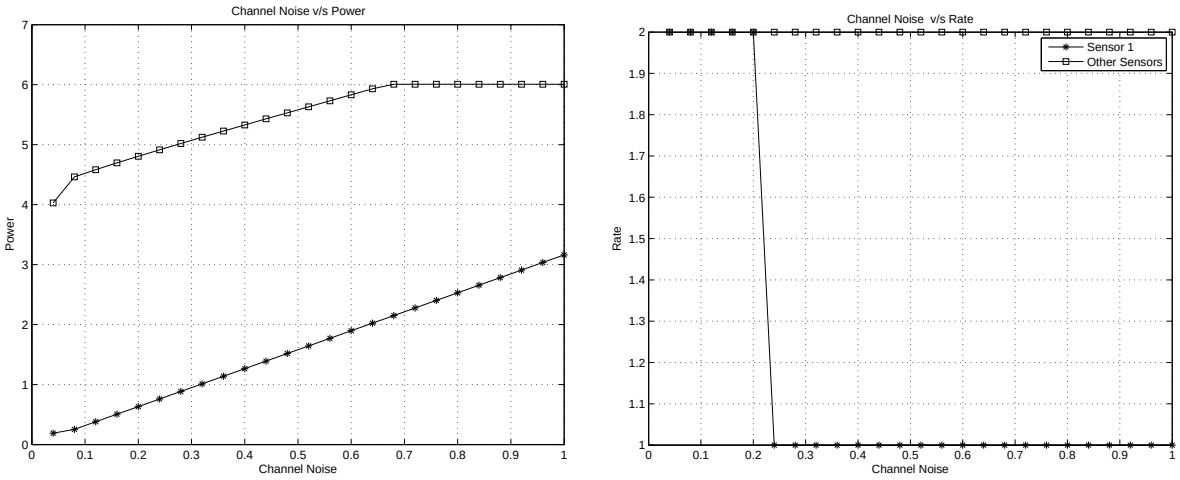
We also observe that -

1. The Lagrange multiplier associated with the BLUE variance constraint η_1 increases with increase in the magnitude of Λ_1. The large magnitude of this multiplier coupled with its sharp increase with Λ_1 explains why other collaborating sensors strongly pitch in to compensate for the deteriorating state of sensor $i = 1$.

2. η_1 is significantly larger than the Lagrange associated with total rate constraint η_{N+2} for Scheme 1 (BPSK). Thus changing the value of $D_{sys}^{\{max\}}$ significantly alters optimal operation of the active sensors while the rate constraint can be changed without incurring any pronounced shift in optimality. However for Scheme 2 (QAM), the Lagrange η_{N+2} for the rate constraint is comparable in magnitude to η_1.

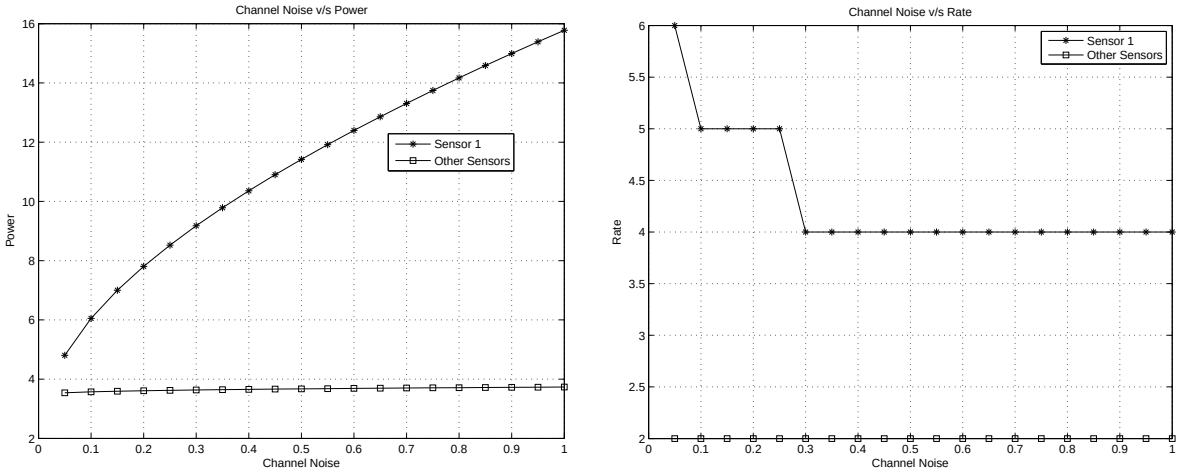
Optimal Actions and Channel Conditions - We vary channel condition $\frac{n_1}{|h_1|^2}$, $(\overline{|h_1|^2} = 1)$ for $i = 1$ and study the variation of $p_i^*, b_i^*, \forall i \in \{1, \ldots, N\}$ for $(\Lambda_1 < \Lambda_i, \forall i \in \{2, \ldots, N\}$; comparable
$\Lambda_i, \forall i \in \{1, \ldots, N\}; \Lambda_1 > \Lambda_i, \forall i \in \{2, \ldots, N\})$. The figures in 4.3 and 4.4 are plotted for comparable values of $\Lambda_i, \forall i \in \{1, \ldots, N\}$ for Scheme 1 (BPSK) and Scheme 2 (QAM) respectively. Optimal Power per bit follows reverse water-filling in general for all cases of $\Lambda_i, \forall i \in \{1, \ldots, N\}$ as the sensor with poor channel condition expends more power to satisfy the BLUE variance constraint at the fusion center. Another interesting aspect is that with deterioration in channel quality, the collaborating sensors step up their total power usage for information transmission unlike in the case of *OR-A*. This trend ensures that the BLUE variance constraint at the fusion center is satisfied even in the face of bad channel conditions experienced by any of the active sensors. Optimal information length b_i^* is seen to inversely depend on the channel condition metric $n_i, \forall i \in \{1, \ldots, N\}$.

Other relevant observations deductions are as follows -



(a) (b)

Figure 4.3: Scheme 1 (BPSK) - (a) Variation of $\frac{n_1}{|h_1|^2}$ with the Power used by Sensor (i=1) and Collaborating Mid-Range Healthy Sensors. (b)Variation of $\frac{n_1}{|h_1|^2}$ with the Bits used by Sensor (i=1) and Collaborating Mid-Range Healthy Sensors.



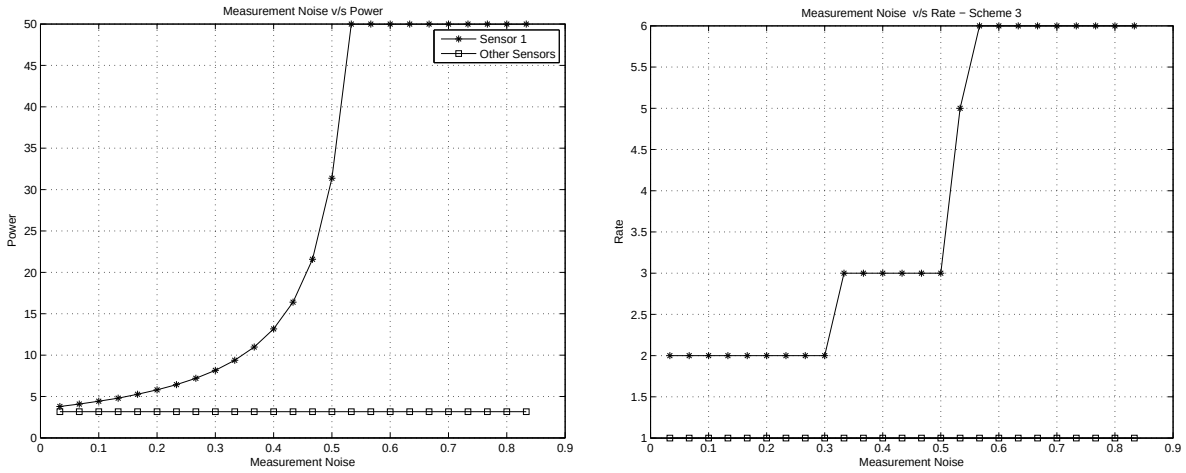
(a) (b)

Figure 4.4: Scheme 2 (QAM) - (a) Variation of n_1 with the Power used by Sensor (i=1) and Collaborating Mid-Range Healthy Sensors. (b)Variation of n_1 with the Bits used by Sensor (i=1) and Collaborating Mid-Range Healthy Sensors.

1. The Lagrange multiplier η_1 associated with the BLUE variance constraint is significantly large, which in turn forces a collaboration behavior as discussed. η_1 is seen to increase with increase in the magnitude of n_i, for any $i \in \{1, \ldots, N\}$.

2. It can be deduced that if any of the active sensors have a poor channel, then there is an increase in the total amount of power expended by all active sensors in the system to meet the BLUE variance requirement at the fusion center.

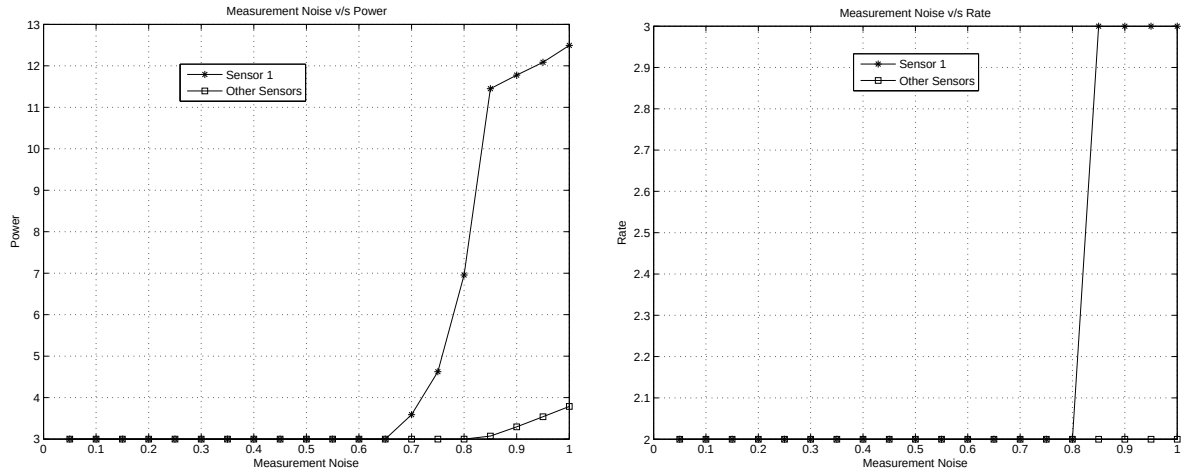
Optimal Actions and Measurement Noise Variance - The measurement noise $R_i, i = 1$ is varied for different operating states of the active sensors ($\Lambda_1 < \Lambda_i, \forall i \in \{2, \ldots, N\}$; comparable $\Lambda_i, \forall i \in \{1, \ldots, N\}; \Lambda_1 > \Lambda_i, \forall i \in \{2, \ldots, N\}$). We observe that for $\Lambda_1 < \Lambda_i, \forall i \in \{2, \ldots, N\}$; comparable Λ_i, the optimal actions of sensor $i = 1$ increases while that of the collaborating sensors remain constant as in figures 4.5 and 4.6 for Scheme 1 (BPSK) and Scheme 2 (QAM) respectively. This reflects the attempt of sensor $i = 1$ to satisfy the BLUE variance constraint by increasing its total amount of power for information transmission given its degrading measurement noise variance R_1.



(a) (b)

Figure 4.5: Scheme 1 (BPSK) - (a) Variation of R_1 with the Power used by Sensor (i=1) and Collaborating Unhealthy Sensors. (b) Variation of R_1 with the Bits used by Sensor (i=1) and Collaborating Unhealthy Sensors.

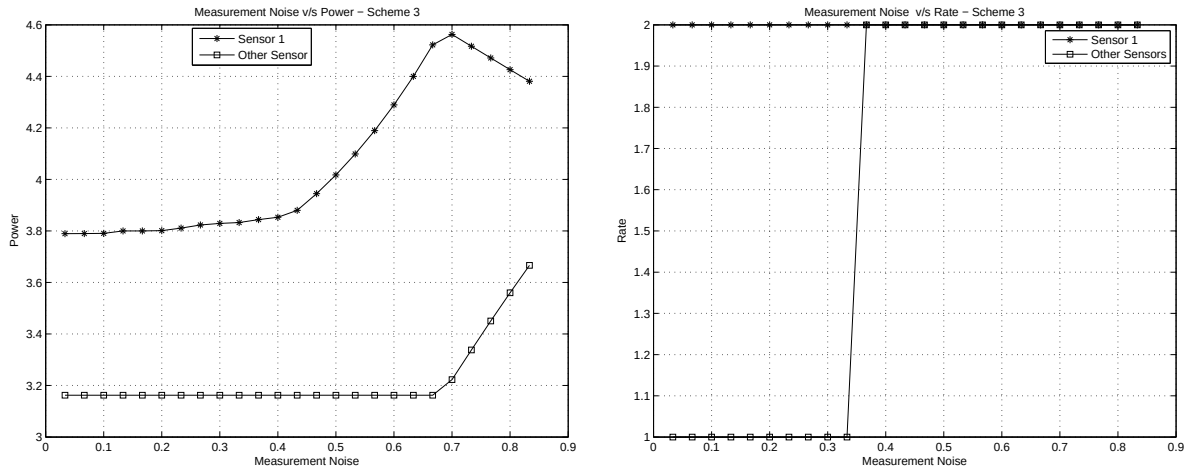
However, for the case $\Lambda_1 > \Lambda_i, \forall i \in \{2, \ldots, N\}$) in Scheme 1 (BPSK), we observe that this behavior reverses where the collaborating sensors begin using more resources causing the resource expenditure of sensor $i = 1$ to decrease. This is



(a) (b)

Figure 4.6: Scheme 2 (QAM) - (a) Variation of R_1 with the Power used by Sensor (i=1) and Collaborating Unhealthy Sensors. (b) Variation of R_1 with the Bits used by Sensor (i=1) and Collaborating Unhealthy Sensors.

illustrated with the figures in 4.7 for Scheme 1.



(a) (b)

Figure 4.7: Scheme 1 (BPSK) - (a) Variation of R_1 with the Power used by Sensor (i=1) and Collaborating Unhealthy Sensors. (b) Variation of R_1 with the Bits used by Sensor (i=1) and Collaborating Unhealthy Sensors.

Unlike for Scheme 2 (QAM), in Scheme 1 (BPSK), we note that η_1 is significant in magnitude compared to η_{N+2}. Also for both schemes the magnitude of η_1

increases sharply with increase in the measurement noise variance of any of the active sensors.

4.2 Summary

In this chapter, we presented an optimization formulation where we minimized the total resource expenditure of all active sensors in the network subject to a BLUE variance constraint at the fusion center. We analyzed the tradeoff between estimation error and resource utilization, and also optimal sensor actions. We investigated the dependency of optimal sensor actions on channel conditions, operating state, and measurement noise. From simulations, we inferred that the nonconvex problem in the original formulation can be approximated as the convex problem in CA-B without affecting optimality.

Chapter 5

Joint Estimation Error and Resource Utilization Minimization

In this chapter, we consider the third optimization formulation for Scheme 1 (BPSK) and Scheme 2 (QAM) - Formulation C. Here we seek to minimize the estimation error at the fusion center and the total amount of resources utilized simultaneously, and determine optimal sensor actions. We adopt the system model presented in chapter 3. We present the problem in section 5.1 followed by KKT analysis, results and discussion.

5.1 Formulation C - Minimize D and J Simultaneously.

We consider the problem of simultaneously minimizing the BLUE variance at the fusion center and the total resources expended by all active sensors in the system. The question that we address is as follows -

(Q3) What are the optimal sensor actions (power per bit and the number of bits) in an information transmission that jointly minimizes the total resource utilization of all active sensors and the variance of the best linear unbiased estimator at the fusion center?

The ensuing formulation is a generalization of $N = 2$ sensor case that is considered in [25] for both Scheme 1 (BPSK) and Scheme 2 (QAM). We use a similar expression for BLUE variance as previously and model resource utilization as $J = \sum_{i=1}^{N} \frac{\Lambda_i p_i b_i}{p_{max}^{(i)} BW}$ for Scheme 1, and $J = \sum_{i=1}^{N} \frac{\Lambda_i p_i c_i}{p_{max}^{(i)} BW}$ for Scheme 2, where, $\Lambda_i, \forall i \in \{1, , \ldots, N\}$ are the scalarization parameters in the multi-objective utility function. It can be treated as a parameter that reflects the resource policy of each sensor based on their operating state: If a sensor is operating on low residual battery power, it can choose a high Λ_i value thereby showing greater affinity for resource economical actions that help in conserving its battery power. Smaller values of Λ_i indicating higher residual battery power imply that the sensors weigh the objective of estimation error reduction over resource optimization. We express the problem in its standard form as -

$$\text{Minimize } J + D$$

$$\text{Subject to}$$

$$\sum_{i=1}^{N} b_i - BW \leq 0, \qquad \text{- Rate Constraint}$$

and, the box constraints from Formulation A follow implicitly. We refer to the above formulation in (5.1) as *OR-C*. We set $x_i = \frac{1}{(R_i + \frac{W^2}{3(2^{b_i}-1)^2} + R_{i,c})}$, and we have the modified formulation written in the D.C. form (*DC-C*) as -

$$\text{Minimize } -\sum_{i=1}^{N} x_i + \sum_{i=1}^{N} \Lambda_i p_i b_i,$$

$$\text{subject to } R_i + R_{i,q} + R_{i,c} - \frac{1}{x_i} \leq 0$$

$$\sum_{i=1}^{N} b_i - BW \leq 0,$$

The convex approximated version *CA-C* of *DC-C* can be written as follows -

$$\text{Minimize } -\sum_{i=1}^{N} x_i + \sum_{i=1}^{N} \Lambda_i p_i b_i,$$
$$\text{subject to}$$
$$R_i + R_{i,q} + R_{i,c} - \left(\frac{1}{x_i^{\{m\}}} + \frac{x_i - x_i^{\{m\}}}{(x_i^{\{m\}})^2} \right) \leq 0$$
$$\sum_{i=1}^{N} b_i - BW \leq 0,$$

where, the first-order Taylor expansion of the function $-\frac{1}{x_i} \forall i \in \{1, \ldots, N\}$ about $x_i^{\{m\}}$ has been used.

5.1.1 Analysis

We first determine the KKT conditions satisfied by the minima for the formulation in *DC-C* for both schemes. The Lagrangian $G_{1.c}$ for Scheme 1 (BPSK) in this formulation is as follows -

$$G_{1.c} = \sum_{i=1}^{N} \Delta_i p_i b_i - \sum_{i=1}^{N} x_i + \sum_{i=1}^{N} \eta_i \left(R_i \right.$$
$$\left. + \frac{W^2}{3(2^{b_i} - 1)^2} + \frac{4W^2}{3} \left(1 - \sqrt{\frac{p_i \overline{|h_i|^2}}{p_i \overline{|h_i|^2} + n_i}} \right) - \frac{1}{x_i} \right) +$$
$$\eta_{N+1} \left(\sum_{i=1}^{N} b_i - BW \right) + \sum_{i=1}^{N} \eta_{N+2+i}(-b_i + 1) +$$
$$\sum_{i=1}^{N} \eta_{2N+1+i}(p_i - p_i^{(max)}) + \sum_{i=1}^{N} \eta_{3N+1+i}(-p_i + p_i^{(min)}),$$

where, $\eta = [\eta_1, \ldots, \eta_{4N+1}]^T$ are the Lagrange multipliers that necessarily exist such that -

$$\begin{aligned}
&\frac{\partial G_{1.c}}{\partial p_k} = \Lambda_k b_k^* - \left(\frac{4W^2}{3} \frac{n_k \sqrt{\overline{|h_k|^2}}}{(p_k^* \overline{|h_k|^2} + n_k)^{3/2} p_k^{*1/2}} \right) \\
&+\eta_{2N+1+k} - \eta_{3N+1+k} = 0 \\
&\frac{\partial G_{1.c}}{\partial b_k} = \Lambda_k p_k^* - \frac{W^2 \ln 2}{3(2^{b_k^*} - 1)^2} \\
&+\eta_{N+1} + \eta_{N+1+k} = 0 \\
&\frac{\partial G_{1.c}}{\partial x_k} = -1 + \frac{\eta_k}{x_k^{*2}} = 0
\end{aligned} \tag{5.1}$$

Similarly for Scheme 2 in *DC-A*, the Lagrangian $G_{2.c}$ is as follows -

$$\begin{aligned}
G_{2.c} &= \sum_{i=1}^{N} \Delta_i B_s K_i a_i \ln 2 p_b (2^{b_i} - 1) - \sum_{i=1}^{N} x_i \\
&+ \sum_{i=1}^{N} \eta_i (R_i + \frac{W^2}{3(2^{b_i} - 1)^2} + \frac{16W^2}{3} \exp \sqrt{\frac{3B_s K_i a_i \log\left(\frac{2}{p_b}\right)}{2N_0}} - \frac{1}{x_i}) + \\
&\eta_{N+1} \left(\sum_{i=1}^{N} b_i - BW \right) + \sum_{i=1}^{N} \eta_{N+1+i} (-b_i + 2),
\end{aligned} \tag{5.2}$$

where, $\eta = [\eta_1, \ldots, \eta_{2N+1}]^T$ are the Lagrange multipliers that necessarily exist such that -

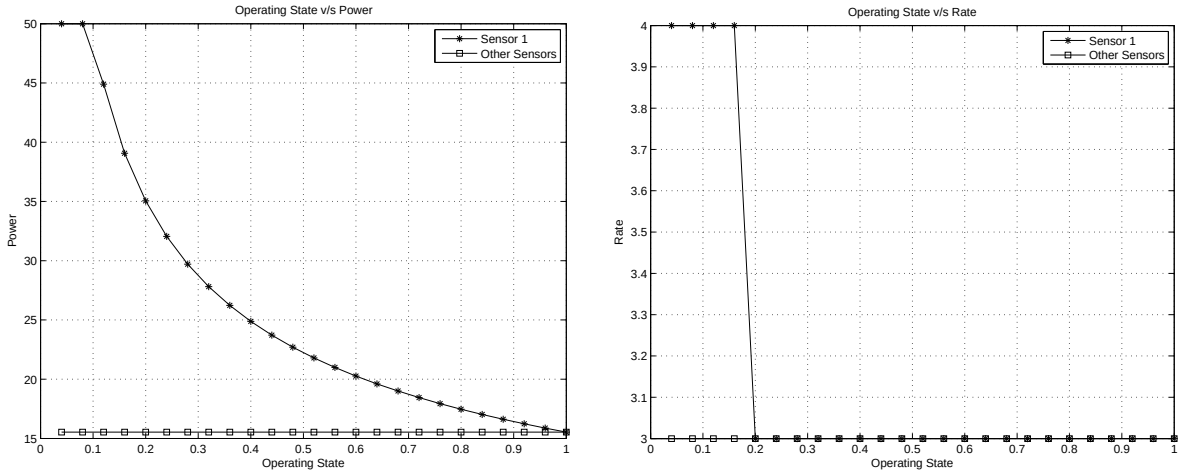
$$\begin{aligned}
&\frac{\partial G_{2.c}}{\partial b_k} = \Delta_k B_s K_k a_k \log\left(\frac{2}{p_b}\right) 2^{b_k^*} \ln 2 \\
&-\eta_k \left(\frac{W^2 \ln 2}{3(2^{b_k^*} - 1)^2} + \frac{16W^2}{3 b_k^{*2}} \exp \sqrt{\frac{3B_s K_i a_i \log\left(\frac{2}{p_b}\right)}{2N_0}} \right) + \eta_{N+1} - \eta_{N+1+k} = 0 \\
&\frac{\partial G_{1.a}}{\partial x_k} = -1 + \frac{\eta_k}{x_k^{*2}} = 0
\end{aligned} \tag{5.3}$$

5.1.2 Results

We consider a network with system parameters similar to that described in the previous chapter and perform numerical simulations for both schemes in formulations *DC-C*, *CA-C* and an exhaustive search for the global minimizers of the problem in *OR-C*.

The minimizers found in all three cases for each scheme are seen to be the same implying that the non-convex formulation in *OR-C* and *DC-C* can be reduced to its convex formulation in *CA-C*. We study the variation of optimal $p_i, b_i, x_i, \forall i \in \{1, \ldots, N\}$ for different values of $\Lambda_i, \forall i \in \{1, \ldots, N\}$. In general, we observe the following -

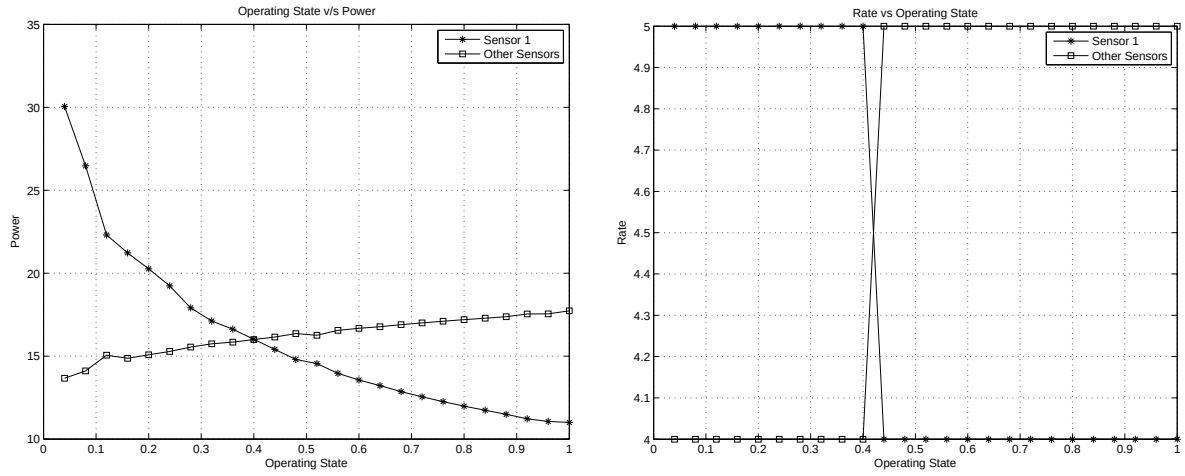
Optimal Actions and Operating State - The optimal actions $p_i, b_i, \forall i \in \{1, \ldots, N\}$ relates to the operating state $\Lambda_i, \forall i \in \{1, \ldots, N\}$ inversely as in figures 5.1 and 5.2 for Scheme 1 (BPSK) and Scheme 2 (QAM) respectively. Note that the figures present the results for mid-range healthy sensors collaborating with sensor $i = 1$; i.e. $\Lambda_i, \forall i \in \{2, \ldots, N\}$ around 0.5.



(a) (b)

Figure 5.1: Scheme 1 (BPSK) - (a) Variation of Λ_1 with the Power used by Sensor (i=1) and Collaborating Mid-Range Healthy Sensors. (b) Variation of Λ_1 with the Bits used by Sensor (i=1) and Collaborating Mid-Range Healthy Sensors.

As Λ_i increases, the corresponding sensor shows greater affinity for reducing resource expenditure by significantly reducing the information length b_i^* and the power associated with each bit in the information p_i^*. Such actions results in increase in variance in the information transmitted due to imperfect channel and quantization thereby increasing the BLUE variance at the fusion center. The converse is observed for lower values of Λ_i when the corresponding sensor prefer actions that minimize the



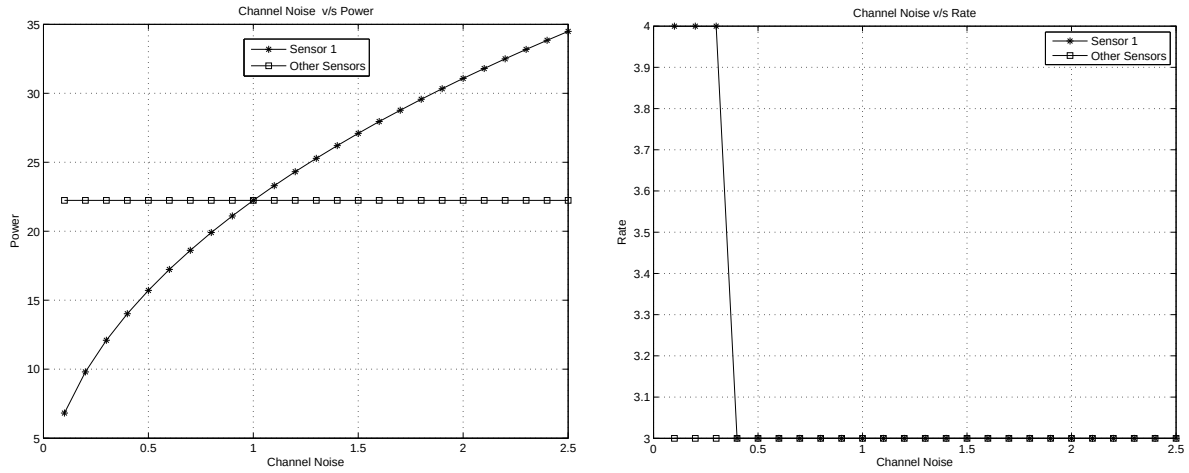
(a) (b)

Figure 5.2: Scheme 2 (QAM) - (a) Variation of Λ_1 with the Power used by Sensor (i=1) and Collaborating Mid-Range Healthy Sensors. (b) Variation of Λ_1 with the Bits used by Sensor (i=1) and Collaborating Mid-Range Healthy Sensors.

variance of the information at the fusion center. In general, there is no collaboration observed for Scheme 1. But for Scheme 2, a decrease in the resource expended by an active sensor (due to its deteriorating operating state) is compensated by an increase in the resource spending of the collaborating sensors only if they are healthy.

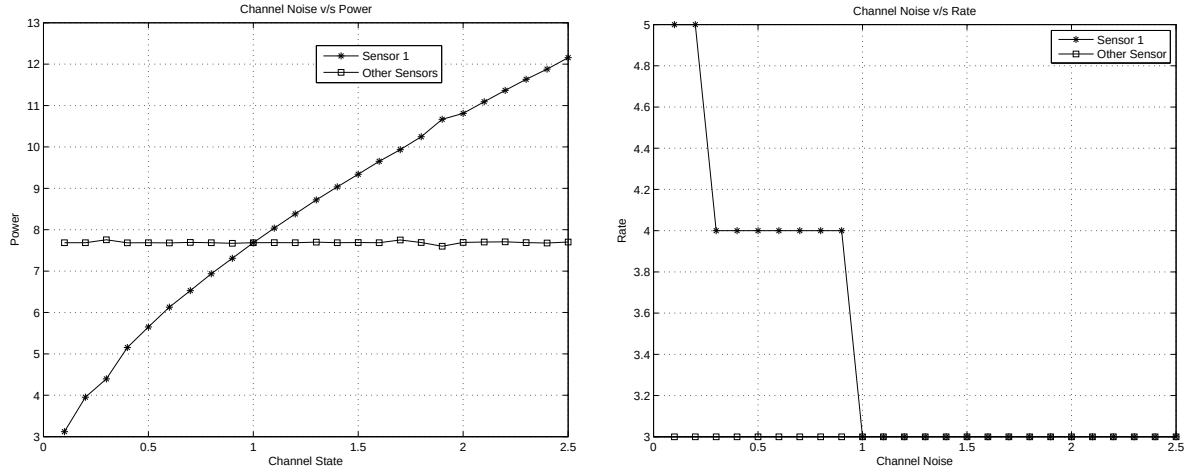
Optimal Actions and Channel Conditions - As expected for Scheme 1 (BPSK), $p_i^*, \forall i \in \{1, \ldots, N\}$ relates to the channel condition metric $\frac{n_i}{|h_i|^2}$ ($\overline{|h_i|^2} = 1$) directly indicating a reverse-water filling tendency as in figure 5.3. The information length $b_i^*, \forall i \in \{1, \ldots, N\}$ is seen to depend inversely on the channel condition as in figure 5.3. Similar trends in the variation of p_i^*, b_i^*, with $n_i, \forall i \in \{1, \ldots, N\}$ are observed for Scheme 2 (QAM) in figure 5.4. In general, there is no collaboration observed for Scheme 1. But for Scheme 2, when an active sensor's channel condition deteriorates, the collaborating sensors increase their resource spending in case they are healthy or $\Lambda_i, i \in \{2, \ldots, N\}$ is close to zero.

Optimal Actions and Measurement Noise Variance - We observe that $p_i, b_i, \forall i \in \{1, \ldots, N\}$ depends on $R_i, \forall i \in \{1, \ldots, N\}$ inversely implying that an increase in the



(a) (b)

Figure 5.3: Scheme 1 (BPSK) - (a) Variation of $\frac{n_1}{|h_1|^2}$ with the Power used by Sensor (i=1) and Collaborating Mid-Range Healthy Sensors. (b) Variation of $\frac{n_1}{|h_1|^2}$ with the Bits used by Sensor (i=1) and Collaborating Mid-Range Healthy Sensors.

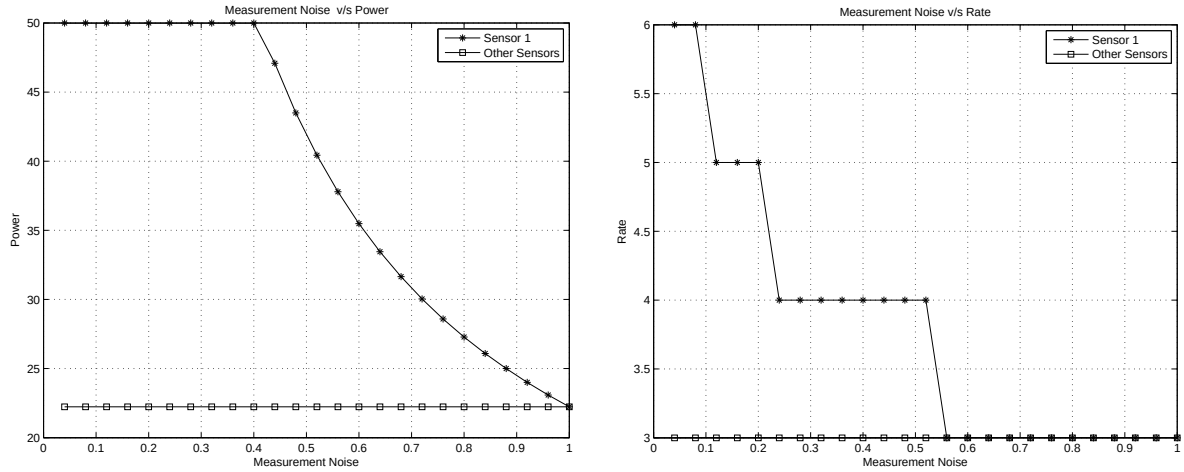


(a) (b)

Figure 5.4: Scheme 2 (QAM) - (a) Variation of n_1 with the Power used by Sensor (i=1) and Collaborating Mid-Range Healthy Sensors. (b) Variation of n_1 with the Bits used by Sensor (i=1) and Collaborating Mid-Range Healthy Sensors.

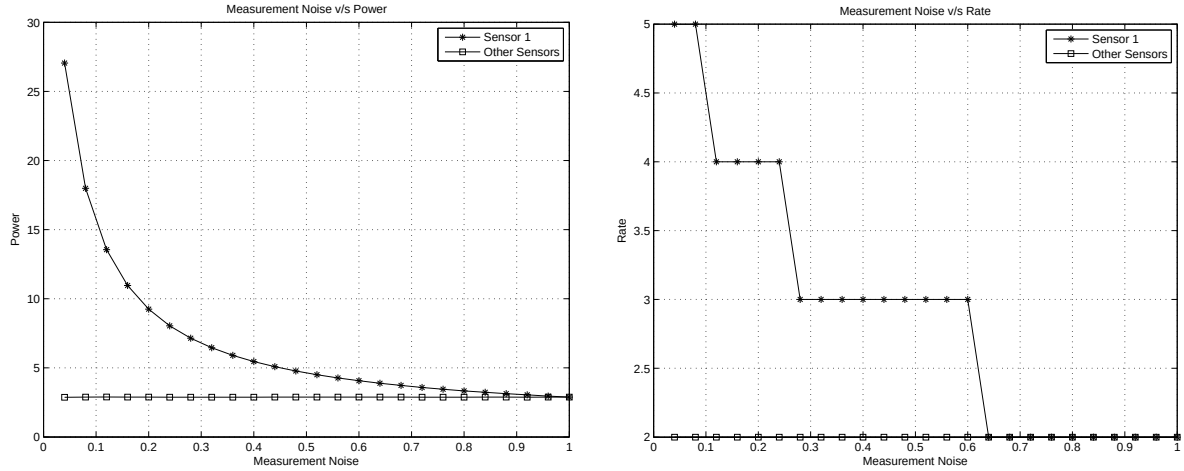
measurement noise variance prompts the corresponding sensor to utilize lesser amount of resources for transmission of information to the fusion center. This is illustrated in figures 5.5 and 5.6 for Scheme 1 (BPSK) and Scheme 2 (QAM) respectively. When

the measurement noise of any of the active sensor increases, we observe no collaboration for Scheme 1; but for Scheme 2, there is an increase in resource expenditure for the collaborating sensors if they are in good health.



(a) (b)

Figure 5.5: Scheme 1 (BPSK) - (a) Variation of R_1 with the Power used by Sensor (i=1) and Collaborating Unhealthy Sensors. (b) Variation of R_1 with the Bits used by Sensor (i=1) and Collaborating Unhealthy Sensors.



(a) (b)

Figure 5.6: Scheme 2 (QAM) - (a) Variation of R_1 with the Power used by Sensor (i=1) and Collaborating Unhealthy Sensors. (b) Variation of R_1 with the Bits used by Sensor (i=1) and Collaborating Unhealthy Sensors.

For Scheme 1 (BPSK), we observe that the magnitude of the Lagrange associated with the total rate constraint η_{N+1} is significantly low implying that BW can increased or decreased without changing the optimal values of $p_i, b_i, x_i, \forall i \in \{1, \ldots, N\}$. As a result, we observe that collaborative behavior among the active sensors is almost absent. This directly motivates the idea of decoupling the problem in *CA-C* for Scheme 1 using dual decomposition and formulating a system with purely distributed autonomous operation -

$$\begin{aligned}
&\text{Minimize } \sum_{i=1}^{N} \Delta_i p_i b_i - \sum_{i=1}^{N} x_i + \sum_{i=1}^{N} \eta_i \left(R_i \right. \\
&\left. + \frac{W^2}{3(2^{b_i}-1)^2} + \frac{4W^2}{3}\left(1 - \sqrt{\frac{p_i|\overline{h_i}|^2}{p_i|\overline{h_i}|^2 + n_i}}\right) - \left(\frac{1}{x_i^{\{m\}}} + \frac{x_i - x_i^{\{m\}}}{(x_i^{\{m\}})^2}\right)\right) + \\
&\eta_{N+1}\left(\sum_{i=1}^{N} b_i - BW\right), \\
&\text{subject to} \\
&-b_i + 1 \leq 0 \\
&p_i - p_{max}^{(i)} \leq 0; \; -p_i + p_{min}^{(i)} \leq 0; \forall i \in \{1, \ldots, N\},
\end{aligned}$$

We set $\eta_{N+1} = 0$ and let $f_i(p_i, b_i, x_i) = \Delta_i p_i b_i - \sum_{i=1}^{N} x_i + \eta_i \left(R_i + \frac{W^2}{3(2^{b_i}-1)^2} + \frac{4W^2}{3} P_b^{(i)} - \left(\frac{1}{x_i^{\{m\}}} + \frac{x_i - x_i^{\{m\}}}{(x_i^{\{m\}})^2}\right)\right)$. This enables to achieve purely autonomous operation of the sensors in the network where each sensor $i \in \{1, \ldots, N\}$ performs local convex optimization as follows -

$$\begin{aligned}
&\text{Minimize } f_i(p_i, b_i, x_i) \\
&\text{subject to} \\
&-b_i + 1 \leq 0 \\
&p_i - p_{max}^{(i)} \leq 0; \; -p_i + p_{min}^{(i)} \leq 0;
\end{aligned}$$

Since the total rate constraint of the system has been removed, the above formulation does not involve a master problem (problem solved at the fusion center in this case)

that updates η_{N+1} and signals the updated η_{N+1} value to the distributed sensors. The problem in *CA1.c* is convex and hence we have that the decomposed problem above in its dual form must has zero duality gap and hence give the same solution as the primal problem. However for Scheme 2 (QAM), the significant magnitude of the lagrange multiplier associated with the total rate constraint sans pure decoupling.

5.2 Summary

In this chapter, we presented an optimization formulation where we minimized the BLUE variance at the fusion center and the total resource expenditure of all active sensors in the network. We analyzed the tradeoff between estimation error and resource utilization, and also optimal sensor actions. We investigated the dependency of optimal sensor actions on channel conditions, operating state, and measurement noise. From simulations, we inferred that the non-convex problem in the original formulation can be approximated as the convex problem in CA-C without affecting optimality. We finally presented the notion of achieving perfectly autonomous operation of the sensors in the network in the case of Scheme 1 (BPSK) since the Lagrange multiplier associated with the total rate constraint is very small in magnitude.

Chapter 6

Comparative Analysis

In this chapter, we comment on the total amount of resources utilized and the BLUE variance in each of the formulations for both Scheme 1 (BPSK) and Scheme 2 (QAM). In section 6.1, we compare the formulations in terms of estimation error achieved and the resources utilized to achieve that error. In section 6.2, we compare the optimal sensor actions, and collaborative behavior among sensors for each of the formulations.

6.1 General Comparison - Formulations

We set $N = 5$, $BW = 20$, minimum system SNR requirement as 5dB, $p_{max}^{(i)} = 20$ units, $\forall i \in \{1, \ldots, N\}$ and $E_{sys}^{\{max\}} = 20$ for Formulation A, and $D_{sys}^{\{max\}} = 0.35$ for Formulation B. We analyze the BLUE variance achieved at the fusion center and the corresponding amount of resources utilized by first varying the operating state Λ_i of sensor $i = 1$, given that the collaborating sensors are healthy and that the channel conditions and measurement noise are maintained constant. This is illustrated in figures 6.1, 6.2, and 6.3 for Scheme 1 (BPSK) using Formulations A, B, and C respectively. It can be seen that Formulation B is the most economical approach in terms of resource consumed for a target BLUE variance. Formulation A

enables achieve high quality estimator but at the cost of excess amount of resources. Formulation C is apparently the least favorable in resources consumed and BLUE variance achieved. For e.g., - It can be seen that for the system under consideration, the minimum variance achieved by Formulation A is $D = 0.359$ corresponding to $J = 20$ (amount of resource used), while in the worst case $D = .404$ is attained for $J = 20$. Formulation B achieves a variance of $D = 0.35$ for $J = 14$ in the best case and $D = .35$ for $J = 19.1$ in the worst case. In Formulation C, Pareto optimality is observed, with the curve ranging from the best case variance of $D = 0.36$ for $J = 24$ to the worst case variance of $D = .393$ for $J = 23$. i.e., in effect, Formulation B uses 42% lesser amount of resources than Formulation A, and 71% lesser amount of resources than Formulation C to achieve the same BLUE variance in the best case scenario. However Formulation C may be suitable for achieving a completely autonomous distributed estimation process as discussed in the previous chapter. We make similar observations when the channel quality of sensor $i = 1$ is varied, keeping the operating state and measurement noise of all sensors a constant; and when measurement noise of sensor $i = 1$ is varied, keeping operating states and channel quality of all sensors a constant.

Using the same system parameters, we analyze all formulations using Scheme 2 (QAM) along the same lines as that of Scheme 1. Figures 6.4, 6.5, and 6.6 are plotted for Formulations A, B, and C respectively, by varying the operating state Λ_i of sensor $i = 1$, given that the collaborating sensors are healthy and that the channel conditions and measurement noise are maintained constant. We make similar conclusions as in Scheme 1. However for Scheme 2, pure decoupling of formulation C is not possible because of the significant magnitude of Lagrange multiplier associated with the total rate constraint.

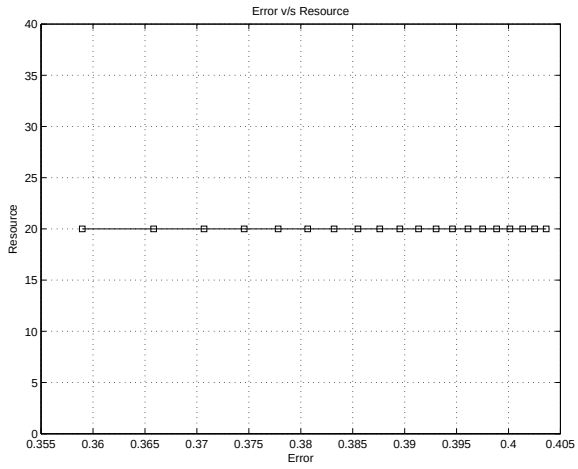


Figure 6.1: Scheme 1 (BPSK) Formulation A - BLUE Variance vs Total Power used by Active Healthy Sensors.

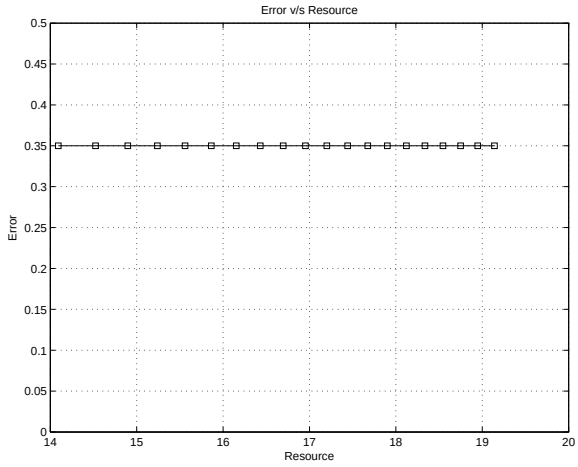


Figure 6.2: Scheme 1 (BPSK) Formulation B - BLUE Variance vs Total Power used by Active Healthy Sensors.

6.2 Comparison - Optimal Actions and Collaboration

Effect of Operating State on Optimal Actions and Collaboration - For both Scheme 1 (BPSK) and Scheme 2 (QAM) in all formulations, the optimal power per transmission and rate of any sensor i depends inversely on its operating state (characterizing

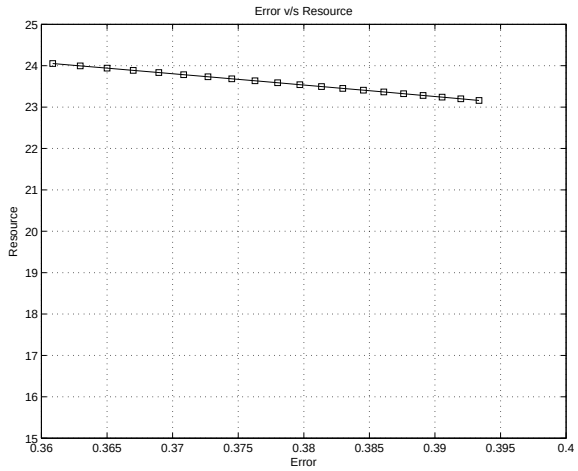


Figure 6.3: Scheme 1 (BPSK) Formulation C - BLUE Variance vs Total Power used by Active Healthy Sensors.

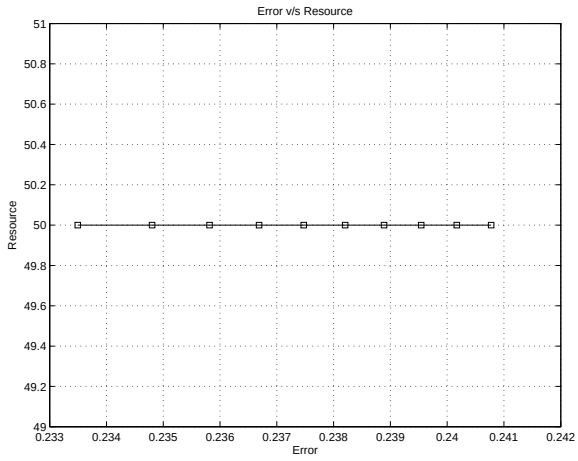


Figure 6.4: Scheme 2 (QAM) Formulation A - BLUE Variance vs Total Power used by Active Healthy Sensors.

its residual battery power) $\Lambda_i, \forall i \in \{1, \ldots, N\}$. In Formulation A, when an active sensor's operating state deteriorates, the collaborating sensors (irrespective of their operating states) decrease their optimal power and rate and hence do not collaborate to minimize BLUE variance at the fusion center. However in Formulation B, when a sensor's operating state deteriorates, the other collaborating sensors collaborate by

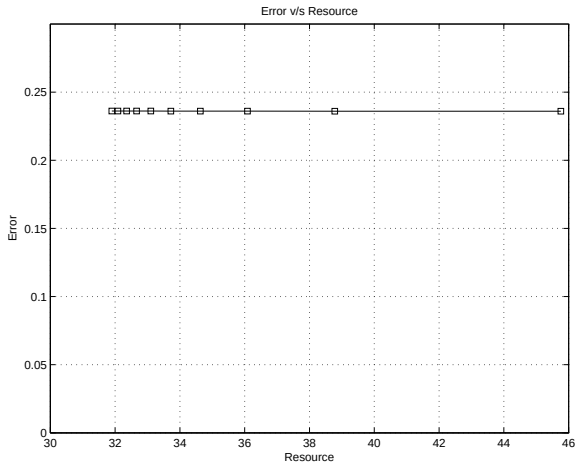


Figure 6.5: Scheme 2 (QAM) Formulation B - BLUE Variance vs Total Power used by Active Healthy Sensors.

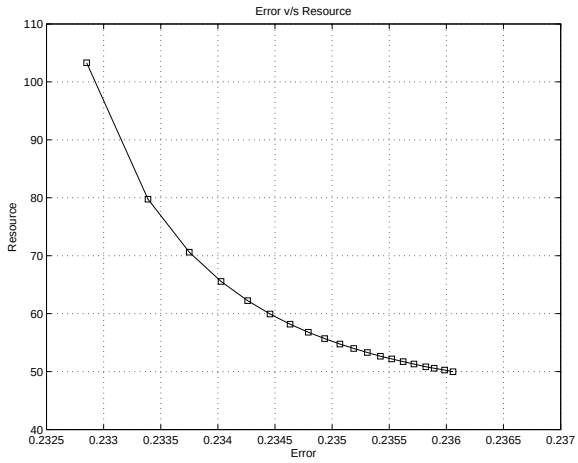


Figure 6.6: Scheme 2 (QAM) Formulation C - BLUE Variance vs Total Power used by Active Healthy Sensors.

expending more resources to minimize the BLUE variance at the fusion center. In Formulation C, there is no collaboration observed for Scheme 1 since the Lagrange multiplier associated with the rate constraint is very small in magnitude. However, for Scheme 2, when the operating state of any of the active sensors deteriorates, we observe an increase in resource expenditure for the collaborating sensors if they are

in good health.

Effect of Channel Condition on Optimal Actions and Collaboration - In all formulations for both schemes, the optimal power per transmission depends directly on channel noise power; i.e. a reverse water-filling trend is observed. However the optimal length of information transmitted (also the rate) depends inversely on channel noise power. In Formulation A, if any of the active sensors experience degraded channel condition, then the collaborating sensors step down their resource utilization. This is unlike in Formulation B, where collaborating sensors increase their resource consumption in order to minimize the BLUE variance at the fusion center. In Formulation C for Scheme 1, no collaboration is observed. But for Scheme 2 in Formulation C, an increase in resource spending of the collaborating sensors is observed when they are healthy and if the channel quality of any of the active sensors degrades.

Effect of Measurement Noise on Optimal Actions and Collaboration - For both schemes in Formulation A, when the measurement noise of an active sensor increases, there is a decrease in its resource spending and an increase in the resources utilization for the collaborating sensors. In Formulation B, when the measurement noise of an active sensor increases, its resource spending increases only if it is healthy while the optimal actions of the collaborating sensors remain unchanged. But if any of the active sensor is unhealthy, then its resource spending decreases coupled with an increase in the resource spending of other collaborating sensors. In Formulation C for Scheme 1, there is a decrease in the resource expended by an active sensor when its measurement noise increases, but there is no change in the actions of the collaborating sensors. However for Scheme 2, there is an increase in the resource spending of the collaborating sensors only if they are healthy.

The optimal actions of the active sensors in the networks for both schemes are summarized in tables 6.1, 6.2, and 6.3 (where $\uparrow$ indicates an increase and $\downarrow$ implies a decrease in value).

Table 6.1: Formulation A - Results

	Operating State $\Lambda_1 \uparrow$	Channel Condition $n_1 \uparrow$	Measurement Noise $R_1 \uparrow$
S-1 Power [1]	$\downarrow$	$\uparrow$	$\downarrow$
S-1 Bits	$\downarrow$	$\downarrow$	$\downarrow$
C-S Power [2]	$\downarrow$	$\downarrow$	$\uparrow$
C-S Bits	$\downarrow$	$\downarrow$	$\uparrow$
Resource [3]	$\uparrow$	$\uparrow$	$\downarrow$
Rate [4]	$\uparrow$	$\uparrow$	$\downarrow$

Table 6.2: Formulation B - Results

	Operating State $\Lambda_1 \uparrow$	Channel Condition $n_1 \uparrow$	Measurement Noise $R_1 \uparrow$
S-1 Power	$\downarrow$	$\uparrow$	$\uparrow$
S-1 Bits	$\downarrow$	$\downarrow$	$\uparrow$
C-S Power	$\uparrow$	$\uparrow$	$\downarrow$
C-S Bits	$\uparrow$	$\uparrow$	$\downarrow$
Variance Cost [5]	$\uparrow$	$\uparrow$	$\uparrow$
Rate Cost	$\uparrow$	$\uparrow$	$\uparrow$

6.3 Summary

In this chapter, we presented a comparative analysis of the estimation error and resource utilization achieved in the three formulations for both Scheme 1 (BPSK) and Scheme 2 (QAM). We also compared optimal sensor actions and collaborative behavior that resulted from the three optimization formulations. In the next chapter, we

[1] Sensor 1
[2] Collaborating Sensors
[3] Lagrange Multiplier associated with the Resource Constraint
[4] Lagrange Multiplier associated with the Rate Constraint
[5] Lagrange Multiplier associated with the BLUE Variance Constraint
[6] Refers to Scheme 2 (QAM), as collaboration is absent in Scheme 1 (BPSK)

Table 6.3: Formulation C - Results.

	Operating State $\Lambda_1 \uparrow$	Channel Condition $n_1 \uparrow$	Measurement Noise $R_1 \uparrow$
S-1 Power	$\downarrow$	$\uparrow$	$\downarrow$
S-1 Bits	$\downarrow$	$\downarrow$	$\downarrow$
C-S Power	$\downarrow$	$\downarrow$	$\uparrow$
C-S Bits [6]	$\downarrow$	$\downarrow$	$\uparrow$
Rate Cost	$\uparrow$	$\uparrow$	$\downarrow$

present a problem in distributed detection and signal processing in wireless biomedical sensors.

Chapter 7

Distributed Detection Application

Corruption of photopleythsmograms (PPGs) by motion artifacts has been a serious obstacle to the reliable use of pulse oximeters for real-time, continuous state-of-health monitoring. In this chapter, we illustrate an application of distributed detection and signal processing by proposing an automated, two-stage PPG data processing system to minimize the effects of motion artifacts. The technique is based on our prior work related to motion artifact detection (stage one) [30] and motion artifact reduction (stage two) [31]. These two steps are fundamental to the realization of a completely automated PPG processing system that would enable reliable continuous state-of-health monitoring of subjects. We introduce the area to the readers in section 7.1 and describe the proposed model in section 7.2.

7.1 Introduction

Photoplethysmography is a noninvasive, optical means to obtain relative blood volume in tissue as a function of time. Since hemoglobin is an optical absorber, light passing through tissue is modulated by each cardiac cycle of the subject and also by other processes like respiration and subject motion. The resulting photoplethysmograms can be acquired with reflectance- or transmittance-mode sensors, and multiple

excitation wavelengths allow waveform features from time-domain PPGs to be converted into values of heart rate and blood oxygen saturation. Corruption of PPGs by motion artifacts has been a significant obstruction to efficient and reliable use of pulse oximeters for continuous real-time health monitoring, especially in ambulatory settings [41]. If PPG data are to be reliably obtained from wearable sensors used for real-time, continuous state-of-health monitoring, then effective algorithms for motion artifact reduction must be employed.

7.1.1 Prior Works

Over the years, most of the PPG enhancement research has focused on motion artifact removal techniques. Various signal processing techniques have been investigated to address the problem of recovering quasiperiodic PPGs from data corrupted with motion artifacts. These include wavelet analysis and decomposition techniques [34] and adaptive filters [35]. However, the study in [36] indicates that both wavelet-transform and adaptive filter techniques introduce phase distortions in PPG data. Work involving analog filters and moving average techniques is presented in [37]. The artifact extraction problem has also been viewed as a blind source separation problem in [38], [41] and [39]. In [38], an enhanced preprocessing unit preceded the Independent Component Analysis (ICA) block. The preprocessing unit consisted of signal period detection using an autocorrelation method followed by a block-interleaving operation. However, this technique relies on the ability of the autocorrelation technique to correctly detect the waveform period and hence provides erroneous results in the presence of extreme motion artifacts. In [39], an improved preprocessing technique is described that employs extrapolation/truncation of each cardiac cycle to the mean of the measured cardiac cycle followed by ICA. This method is highly prone to errors and inconsistencies, since accurate cardiac cycle measurements become difficult in the presence of extreme motion. While removing motion artifact from PPGs is critical,

detecting its presence is a key task that needs to be addressed first. A reliable motion artifact detection technique lays the foundation for a completely automated PPG data processing system that can identify PPG data frames contaminated with artifacts and further process them for motion artifact removal. Some work has addressed the detection issue by correlating a PPG data frame with a clean reference signal to detect motion artifact [39]. However, such techniques are unsuitable for robust continuous monitoring where clean PPG signals and motion-corrupted signals are to be identified automatically in real-time implementation.

7.2 System Model

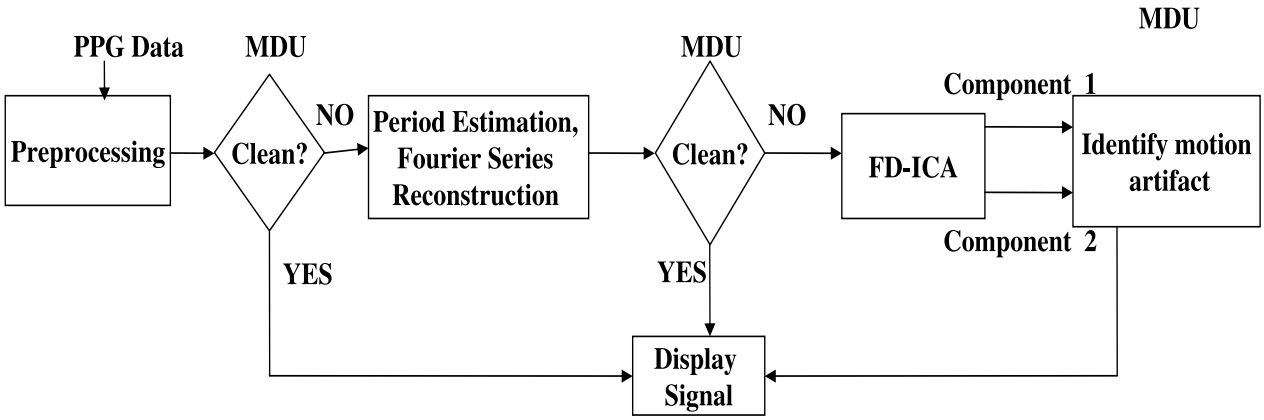


Figure 7.1: PPG Data Processing - System Model

The model for PPG data processing system is illustrated in figure 7.1 and is based on motion artifact detection (stage one), and motion artifact reduction (stage two).

Motion Artifact Detection - Stage One : Both clean and corrupt PPG data are analyzed in the time and frequency domains. In the time domain, skew and kurtosis measures of the signal are used as distinguishing metrics between clean and motion-corrupted data. In the frequency domain, the presence of random components due

to motion artifact is analyzed using a frequency domain kurtosis measure. Additionally, bispectral analysis of PPG data indicates the presence of strong quadratic phase coupling (QPC) and more specifically self coupling in the case of clean PPG data. Though quadratic phase coupling is found in data corrupted by motion artifact, the self coupling feature is absent. Then, a Neyman-Pearson (NP) detection rule is formulated for each of the measures. Additionally, treating each of the measures as observations from independent sensors, soft decision fusion from [33] and hard-fusion (Varshney-Chair rule) from [32] are used to fuse individual decisions to form a global system decision.

Motion Artifact Reduction - Stage Two : This stage involves an enhanced preprocessing unit consisting of a motion detection unit (MDU, based on stage one and developed in chapter 8), period estimation unit, and a Fourier series reconstruction unit. The MDU aids in identifying clean data frames versus those corrupted with motion artifacts. The period detection unit is used to determine the fundamental frequency of a corrupt frame. The Fourier series reconstruction unit reconstructs the final preprocessed signal. The reconstruction process primarily utilizes the spectrum variability of the pulse waveform. Preprocessed data are then fed to a magnitude-based frequency domain Independent Component Analysis (FD-ICA) unit. This helps reduce motion artifacts present at the frequency components chosen for reconstruction.

System Operation : PPG data obtained from a pulse oximeter are first filtered and detrended, which is fed as input to an enhanced preprocessing unit consisting of a motion detection unit (MDU, presented in chapter 8), period estimation unit, and Fourier series reconstruction unit. The MDU identifies clean PPG data frames versus those corrupted with motion artifacts. If the data are found to be clean, no further cleansing operations are imposed. If the data are corrupt, they are fed into the period estimation unit that determines the fundamental frequency of the corrupt frame. The

Fourier series reconstruction unit then reconstructs the final preprocessed signal by utilizing the spectrum variability of the pulse waveform yielding a signal composed primarily of its fundamental frequency component and harmonics. We then use the MDU to determine whether the reconstructed PPG signal is stained with motion artifacts. If the data are not free of motion artifacts, the magnitude-based frequency domain Independent Component Analysis (FD-ICA) routine is applied to estimate the blood volume pulsation and motion artifact components as described in chapter 9. This helps reduce motion artifacts present at the frequencies of the reconstruction components.

Chapter 8

Stage One - Motion Detection

In this chapter, we present methods for detecting the presence of motion artifact in photoplethysmographic (PPG) measurements based on higher order statistical information present in the data. The theory of HOS measures considered in this work is briefly discussed in 8.1. The results of PPG data analyses (with and without motion artifact) based on the above measures are presented in section 8.2. Based on these results, a Neyman-Pearson detection (NP) rule is formulated for each of the measures and discussed in section 8.3. Section 8.3.2 addresses the combination of all of these measures to formulate overall system decision.

8.1 Theory

The following subsections review the HOS measures applied to these PPG data [43]-[47]:

1. *Skew and Kurtosis* - Skew is a measure of the symmetry (or the lack of it) of a probability distribution, while the kurtosis measure indicates a heavy tail and peakedness *OR* a light tail and flatness of a distribution relative to the normal distribution. This measure captures the random variations of data from the

mean. The skew and kurtosis of a random variable x are given by

$$\begin{aligned} C_{3x}(0,0) &= \frac{\mu_3}{\sigma^{3/2}} \qquad \text{(skew)} \\ C_{4x}(0,0,0) &= \frac{\mu_4}{\sigma^4} - 3 \qquad \text{(kurtosis)} \end{aligned} \tag{8.1}$$

where σ is the standard deviation; μ_3 and μ_4 are the third and fourth moments.

2. *Bispectrum* - The third-order polyspectrum of a random variable x is defined as the Fourier transform of its third cumulant sequence :

$$S_{3x}(f_1, f_2) = \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} [C_{3x}(k,l) \exp(-j2\pi(f_1 k + f_2 l))] \tag{8.2}$$

where $C_{3x}(k,l)$ is the third cumulant sequence of x. The third-order polyspectrum, or the power spectrum, suppresses all phase information in a random process, while the bispectrum does not. When the harmonic components of a process interact, definitive phase relations also exist, in addition to the contribution of power at their sum and difference frequencies; this is called *Quadratic Phase Coupling* (QPC). For example, consider the following process:

$$X_1(k) = \cos(\lambda_1 k + \phi_1) + \cos(\lambda_2 k + \phi_2) + \cos(\lambda_3 k + \phi_3) \tag{8.3}$$

where $\lambda_3 = \lambda_1 + \lambda_2$, indicating that λ_1, λ_2 and λ_3 are harmonically related. If ϕ_1, ϕ_2 and ϕ_3 in (8.3) are independent random variables uniformly distributed in the range [0,2π], then (λ_3, ϕ_3) is an independent harmonic component. However, if $\phi_3 = \phi_1 + \phi_2$ in (8.3), then (λ_3, ϕ_3) is the result of quadratic coupling between (λ_1, ϕ_1) and (λ_2, ϕ_2).

8.2 PPG Data Analysis

PPG data analysis is performed to understand and extract features that can be used as distinguishing metrics between clean and motion-corrupted data. Initially, frames

of data are collected from healthy subjects in the age group of 22-24 years using a reflectance pulse oximeter [41]. The subjects follow the same motion patterns as in [41]:

1. Stationary Position: The subjects remain still with no wrist, finger or elbow movement.

2. Finger movements (three cases): left-right (swinging), up-down (bending), and arbitrary finger movements while keeping the wrist and elbow stationary.

3. Wrist movements: The wrist is rotated and arbitrarily moved, keeping the elbow and fingers stationary.

4. Elbow movements: The elbow is bent and extended, keeping the wrist and fingers stationary.

Data are fed into a MATLAB script that partitions the entire data segment into short frames of equal length. First, each frame is passed through a bandpass filter (0.3-12 Hz). Here, the design of the filter is critical, as the phase information in the data needs to be preserved to retain the shape of the PPG waveform. For this purpose, a zero-phase forward-reverse filter of order four in both directions is chosen. After filtering, the baseline trend associated with each data frame is removed by extracting an appropriately fitted polynomial curve. Each data frame is then inspected in the time and frequency domains, and the HOS properties are characterized.

8.2.1 Time Domain Analysis

In the time domain, we analyze the skew and kurtosis measure of the time variation of the amplitude of the PPG signal in each frame considered. This is done by evaluating equation (8.1) for each data frame. It is important to note that these measures will vary with age and health condition. It is observed that the skew and kurtosis measured

for the case of motion-corrupted data are much higher in magnitude when compared to the skew and kurtosis for clean data. Therefore, these measures could serve as candidate features for motion detection.

8.2.2 Frequency Domain Analysis

In the frequency domain, the kurtosis measure is computed for the magnitude of the Fourier spectrum for each data frame. This measure considers the magnitude of the components present at each frequency sampled by the Discrete Fourier Transform (DFT) operation. It is seen that kurtosis is lesser in magnitude for frames corrupted with motion artifact versus frames with clean data. This means that a Fourier spectrum of clean data has a lower number of significant frequency components (since only the harmonic components are prominent) compared to a spectrum of motion corrupted data (that consists of harmonic and random spectral components).

8.2.3 Bispectral Analysis and Quadratic Phase Coupling

The bispectrum and the bicoherence of each data frame are analyzed using the MATLAB Higher-Order Spectral Analysis Toolbox [45]. Significant peaks at non-zero frequencies are observed in the bispectrum diagonal slice plots for clean PPG data, thereby confirming the presence of strong quadratic phase coupling. In the case of clean PPG data from the initial subjects, Table 8.1 indicates peaks at tf_0 Hz, $t = 1, 2, 3$, where $f_0 = 1.54$ Hz is the most dominant frequency being coupled, indicating the presence of *self coupling* between frequencies (we have $f_0 + f_0 = 2f_0$ and $f_0 + 2f_0 = 3f_0$ and so on, indicating equally-spaced peaks in the diagonal slice plot). However, in the case of corrupt PPG data, QPC is observed to occur between random frequency components, and the phenomenon of self coupling is absent, as illustrated in Table 8.2. The features used for motion detection are summarized below:

1. *Time-Domain Features*: Skew and kurtosis measures that provide information

on the distribution of data. They contain information regarding amplitude variation of the PPG waveform.

2. *Frequency-Domain Feature*: Frequency domain kurtosis measure that indicates the presence of random components in the Fourier spectrum which are not present in the spectrum of a clean signal that contains only the main harmonics.

3. *Bispectral Feature and Quadratic Phase Coupling*: Clean PPG data are characterized by the presence of strong *self coupling* between the fundamental components of the frequency spectrum. This is absent in artifact-corrupted measurements where quadratic phase coupling between random frequency components is observed.

8.3 Motion Detection Unit (MDU)

8.3.1 Methods for Motion Artifact Detection

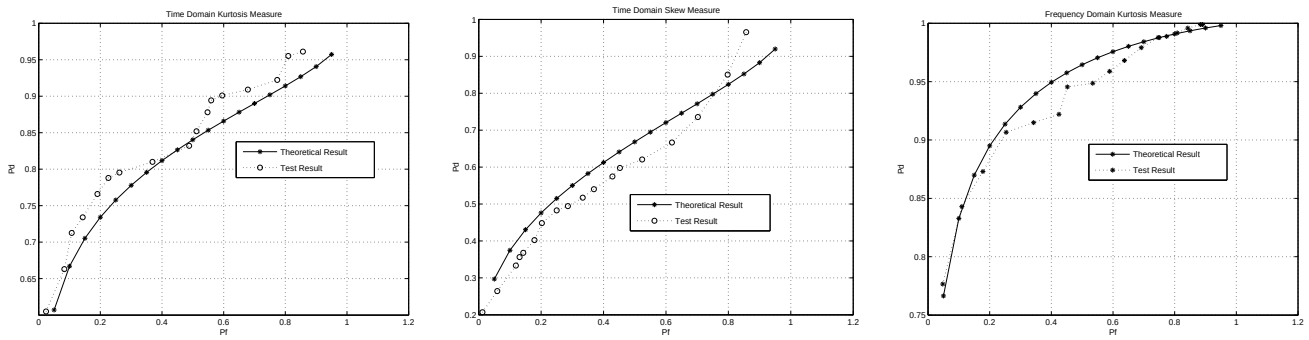
Neyman-Pearson (NP) Detection Rule Formulation: PPG data were collected from 10 healthy male and female subjects, in the age group of 22-30 years (different subjects from those considered for preliminary analyses and feature extractions in section 8.2),

Table 8.1: Bispectrum Plot Results - Clean Data

Coupling Frequency (f) Hz	Coupling Magnitude
1.54	0.1565
3.08	0.0211
4.62	0.0037

Table 8.2: Bispectrum Plot Results - Corrupt Data

Coupling Frequency (f) Hz	Coupling Magnitude
0.74	0.2839
2.74	0.0081
4.41	0.0009



(a) (b) (c)

Figure 8.1: Receiver Operating Characteristic (ROC) curves for the (a) time-domain kurtosis measure, (b) time-domain skew measure, and (c) frequency-domain kurtosis measure.

in order to formulate the hypotheses for the NP detection rule. The subjects followed the same motion routines as detailed in section 8.2. Based on the resulting data, distinguishing measures were computed for each data frame as described in section 8.2. For each of the measures, let H_0 denote the null hypothesis corresponding to the region for clean data and H_1 denote the alternative hypothesis corresponding to the region for corrupt data. Under the hypotheses H_0 and H_1 the time-domain kurtosis, skew and frequency-domain kurtosis measures are distributed as

$$\begin{aligned} H_0: &\quad y_i \sim \mathcal{N}(\mu_{0i}, \sigma_{0i}^2) \\ H_1: &\quad y_i \sim \mathcal{N}(\mu_{1i}, \sigma_{1i}^2) \qquad \forall i \in \{1,2,3\}, \end{aligned} \tag{8.4}$$

where, $\mathcal{N}(\mu, \sigma^2)$ is a Gaussian distribution with mean μ and variance σ^2, and i corresponds to each of the distinguishing metrics. y_i is the observation corresponding to the time-domain kurtosis ($i = 1$), skew ($i = 2$) and frequency-domain kurtosis ($i = 3$) measures. Based on values of the time-domain kurtosis and skew for each frame, local decisions $\delta_i \in \{-1, 1\}$ are made according to

$$\delta_i = \begin{cases} 1 & \text{if } y_i \geq \eta_i \\ -1 & \text{if } y_i < \eta_i \qquad \text{for } i \in \{1,2\} \end{cases} \tag{8.5}$$

where $\eta_i = \sigma_{0i}\mathcal{Q}^{-1}(1-P_{F_i})+\mu_{0i}$, $\delta_i = -1$ corresponds to the null hypothesis, and $\delta_i = 1$ corresponds to the alternative hypothesis. Here P_{F_i} is the false-alarm probability. For frequency-domain kurtosis, a decision is made according to

$$\delta_i = \begin{cases} 1 & \text{if } y_i \leq \eta_i \\ -1 & \text{if } y_i > \eta_i \qquad \text{for } i = 3 \end{cases} \tag{8.6}$$

where $\eta_i = \sigma_{0i}\mathcal{Q}^{-1}(P_{F_i}) + \mu_{0i}$. It can be easily shown for the time-domain kurtosis and skew measures that

$$\begin{aligned} P_{F_i} &= 1 - \mathcal{Q}\left(\frac{\eta_i - \mu_{0i}}{\sigma_{0i}}\right) \text{ and} \\ P_{D_i} &= 1 - \mathcal{Q}\left(\frac{\eta_i - \mu_{1i}}{\sigma_{1i}}\right), \end{aligned} \tag{8.7}$$

where P_{D_i} is the corresponding probability of detection for each measure, and $\mathcal{Q}^{-1}$ is the inverse $\mathcal{Q}$-function. For the frequency-domain kurtosis measure, the corresponding P_{F_i} and P_{D_i} are given by

$$\begin{aligned} P_{F_i} &= \mathcal{Q}\left(\frac{\eta_i - \mu_{0i}}{\sigma_{0i}}\right) \text{ and} \\ P_{D_i} &= \mathcal{Q}\left(\frac{\eta_i - \mu_{1i}}{\sigma_{1i}}\right). \end{aligned} \tag{8.8}$$

The tests in (8.5) and (8.6) are applied to data obtained from three healthy test subjects of 22-30 years of age (different subjects from those considered for formulating the hypotheses in (8.4)). The performance of the motion detectors on test data and theoretical receiver operating characteristic (ROC) curves for the tests are displayed in figure 8.1. It is important to note that the performance on test data conforms to that expected in theory, assuming a Gaussian distribution for y_i. The kurtosis measures in the time and frequency domains perform better than the skew measure in the time domain. This is because the skew measure indicates the symmetry (or the lack of it) of the distribution of the data about the mean and is thus more characteristic of the PPG waveform (or the subject), while the kurtosis measure captures random variations from the mean.

Self Coupling Detection Rule: As concluded earlier, clean PPG data are characterized by self coupling, which is absent from data containing motion artifact (though QPC between random frequency components is present). That is, self coupling implies that the data are clean or contain insignificant amount of motion artifacts. Hence, to determine the presence/absence of self coupling, the frequencies being coupled are noted for each data frame and a decision is made as follows:

$$\delta_i = \begin{cases} 1 & \text{Self coupling} \Rightarrow \text{clean data} \\ -1 & \text{No self coupling} \Rightarrow \text{corrupt data} \end{cases} \tag{8.9}$$

The P_D and P_F measures related to the self coupling measure are directly computed from the initial training set. The P_F value is found to be 0.0420, while the P_D value is found to be 0.8932 for this training set.

8.3.2 Decision Fusion

The time-domain measures (kurtosis, skew) and the frequency-domain measures (QPC, kurtosis) are modeled as four individual sensors whose independent decisions can be fused to detect the presence of motion artifact in a given data frame. To implement this sensor decision fusion, we employ two methods: hard-decision fusion presented in [32] and soft-decision fusion developed in [33].

The hard-decision fusion technique fuses individual sensor decisions while minimizing the probability of error for the overall detection system. Weights or reliability measures that are a function of individual P_{F_i} and P_{D_i} values are associated with the decisions made by the individual sensors, and the fused global decision is given as follows:

$$f(\delta_1, \ldots, \delta_n) \;=\; \begin{cases} +1 & \text{if } a_0 + \sum_{i=1}^{4} a_i \delta_i > 0 \\ -1 & \text{otherwise,} \end{cases} \quad . \tag{8.10}$$

where $\delta_i = +1$ and $\delta_i = -1$ $\forall i \in \{1, 2, 3, 4\}$ are the decisions made by the individual sensors corresponding to the presence/absence respective of motion artifact based on

the detection rules developed in section 8.3. The weights a_i are defined as

$$\begin{aligned} a_0 &= 0 \\ a_i &= \log\left(\frac{P_{D_i}}{P_{F_i}}\right) \text{ if } \delta_i = +1 \\ a_i &= \log\left(\frac{1-P_{F_i}}{1-P_{D_i}}\right) \text{ if } \delta_i = -1 \end{aligned} \tag{8.11}$$

assuming uniform cost assignment and equal prior probabilities for both hypotheses in (8.4).

The tests in (8.5), (8.6) and (8.9) are applied to data obtained from three test subjects as described in the previous section to obtain $\delta_i \ \forall i \in \{1,2,3,4\}$. We select thresholds in (8.5) and (8.6) to yield $P_{F_i} = 0.2$. We then evaluate the individual P_{D_i} and their respective probability of error, P_{error_i}. The weights are computed as in (8.11) and the fused decision is formed using (8.10). This is repeated for $P_{F_i} = 0.4$.

Under the same assumption of uniform costs and equal prior probabilities, we employ the soft-fusion technique as in [33]. In this technique, we partition each of the hypothesis regions H_0 and H_1 into mutually exclusive sub-regions and associate a level of confidence with each of them. The level of confidence depends on the distance of the local decision statistic from the decision threshold and hence is also a function of the probability of false alarm and the probability of detection associated with the sub-region. Optimal partitioning of the local decision space is achieved by partitioning the probability of false alarm and the probability of detection based on a J-divergence maximization criterion. For the time-domain kurtosis and skew measures, P_{F_i} and P_{D_i}have a functional relationship from (8.7) of

$$P_{D_i} = 1 - \mathcal{Q}\left(\frac{\sigma_{0i}\mathcal{Q}^{-1}(1-P_{F_i}) + \mu_{0i} - \mu_{1i}}{\sigma_{1i}}\right) \qquad \text{for i} \in \{1,2\} \tag{8.12}$$

For the frequency-domain kurtosis measure, the relationship between $P_{F_i}$and $P_{D_i}$from (8.8) is as follows-

$$P_{D_i} = \mathcal{Q}\left(\frac{\sigma_{0i}\mathcal{Q}^{-1}(P_{F_i}) + \mu_{0i} - \mu_{1i}}{\sigma_{1i}}\right) \qquad \text{for i} = 3 \tag{8.13}$$

The soft-decision-fusion technique is performed for $P_{F_i} = 0.2$ and $P_{F_i} = 0.4$ for $i \in \{1, 2, 3\}$. We refer the readers to [33] for a detailed qualitative treatment.

The results are summarized in Table 8.3. It can be easily seen that for both values of P_{F_i}, the fused decision, and in particular the soft fusion technique, provides a better probability of detection of motion artifact than the individual sensors.

Table 8.3: Sensor Decision Fusion Results

Sensor	P_F	P_D	P_{error}	P_F	P_D	P_{error}
Kurtosis	0.2	0.78	0.16	0.4	0.85	0.19
Skew	0.2	0.42	0.40	0.4	0.58	0.47
QPC	0.04	0.89	0.08	0.04	0.89	0.08
FDK	0.2	0.9	0.2	0.4	0.92	0.25
Fused Decision (hard)	0.07	0.91	0.07	0.2	0.94	0.11
Fused Decision (soft)	0.06	0.92	0.06	0.22	0.96	0.14

8.4 Summary

In this chapter, we formulated an HOS-based motion detection algorithm that is a consistent and reliable method to identify corrupt data frames that can be further processed for motion artifact removal. In the time domain, we observe that the skew and kurtosis measures associated with the motion-corrupted PPG data are much higher in magnitude than the same measures for clean PPG data. The frequency-domain kurtosis measure is much smaller for the corrupt data frames than for the clean frames. Bispectral analyses of PPG data indicate the presence of strong quadratic phase coupling (QPC) and, more specifically, self-coupling in the case of clean PPG data. Though quadratic phase coupling is found in data corrupted by motion artifacts, the self coupling feature for the desired PPG is absent. Based on all of these observations, Neyman-Pearson (NP) rules are formulated for each of the measures. It is understood that kurtosis-based detection is more reliable than the skew measure.

It is seen that soft decision fusion based on individual measures further enhances the overall detection capability.

Chapter 9

Stage Two - Motion Reduction

In this chapter, we present a motion artifact reduction *methodology* that is effective even in the case of severe subject movement. The enhanced preprocessing unit and the frequency domain ICA routine are detailed in section 9.1. The efficacy of the proposed methodolgy is demonstrated in section 9.3 by applying the proposed technique to real PPG data corrupted by significant motion artifacts.

9.1 Motion Artifact Reduction Method

The proposed motion artifact reduction method consists of a preprocessing unit and an FD-ICA unit. The preprocessing unit employs a Fourier series reconstruction, and its output is fed into the FD-ICA unit. The FD-ICA unit separates out the motion artifacts present at the frequency components chosen for reconstruction.

9.1.1 Preprocessing Unit

The preprocessing unit consists of the MDU, period estimator, and Fourier series reconstructor as in figure 9.1. The detrending and MDU are similar to those implemented in Sections 8.2 and 8.3 respectively.

Period Estimation: The period of the PPG signal can be estimated by an

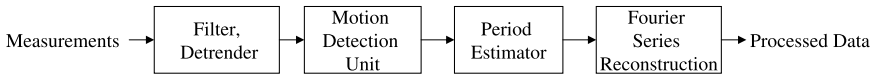


Figure 9.1: Preprocessing Unit.

autocorrelation operation. However, accurate and consistent prediction of the period is not possible by this method in the presence of extreme motion artifacts, where the PPG data are completely buried in noise. Hence, an indirect method for the computation of the period is considered (using the MDU from Section **??**). When a frame with motion artifact is identified as corrupt, the most recent frame with clean data is identified. The most significant frequency component from the clean frame is identified from its Fourier spectrum and also used as the fundamental period for the corrupt frame. This is a reasonable assumption, since moment-to-moment changes in heart rate should be minor.

Fourier Series Reconstruction: Upon obtaining the fundamental frequency and its harmonics for a corrupt frame, a Fourier series reconstruction of the signal is performed. In this reconstruction process, frequencies in the neighborhood of the harmonics are also used to account for the spectral variability of the PPG data. This is motivated by the inherent quasiperiodic nature of PPG signals. Additionally, in the presence of motion artifact, we expect a possible doppler spread around the fundamental frequency and its harmonics. Hence, we must account for that frequency-dependent spread of energy. The technique is illustrated in figure 9.2.

9.1.2 Frequency Domain ICA Unit

In this work, we assume that motion artifacts and PPG signal sources mix linearly with a mixing matrix in the time domain to form the observables (measurements). The observables, denoted by $x_1(t)$ and $x_2(t)$, are the preprocessed measurements due to the near-infrared and red excitation sources, respectively. For a measurement time

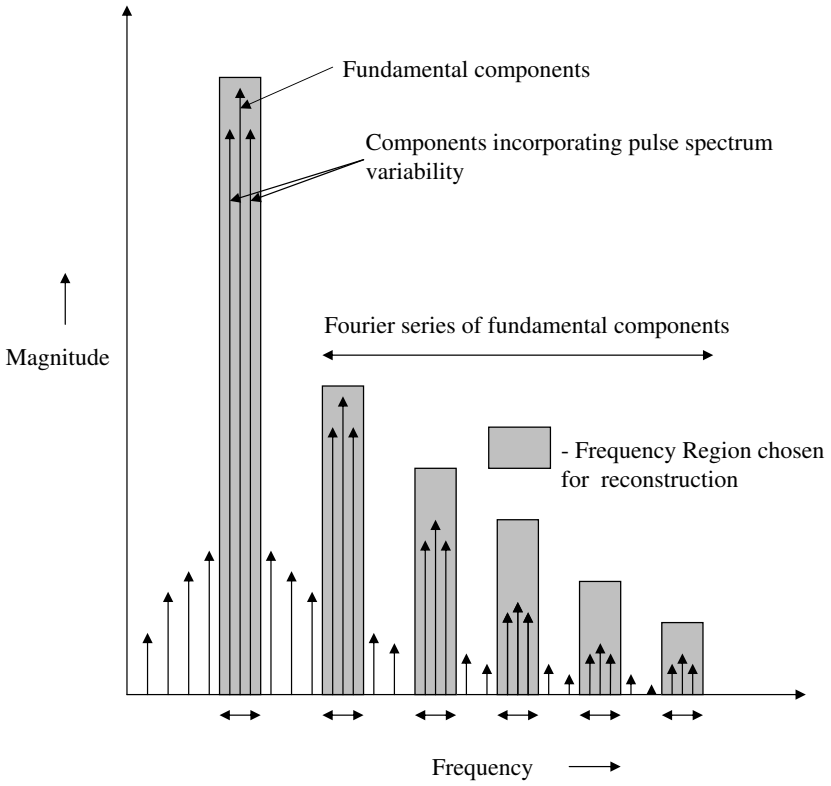


Figure 9.2: Fourier series reconstruction of a motion-corrupted frame.

frame, τ, this instantaneous mixing is given by

$$\begin{aligned} x_1(t) &= a_{11}s_1(t) + a_{12}s_2(t) \\ x_2(t) &= a_{21}s_1(t) + a_{22}s_2(t) \end{aligned} \tag{9.1}$$

where sources $s_1(t)$ and $s_2(t)$ denote the time-domain PPGs and motion artifacts, respectively. The mixing matrix $\mathbf{A}$ is assumed to be constant over the time frame τ. Hence, the frequency domain representation of (9.1) is

$$\begin{aligned} X_1(f) &= a_{11}S_1(f) + a_{12}S_2(f) \\ X_2(f) &= a_{21}S_1(f) + a_{22}S_2(f) \end{aligned} \tag{9.2}$$

where $X_i(f)$ and $S_i(f)$ are the Fourier transforms of $x_i(t)$ and $s_i(t)$ for $i = 1$ and 2. Now, considering the magnitude of $X_i(f)$ and using the triangle inequality, we can

write (9.2) as

$$\begin{aligned} |X_1(f)| &\leq a_{11}|S_1(f)| + a_{12}|S_2(f)| \\ |X_2(f)| &\leq a_{21}|S_1(f)| + a_{22}|S_2(f)| \end{aligned} \tag{9.3}$$

Both heart activity and motion artifacts affect the blood vessel volume at the tip of the finger, the acquisition point for PPG data in this study. Variations in blood volume due to source interference are observed to result in corrupt PPG data. However, interference between the two source signals is maximal when they are aligned in the same direction in the signal space. That is, for the case of maximal interference, $S_1(f)$ and $S_2(f)$ exhibit linear dependence, thus equality in (9.3) can be considered. Therefore,

$$\begin{aligned} |X_1(f)| &\approx a_{11}|S_1(f)| + a_{12}|S_2(f)| \\ |X_2(f)| &\approx a_{21}|S_1(f)| + a_{22}|S_2(f)| \end{aligned} \tag{9.4}$$

Hence, the Fourier magnitude spectrum of the corrupt PPG data can be modeled as motion artifacts and pulsatile blood volume components linearly mixing with an unknown mixing matrix. Since these pulsatile signals and motion artifacts are assumed to be statistically independent, we can employ ICA in either the time or frequency domain. From (9.4), ICA can be performed on the magnitude spectrum of $x_1(t)$ and $x_2(t)$ using the fastICA MATLAB package based on the fast ICA algorithm [51]. After applying the ICA routine to these magnitude spectra, we obtain an estimate of $\hat{S}_1(f)$ and $\hat{S}_2(f)$ that represents blood volume pulsation and motion artifact magnitude information, respectively. Utilizing the phase information of the original PPG data, we then reconstruct the clean PPG data and the motion artifacts.

In traditional frequency domain ICA approaches [50], ICA is performed on complex data under the assumption that the mixing matrix is different for each frequency bin. In another frequency domain approach applied to speech recognition [52],

Table 9.1: Correlation Coefficient (CC) for quantitative comparison of different techniques

	Wrist Movement	Finger (left/right)	Finger (up/down)	Elbow Movement	Random Movement
Time Domain ICA	0.6910	0.6681	0.6301	0.6557	0.6583
Magnitude-based FD-ICA	*0.7642*	*0.7238*	*0.6916*	*0.7119*	*0.7378*
Complex ICA	0.6141	0.5731	0.4297	0.5707	0.5924
Fourier Preprocessing	0.5220	0.5040	0.4116	0.4976	0.5193

the energy of the observables in the frequency domain is considered, but the unmixing matrix varies for different frequency bins. As a result, conventional approaches suffer from permutation problems and gain issues as discussed in [50]. However these issues do not exist in our approach since all of the frequencies selected by the preprocessing unit are treated as a single bin for which an unmixing matrix is computed; that is, the mixing/unmixing matrix is treated as frequency non-selective (constant for all frequencies). The gain issue is tackled by normalizing the determinant of the unmixing matrix to unity as discussed in [50], then the power of the recovered PPG source signal (obtained after the FD-ICA process) is scaled to the original PPG measurement data.

9.2 Methods

Data were collected from ten healthy subjects in the age group of 22-30 years who were subjected to the same motion routine as mentioned in Section 8.2. Data from each subject were fed to a MATLAB script that dissects the entire segment into

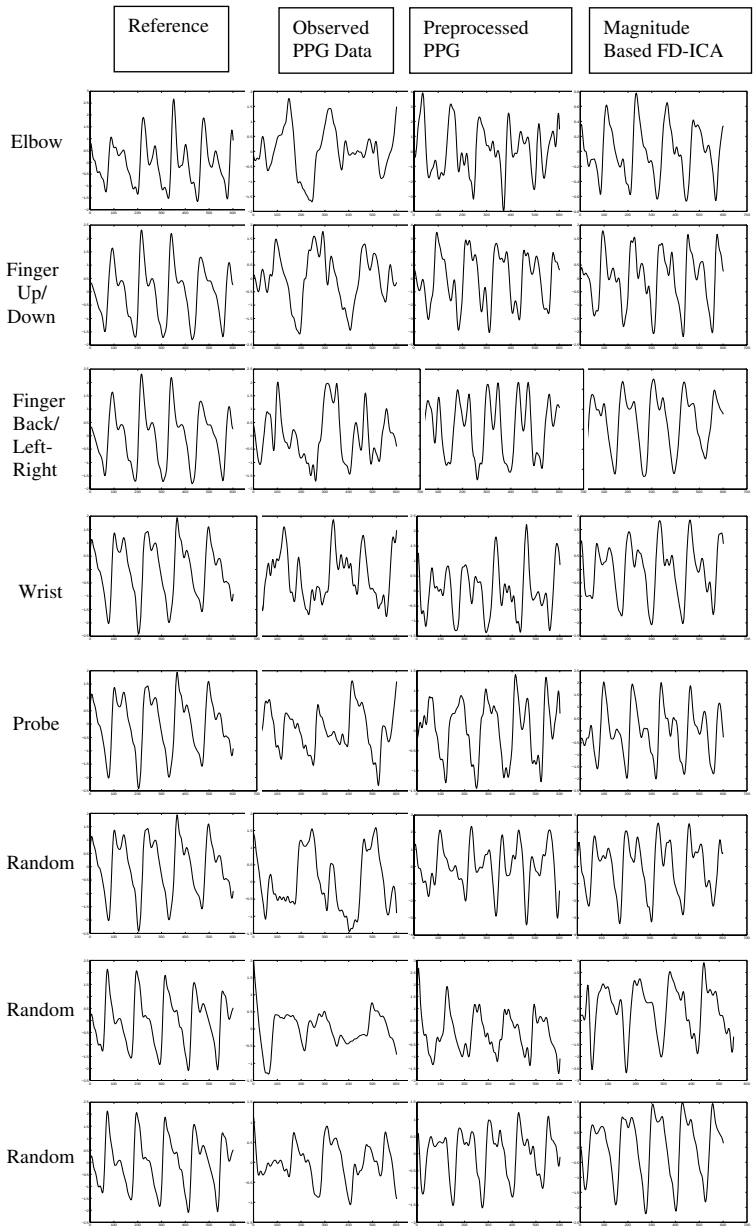


Figure 9.3: Separation results using the new technique

short equal-length frames. These frames were fed to the MDU to detect the presence of motion artifact. The frames identified as corrupt were then processed using the technique described in Section 9.1.

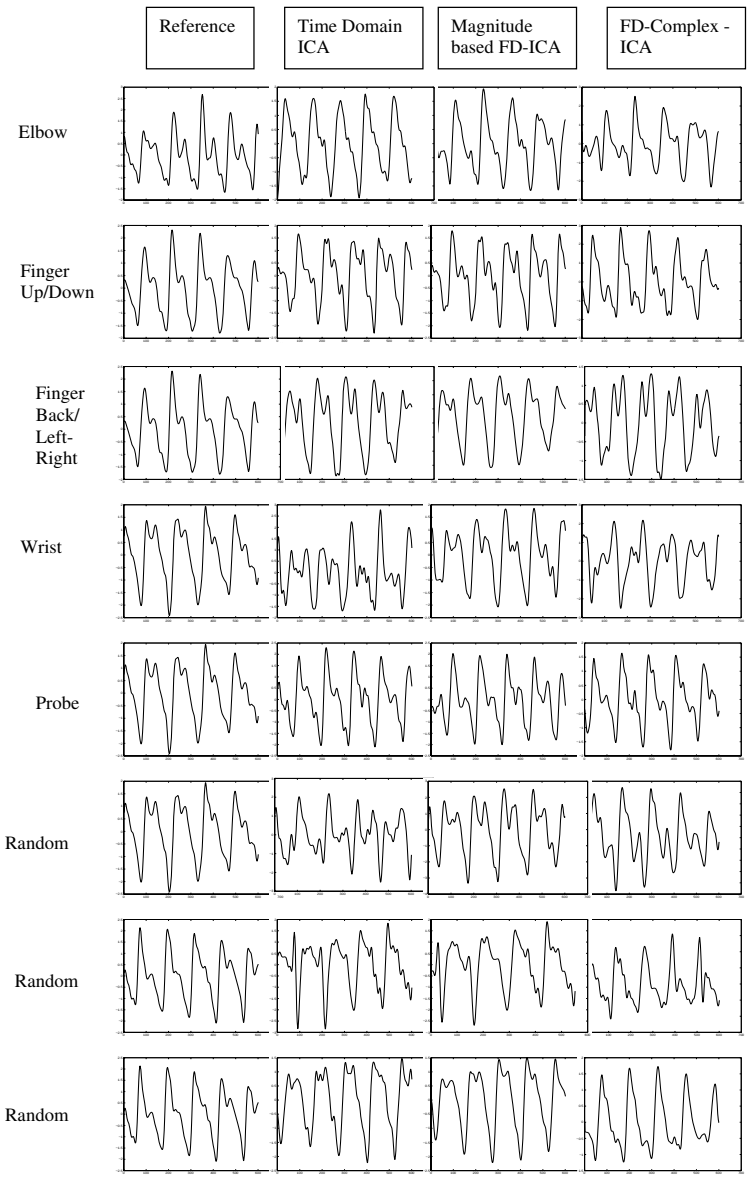


Figure 9.4: Comparison of the FD-ICA techniques with the time domain ICA and complex ICA approaches.

9.3 Results and Discussion

The results of the separation process for a single subject are presented in figure 9.3. Each recovered signal segment is visually compared (shape and peak-to-peak amplitude) with the most recent clean frame (the Reference). The proposed technique

(magnitude-based FD-ICA) is effective even in the presence of significant motion artifacts.

A quantitative comparison between various techniques as applied to different types of motion artifacts is presented in Table 9.1. A correlation coefficient (normalized to unity) is obtained by identifying the maximum of all peaks that appear in the normalized cross-correlation plot between the chosen reference signals and the output signals recovered by each of the techniques.

The same separation routine, when applied to data obtained from the rest of the ten subjects, effectively recovers the clean PPG data from the corrupt frames in all cases. However, the efficacy of the whole routine depends primarily on the preprocessing phase, in particular the accuracy in determining the fundamental frequency of the corrupt frame. This can be seen from Table 9.1, where the correlation coefficients between the recovered signals and their corresponding references drop when recovery from the Fourier preprocessing phase is not effective enough. Hence it can be understood that the clean signal are recoverable if the fundamental frequency, determined from the most recent clean frame (reference), matches the actual fundamental frequency of the corrupt frame.

9.3.1 Comparison Between FD-ICA and Time Domain ICA Methods

Assuming the independence of source signals and their linear mixing with an unknown mixing matrix in the time domain as in (9.1), the preprocessed observables were sent to a time domain ICA routine. The results obtained from the time domain ICA routine were then visually compared with those from the FD-ICA technique and the reference signal (shape and peak-to-valley height) as in figure 9.4. The FD-ICA technique outperforms the time domain ICA process. This is also apparent from Table 9.1: the correlation coefficient associated with the FD-ICA routine is much higher than its time domain counterpart for all cases of movement. It may be noted that

when performing an ICA based on the assumption of a constant mixing/unmixing matrix, the duration of the data frame should be short enough for the assumption to hold. Using long data frames would imply that the mixing/unmixing matrix is not a constant and hence leads to inaccurate estimation of the sources.

9.3.2 Comparison Between FD-ICA and Complex ICA Methods

In the complex ICA formulation of the problem, it is assumed that the sources mix with the unknown mixing matrix in a convolutive manner in the time domain. This directly translates to the instantaneous mixing of the sources with the mixing matrix in the frequency domain as in [50]. Here, the assumption of independence between the sources in the frequency domain is considered. For fairness in comparison, all frequencies selected by the preprocessing unit are treated as a single frequency bin. The mixing matrix is assumed to be frequency non-selective, and hence only one mixing matrix is computed for all the frequencies selected during preprocessing, unlike the traditional practice adopted in the complex frequency domain ICA approach. The complex ICA routine described in [53] was implemented for the complex data obtained by the Fourier transform of each preprocessed signal. The results obtained from the complex frequency domain ICA routine were visually compared with those from the FD-ICA technique and the reference signal (shape and peak-to-valley height) in figure 9.4. It can be easily seen that the newly proposed FD-ICA routine outperforms the complex ICA routine. The superiority of the proposed technique is quantified in Table 9.1: the correlation coefficient for the FD-ICA method is consistently higher than that of the complex ICA approach for all cases of movement.

9.4 Summary

In this chapter, we presented a motion artifact reduction methodology that consists of an enhanced PPG preprocessing routine and magnitude-based frequency domain ICA routine. We readily observe that this processing routine effectively reduces motion artifacts in corrupt data frames even in the event of significant motion. The FD-ICA routine proposed in this paper is compared with time domain ICA and complex ICA routines and is shown to be more effective in recovering clean PPG data. The efficacy of the method depends heavily on the ability of the MDU to identify corrupt/clean data segments and estimate the period of the waveform. More accuracy in the fundamental period estimation of the corrupt frame helps the FD-ICA routine to more effectively separate motion artifacts from desired data.

Chapter 10

Conclusions and Future Work

In this chapter, we present a summary of key contributions from our work in distributed estimation in WSNs in section 10.1. In section 10.2, we discuss plausible directions for future work and extensions.

10.1 Summary of Key Contributions

We derived the variance and its lower bound for BLUE-1 for any modulation scheme in general, and specifically for BPSK and uncoded QAM schemes. We observed that the bound is an additive factor away from the estimator variance in the case of perfect sensor channels. From simulations, we deduced that the bound is tight when the channel variances of the participating sensors are comparable. The upper and lower bound for variance of BLUE-2 were observed to be an additive factor away from the BLUE variance evaluated for perfect sensor channels was found to depend on the sensors with the worst and best channels respectively. For BLUE-3, the upper bound was a multiplicative factor away from the BLUE variance in the case of ideal sensor channels. We observed that the new upper bound is tighter than the bound in [15] for lower values of measurement noise. Finally, for all the estimators considered, deviation of the bound was observed to be more pronounced when the channel noise

variances of the participating sensors vary from each other.

Using three optimization formulations or system design approaches, we investigated and developed distributed estimation techniques in WSNs using BLUE-1. We analyzed the tradeoff between estimation error at the fusion center, resource utilization of the sensors to achieve that accuracy and implementation complexity to achieve that accuracy. We saw that the original formulation that is intricately non-convex can be transformed to the more well studied D.C. form. Further, the convex approximation of all the three formulations yielded the same solutions as in the original problem and its D.C. version. This assures that all the original problems can be solved in its convex approximated form without compromising the optimality of the solutions attained.

We introduced the notion of determining optimal sensor actions that depend on their operating states or residual battery power. For each of the formulations, we showed that optimal sensor actions depend on the operating state, channel conditions and measurement noise. In effect, we saw that the amount of error regarding target state introduced in the system not only depended on the nodes' knowledge about the parameter, but also its operating state and channel condition. We also studied the collaborative behavior achieved between the active sensors for each of the formulations. We highlighted the drawback in all the approaches related to the handling of the integer relaxation of b_is, in that determining the optimal b_i involves a brute-force exhaustive search in the immediate integer neighborhood.

We observed that the approach in Formulation B is most economical in terms of resource consumed for a target BLUE variance. Though Formulation A rendered a high quality estimator, the amount of resources utilized to realize the same depends on the constraint set on total resource expenditure. Formulation C was apparently the least favorable in resources consumed and BLUE variance achieved but was found desirable in case a completely autonomous distributed estimation process is desired

for BPSK modulation scheme.

10.2 Future Directions

Plausible future directions and extensions of our work in distributed estimation in WSNs are as follows -

10.2.1 Extensions

The immediate extensions of our work in distributed estimation in WSN are as follows -

- All the formulations for Scheme 2 (QAM) involved optimization of only one variable $b_i, \forall i \in \{1, \ldots, N\}$. However from the resource model for Scheme 2, we have $J = \sum_{i=1}^{N} \Lambda_i p_i \frac{l_i}{b_i}; p_i = B_s K_i a_i \log\left(\frac{2}{p_b}\right)(2^{b_i} - 1)$. It is desirable to solve the formulations for Scheme 2 considering b_i, p_i, l_i as problem variables.

- Finding optimal b_i first involved a continuous relaxation followed by an exhaustive search in the neighborhood. It is reasonable to find the conditions on the continuous variable b_i as in [19], which can be used to determine its optimal integer value.

10.2.2 Analysis based on Dynamical WSN Model

As a next step, it becomes imperative to investigate the tradeoff between WSN system performance and resource efficiency (or estimation error control algorithms) based on analytical models that capture the dynamics of each sensor node's operating state or health (characterized by residual battery), the wireless channel, and measurement noise. It may be noted that the seminal work of [54] inspired optimization techniques as a method to design fair and efficient resource allocation algorithms in general networks. More specifically, [54] considered the problem of dynamically allocating

available bandwidth to competing users with the help of rate control algorithms. It was proved that the complex dynamical system defined by the rate control algorithms and its stability can be achieved by appropriate optimization formulation (in its primal or dual form) that in turn provides the Lyapunov function for the dynamical system. It is of theoretical interest to examine conditions of stability, fairness, and optimality of estimation error control algorithms for such models and their mapping to the optimization framework.

10.2.3 Optimization Decomposition and Autonomy

Another interesting direction is to explore optimization decomposition for the formulations considered. We noted in 5 that applying decomposition to Formulation C for Scheme 1 (BPSK) renders system operation that is almost autonomous (with minimum signalling between fusion center and the sensor nodes). Achieving autonomous operation and performing only local optimization at each sensor node is desirable given that WSNs are generally resource constrained and delay-sensitive. In this context, it is required to analyze the conditions for optimality, and stability of local optimization given that each sensor has partial, or incomplete knowledge of the parameters (Lagrange Multipliers and the other coupling variables) that need to be signalled from the fusion center.

Appendix A - Throughput in Cooperative Wireless Relay Network

In this chapter, we analyze a wireless ad hoc/sensor network where nodes are connected via random channels and information is transported in the network in a cooperative multihop fashion using amplify and forward relay strategy. We characterize the network by studying important parameters such as: (1) SNR degradation with hop, (2) outage probability, (3) maximum permissible number of hops, and (4) maximum permissible number of simultaneous transmissions. We then devise a method for node selection and transmission of information across the network over disjoint routes between source and destination nodes by employing standard constructs from graph theory. Based on the above results, we investigate the throughput achievable in the network and its asymptotic scaling as function of the number of nodes in the network for a given channel distribution.

A-1 Introduction

Capacity and throughput analysis of wireless ad hoc networks have received considerable attention in recent times, primarily triggered by the pioneering work in [56]. The work in [56] essentially spurred research in the direction of determining theoretical

bounds on the capacity of wireless networks as a function of the number of nodes for umpteen variants of the wireless network model. Protocol and physical model [56, 57], different traffic (random, symmetric, asymmetric patterns) models [58, 59], dense and extended network models [61, 60] are a few of the many instances that have been looked upon in the past.

In parallel, it has been understood that the use of multiple-input-multiple-output (MIMO) systems for wireless communication enhances spectral efficiency and link dependability significantly by virtue of spatial diversity and space-time coding. However, the direct application of MIMO systems to wireless nodes that operate in ad hoc environments has been an unattractive choice due to their inherent size and processing requirements. Instead, spatially distributed configuration of nodes has been shown to mimic a MIMO system [67]. This has been demonstrated to be possible by the formation of virtual antenna arrays through distributed transmission and signal processing, thus resulting in a form of spatial diversity, formally referred to as cooperative diversity in [62]. In [62], the performance of classical relay systems as distributed antenna arrays, using different relaying strategies and protocols, is investigated. Capacity analysis for wireless networks employing cooperative diversity (more specifically amplify-and-forward) have been analyzed exhaustively from the perspective of [56] in [63, 64].

In this work, we consider a cooperative amplify-and-forward multihop network where the constituting nodes are connected to each other by channel strengths that are identical and independent random variables drawn from an arbitrary probability density. We base our study on the supposition that the strength of a signal received at any node in a small network is governed by random fluctuations that are not captured in the deterministic geometric models. We characterize such a random network by analyzing and evaluating important parameters like: (1) SNR degradation with hop, (2) outage probability, (3) maximum permissible number of hops and (4)

source-destination node pairs that communicate with each other simultaneously. We formulate the basic operation of the network by demonstrating a scheme for choosing appropriate nodes for relaying information over disjoint routes between all source-destination node pairs. We investigate the condition of existence of such disjoint paths between all source and destination nodes in the network and their characteristics. Based on all of the above, we evaluate the achievable throughput of the network and its asymptotic scaling for channel strengths drawn from an exponential density. To the best of our knowledge, the premise adopted in this paper follows the model in [58], where a traditional multihop scheme using decode and forward relaying in a random network is considered. However, our work considers a cooperative amplify-and-forward multihop mode of communication as will be detailed in the next section. Capacity analysis for such multi-level amplify-and-forward system has been analyzed in [66] for a fixed set of relay levels. Also in [65], the optimality of amplify-and-forward relaying strategy and the interplay between network rate, diversity and size under high SNR condition for a multi-level AF network is presented.

A-2 System Model

Consider a wireless ad hoc network consisting of n nodes, each equipped with single transmit-receive antenna communicating with each other in a wideband regime employing DS-CDMA. Let x_i denote the encoded message that a source node i wishes to transmit to a specific destination node j, where $x_i \sim \mathcal{CN}(0,1)\ \forall\ i,j \in \{1,\ldots,n\}, i \neq j$ and all x_i's are independent and identically distributed (i.i.d). All nodes in the network are constrained to transmit their message to any other node with a maximum power of P watts. Also, all nodes operate in a half-duplex mode where they can only either transmit or receive message at any instant of time.

Channel and Noise Model: Every pair of nodes $\{i,j\}, \forall i,j \in \{1,\ldots,n\}, i \neq$

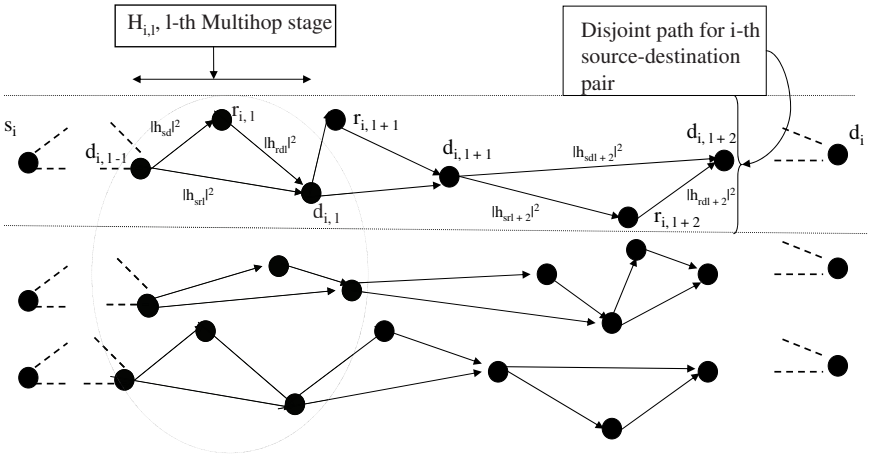


Figure A-1: System Model

j, is connected by a symmetric frequency fading channel, denoted by $h_{i,j} = h_{j,i} = h_{ij}$, which is an i.i.d random variable. Let the channel strength $|h_{ij}|^2$ be drawn from an arbitrary probability density $f(h)$ with zero mean and variance μ. Let $n_i \sim \mathcal{CN}(0, \sigma_{wn}^2)$ denote the temporally and spatially white noise at the terminal of node i. If k source nodes, randomly chosen from $\{1, \ldots, n\}$ and denoted by set T simultaneously transmit k distinct message $x_i, i \in T$ at maximum power P, then the instantaneous signal received at the receiver of a randomly chosen destination node $j \in D = \{1, \ldots, n\} \setminus T$ is given by

$$y_j = \sum_{i=1, i \in T}^{k} \sqrt{P} h_{i,j} x_i + n_j, \quad j \in D \tag{A-1}$$

Here, all the signals received at j from undesirable source nodes constitute the instantaneous interference noise. Note that the interfering signals and n_j are statistically independent of the desired signal received at j.

A-2.1 Network Operation

Cooperative Multihop Communication: Consider k source nodes denoted by $s_i, i \in T$ that attempt to communicate k distinct messages $x_i, i \in T$ with k destination nodes

$d_i, i \in D$. We assume that for a source-destination pair $\{s_i, d_i\}$, s_i communicates with d_i through L cooperative multihop stages denoted by $H_{i,1}, \ldots, H_{i,l}, \ldots, H_{i,L}$ as shown in figure A-1. Then in the first multihop stage $H_{i,1}$, s_i communicates with an intermediate destination $d_{i,1}$ with the aid of a relay node $r_{i,1}$. In the second multihop stage, $d_{i,1}$(intermediate source) communicates with the next intermediate destination $d_{i,2}$ assisted by a relay node $r_{i,2}$ and so on, till the final destination d_i is reached after L multihop stages. Hereafter, in general we would be addressing all (intermediate and otherwise) source, relay and destination nodes as S, R, and D respectively, unless explicitly specified. We naturally impose that each cooperative multihop stage $H_{i,l}$, $l \in \{1, \ldots, L\}, L > 1$ occurs in discrete time slots $t \in \{1, \ldots, L\}$, implying that each multihop stage indexed by l is synonymous to being indexed by time slots t. Hence s_i conveys its message to the desired destination d_i in L time slots.

Protocol and Relay Strategy: In each multihop stage or time slot, S first transmits a message to R with power P. Then, both S and R simultaneously transmit to D with power P, thereby creating a virtual MISO [67]. The nodes involved in all the cooperative multihop stages for any source-destination (s_i, d_i), $\forall i \in T$ pair employ an amplify and forward relay strategy to carry message from S to D. A condition we impose on each node in the network is that it can assist only one source-destination pair communication in all the L multihop stages or time slots; i.e., we require k disjoint paths between all the k source-destination pairs for transporting information between them. For convenience, in general for any of the disjoint paths between source-destination pairs, we denote the channel strength between S and D, R an D, and S and R in the $l-$th stage as $|h_{sdl}|^2, |h_{rdl}|^2$ and $|h_{srl}|^2$, respectively. This is illustrated in figure A-1.

Successful Communication and Throughput: Suppose S wishes to communicate with D with the assistance of R; then the communication is regarded as successful

if and only if the SNR received at D is equal to or greater than a threshold SNR value ρ_0. That is, communication between any source-destination pair is successful if and only if all the constituting cooperative multihop sequences produce an SNR of at least ρ_0 at all intermediate destination nodes and the final destination node. The message is dropped in the event of unsuccessful communication in any of the multihop stages and let ϵ denote the fraction of messages dropped.

The total aggregate throughput considering k source-destination pairs is defined as [58]:

$$C = (1-\epsilon)\frac{k}{L}\log(1+\rho_0) \tag{A-2}$$

Thus, we seek to evaluate ϵ, maximum permissible value of k, L, and ρ_0 to obtain an achievability result for the throughput of the system.

A-3 Main Result

Theorem 1: Consider a network consisting of n nodes, where any pair of nodes $\{i, j\}$ is connected by a channel strength $|h_{i,j}|^2$ that is drawn i.i.d. from an arbitrary probability density $f(h)$. Let there be k source-destination pairs communicating simultaneously with each other over k disjoint paths via L cooperative multi-hops using amplify and forward relay strategy. For any source-destination pair (s_i, d_i), $\forall i \in T$, let any S choose R and D in the l-th stage, $\forall l \in \{1, \ldots, L\}$, based on the rule:

$$\begin{aligned}&\left(|h_{sdl}|^2 + \frac{P|h_{rdl}h_{srl}|^2}{P|h_{srl}|^2 + N_l}\right)\frac{1}{P|h_{rdl}|^2 + P|h_{sdl}|^2 + N_l} \geq h_s \\ &\bigcap\left(1 + \frac{P|h_{rdl}|^2}{P|h_{srl}|^2 + N_l}\right)\frac{1}{P|h_{rdl}|^2 + P|h_{sdl}|^2 + N_l} \leq h_n,\end{aligned} \tag{A-3}$$

where, N_l denotes the sum of variance due to noisy channel and total interference noise received at D in the $l-$th stage. If $p_{i,l}$ be the probability that the condition in

(A-3) is satisfied, then for $p_n = p_{i,l}p_{i,l-1} \geq \frac{\ln n+\omega_n}{n}, \forall i \in T, \forall l \in \{1,\ldots,L\}$, $\omega_n \to \infty$ as $n \to \infty$, $\exists \alpha > 0$, such that a throughput of

$$C = (1-\epsilon) \times \frac{k}{L} \times \log\left(1 + \frac{h_s}{h_n\left(2\mu\rho^2(k-1) + \frac{\sigma_{wn}^2}{P}\right)}\right) \tag{A-4}$$

is achievable where

$$\epsilon = \sum_{l=1}^{L} Pr\left\{k' < \left(\frac{(Ph_s)^l(1-Ph_s)}{P\mu\rho^2 h_n \mu\rho_0(1-(Ph_s)^l)} - \frac{\sigma_{wn}^2}{P\mu\rho^2}\right)\right\} \to 0 \tag{A-5}$$

$$k-1 \leq k' \leq 2(k-1); k \leq \frac{\alpha n \ln np_n}{\ln n} \tag{A-6}$$

$$L \leq \min\left\{\frac{\log\left(\frac{1-(Ph_s)}{Nh_n\rho_0}+1\right)}{\log\left(\frac{1}{Ph_s}\right)}, \frac{\ln n}{\alpha \ln np_n}\right\} \tag{A-7}$$

and, $\rho \in \{0,1\}$ is the orthogonality factor. It may be noted here that the maximum attainable throughput $\sup\{C\}$ for the network considered is attained upon determining optimal values of p_n, h_s, and h_n (that in turn decides k and L).

Proof Summary - In the ensuing section, we first derive the condition in (A-3). Then using graph transformation techniques, the maximum number of simultaneous transmissions permissible (A-6) is derived. Further, we derive the maximum value of the minimum number of hops, and the outage probability (A-5) based on which the maximum allowable value of ρ_0 is obtained. Putting all these results together, we have the main result presented in (A-4).

A-4 Analysis

SNR Analysis: Based on the relaying protocol and strategy discussed in the previous section, we seek to investigate as to how the received signal and hence the SNR at D evolves with each cooperative multihop stage or time slot. Consider a source-destination pair (s_i, d_i). For analysis, we assume that, based on some node selection rule and routing strategy, the nodes that participate in all the cooperative multihop

stages or time slots $H_{i,1}, \ldots, H_{i,l}, \ldots, H_{i,L}$ have been identified. In the first cooperative multihop stage, s_i communicates its encoded message x_i with relay node $r_{i,1}$. If h_{sr1} denotes the channel connection between s_i and $r_{i,1}$, n_{r1} being the noise at $r_{i,1}$, and $h_{srj1}, j \in T \backslash \{i\}$ represents the channel between the interfering sources and $r_{i,1}$, then the signal received at the relay is -

$$y_i^{(r1)} = \sqrt{P} h_{sr1} x_i + n_{r1} + \sqrt{\rho^2} \sum_{j=1, j \neq i}^{k} \sqrt{P} h_{srj1} x_j, \tag{A-8}$$

The signal in (A-8) is normalized to unit average energy as -

$$y_{i,nor}^{(r1)} = \frac{y_i^{(r1)}}{\sqrt{P|h_{sr1}|^2 + \sigma_{wn}^2 + \rho^2 \sum_{j=1, j \neq i}^{k} P|h_{srj1}|^2}} \tag{A-9}$$

This is followed by a simultaneous transmission by s_i and $r_{i,1}$ to the intermediate destination $d_{i,1}$ (all with power P). If h_{sd1} and h_{rd1} denote the channel connections between s_i and $d_{i,1}$, and $r_{i,1}$ and $d_{i,1}$ respectively, n_{d1} be the noise at $d_{i,1}$, and $h_{sd1}^{(j)}$ represents the channel connection between all other sources and relays transmitting in the first stage (j indexes all other interfering sources and relays) then the signal received at $d_{i,1}$ is -

$$\begin{aligned} y_i^{(d1)} = \sqrt{P} h_{sd1} x_i + \sqrt{P} h_{rd1} y_{i,nor}^{(r1)} + \\ n_{d1} + \sqrt{\rho^2} \sum_{j=1}^{k, j \neq i} \sqrt{P} h_{sd1}^{(j)} x_j \end{aligned} \tag{A-10}$$

Assuming that $d_{i,1}$ performs an equal gain combining (EGC) of the signals received from s_i and $r_{i,1}$, the amount of desired signal power $S_{i,1}$ and unwanted noise power $N_{i,1}$ at $d_{i,1}$ are as follows -

$$\begin{aligned} S_{i,1} &= P|h_{sd1}|^2 + \frac{P^2 |h_{rd1} h_{sr1}|^2}{P|h_{sr1}|^2 + \sigma_{wn}^2 + \rho^2 \sum_{j=1, j \neq i}^{k} P|h_{jr1}|^2} \\ N_{i,1} &= \frac{P|h_{rd1}|^2 (\sigma_{wn}^2 + \sum_{j=1, j \neq i}^{k} P|h_{jr1}|^2)}{P|h_{sr1}|^2 + \sigma_{wn}^2 + \rho^2 \sum_{j=1, j \neq i}^{k} P|h_{jr1}|^2} + \\ & \quad \sigma_{wn}^2 + \rho^2 \sum_{j=1}^{k', j \neq i} P|h_{jd1}|^2 \end{aligned} \tag{A-11}$$

Instead of accounting for the instantaneous interference power at each receiver, we consider the expected value of the instantaneous interference noise power at each receiver terminal conditioned on the total number of simultaneous transmissions that occur in the stage -

$$\begin{aligned} I_1 &= \mathbb{E}\left[\rho^2 \sum_{j=1}^{k',j\neq i} P|h_{jr1}|^2 \ |k'=K\right] \\ &= \mathbb{E}\left[\rho^2 \sum_{j=1}^{k',j\neq i} P|h_{jd1}|^2 \ |k'=K\right] \end{aligned} \tag{A-12}$$

We regard the number of simultaneous transmissions that occur in each stage as a random variable denoted by k' that varies between $K=(k-1)$ and $K=2(k-1)$. By using conditionally averaged interference noise, equation (A-13) simplifies to -

$$\begin{aligned} S'_{i,1} &= P|h_{sd1}|^2 + \frac{P^2|h_{rd1}h_{sr1}|^2}{P|h_{sr1}|^2+\sigma_{wn}^2+I_1} \\ &= S_0\left(|h_{sd1}|^2 + \frac{P|h_{rd1}h_{sr1}|^2}{P|h_{sr1}|^2+N_1}\right) \\ N'_{i,1} &= \sigma_{wn}^2 + I_1 + \frac{P|h_{rd1}|^2(\sigma_{wn}^2+I_1)}{P|h_{sr1}|^2+\sigma_{wn}^2+I_1} \\ &= N_1\left(1+\frac{P|h_{rd1}|^2}{P|h_{sr1}|^2+N_1}\right) \end{aligned} \tag{A-13}$$

Where $S_0 = P$ and $N_1 = I_1 + \sigma_{wn}^2$. Using EGC, the SNR at $d_{i,1}$ after the first cooperative multihop is -

$$SNR_{i,1} = \frac{S_0\left(|h_{sd1}|^2 + \frac{P|h_{rd1}h_{sr1}|^2}{P|h_{sr1}|^2+N_1}\right)}{N_1\left(1+\frac{P|h_{rd1}|^2}{P|h_{sr1}|^2+N_1}\right)} \tag{A-14}$$

Node $d_{i,1}$ normalizes the received signal to unit energy and then scales it by power P following which the second cooperative multihop stage $H_{i,2}$ happens as discussed in the previous section. The desired signal power and noise power at $d_{i,1}$ before transmission (after normalization and scaling to power level P) in the second time

slot are as follows -

$$\begin{aligned} S_{i,1} &= S'_{i,1} \times \frac{P}{P|h_{rd1}|^2 + P|h_{sd1}|^2 + N_1} \\ N_{i,1} &= N'_{i,1} \times \frac{P}{P|h_{rd1}|^2 + P|h_{sd1}|^2 + N_1} \end{aligned} \tag{A-15}$$

Similarly, the signal and noise power received at $d_{i,2}$ following stage $H_{i,2}$ (after normalization and scaling) is -

$$\begin{aligned} S_{i,2} &= S_{i,1}\left(|h_{sd2}|^2 + \frac{P|h_{rd2}h_{sr2}|^2}{P|h_{sr2}|^2 + N_2}\right) \\ &\quad \times \frac{P}{P|h_{rd2}|^2 + P|h_{sd2}|^2 + N_2} \\ N'_{i,2} &= N_2\left(1 + \frac{P|h_{rd2}|^2}{P|h_{sr2}|^2 + N_2}\right) \\ &\quad + N_{i,1}\left(|h_{sd2}|^2 + \frac{P|h_{rd2}h_{sr2}|^2}{P|h_{sr2}|^2 + N_2}\right) \\ N_{i,2} &= N'_{i,2} \times \frac{P}{P|h_{rd2}|^2 + P|h_{sd2}|^2 + N_2}, \end{aligned} \tag{A-16}$$

where, $N_2 = I_2 + \sigma^2_{wn}$. By proceeding in the same fashion, we readily observe that after l cooperative multihop stages, the signal and noise power (after normalizing and scaling) are as follows -

$$\begin{aligned} &S_{i,l} = S_0 \prod_{x=1}^{l}\left(|h_{sdx}|^2 + \frac{P|h_{rdx}h_{srx}|^2}{P|h_{srx}|^2 + N_x}\right) \\ &\times \frac{P}{P|h_{rdx}|^2 + P|h_{sdx}|^2 + N_x}, \\ &N'_{i,l} = N_l\left(1 + \frac{P|h_{rdl}|^2}{P|h_{srl}|^2 + N_l}\right) + \sum_{x=1}^{l-1}\prod_{x=1}^{l-1} \\ &\left(|h_{sdx}|^2 + \frac{P|h_{rdx}h_{srx}|^2}{P|h_{srx}|^2 + N_x}\right)\frac{PN_{i,x}}{P|h_{rdx}|^2 + P|h_{sdx}|^2 + N_x}, \\ &N_{i,l} = N'_{i,l} \times \frac{P}{P|h_{rdl}|^2 + P|h_{sdl}|^2 + N_l}, \end{aligned} \tag{A-17}$$

where, $S_0 = P$ and $N_l = I_l + \sigma^2_{wn}$. Hence the SNR at D after l hops is given as $SNR_{i,l} = \dfrac{S_{i,l}}{N_{i,l}}$.

Link Formation: For any source-destination pairs (s_i, d_i), all the constituting cooperative multihop sequences must produce a minimum target SNR ρ_0 at all

intermediate nodes and the final destination node. In order to meet this requirement, we require that any S selects R and D in the $l-$th stage $\forall\ l \in \{1, \ldots, L\}$ based on the following rule:

$$\left(|h_{sdl}|^2 + \frac{P|h_{rdl}h_{srl}|^2}{P|h_{srl}|^2 + N_l}\right) \frac{1}{P|h_{rdl}|^2 + P|h_{sdl}|^2 + N_l} \geq h_s$$
$$\bigcap \left(1 + \frac{P|h_{rdl}|^2}{P|h_{srl}|^2 + N_l}\right) \frac{1}{P|h_{rdl}|^2 + P|h_{sdl}|^2 + N_l} \leq h_n, \tag{A-18}$$

where, h_s and h_n (both with units $(watt)^{-1}$) are design parameters that determine the level of connectivity and quality of connections between the nodes in the network. Note that index i has been dropped from equation (A-18) for convenience. Intuitively h_s gives a measure of the minimum acceptable signal power that reaches D from S in any stage; whereas h_n gives a measure of the maximum permissible noise power added in each stage. Let $p_{i,l}$ denote the probability of link formation in the l-th stage, $\forall i \in T, \forall l \in \{1, \ldots, L\}$ and hence also the probability that (A-18) is satisfied. Here, a link implicitly refers to a favorable S-R-D connection (a triangle in the graphical sense). Since there are k source-destination pairs communicating with each other simultaneously in any stage l, $\forall l \in \{1, \ldots, L\}$, we consider the formation of k non-overlapping favorable S-R-D connections with probability $p_{i,l}, \forall i \in T$ in each stage.

Maximum Number of Source-Destination Pairs: We define our network of n nodes as a random graph $G(n, p'_n)$, where p'_n is the probability of link formation between any two nodes in the network. In the l-th stage of operation, $\forall l \in \{1, \ldots, L\}$, we reduce each of the S-R-D connections (triangles) participating in the $i-$th source-destination communication path to a super-node (represented by black squares) as in figure A-2. Effecting the same transformation for every favorable S-R-D connections $\forall i \in T$ in all stages $\forall l \in \{1, \ldots, L\}$, we obtain a reduced graph $G'(n', p_n)$, where, $n' = k.L$ and p_n is obtained as $p_n = p_{i,l}.p_{i,l-1}, \forall i \in T, \forall l \in \{1, \ldots, L\}$. In order to achieve the network operation detailed in the previous section, it is required that we

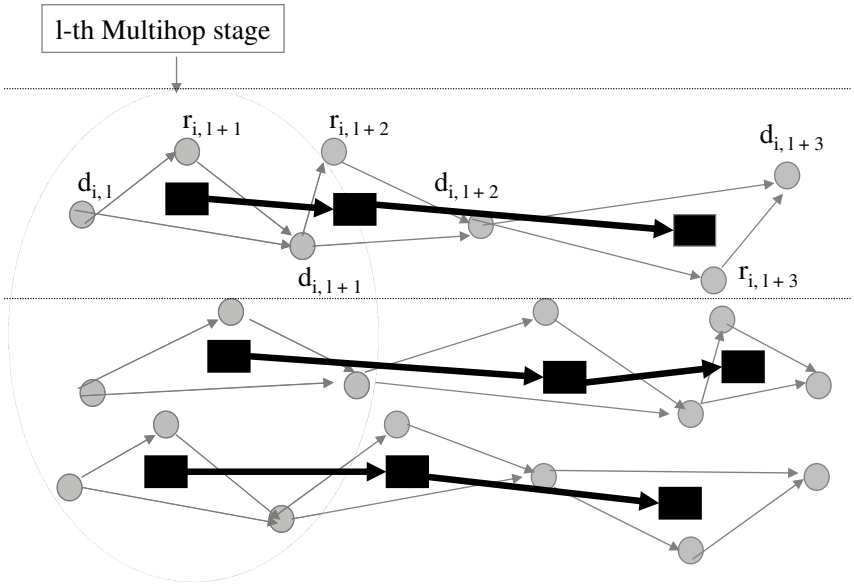


Figure A-2: Transformed Network Graph

establish vertex-disjoint paths between all source-destination nodes in the transformed graph. The conditions for existence of such paths for k disjoint pairs of vertices (s_i, d_i) is proved in [68]. Applying the result of the paper to the transformed graph $G'(n', p_n)$, we have -

Lemma 1: Suppose that $G'(n, p_n)$ and $p_n \geq \frac{\ln n + \omega_n}{n}$, where $\omega_n \to \infty$ as $n \to \infty$, then $\exists$ a constant $\alpha > 0$ such that, with high probability, there are vertex-disjoint paths connecting (s_i, d_i) for $\forall i \in T$, such that the cardinality of set T, $|T| = k \leq \frac{\alpha n \ln np_n}{\ln n}$. Also the length of almost all k vertex-disjoint paths are at most $\frac{\ln n}{\alpha \ln np_n}$.

Thus we have the result in (A-6).

Maximum value of the minimum number of hops: We seek to evaluate an upper bound for the number of hops permissible in the system under worst case noise scenario (N) that would ensure an SNR of at least ρ_0 at D for any stage l, $\forall\ l \in \{1, \ldots, L\}$. Here, the worst case noise scenario corresponds to each node in the system experiencing maximum interference noise due to $2(k-1)$ simultaneous

transmissions. First, we set -

$$\left(|h_{sdl}|^2 + \frac{P|h_{rdl}h_{srl}|^2}{P|h_{srl}|^2 + N_l}\right) \frac{1}{P|h_{rdl}|^2 + P|h_{sdl}|^2 + N_l} = h_s$$
$$\left(1 + \frac{P|h_{rdl}|^2}{P|h_{srl}|^2 + N_l}\right) \frac{1}{P|h_{rdl}|^2 + P|h_{sdl}|^2 + N_l} = h_n \quad \text{(A-19)}$$

$\forall l \in \{1, \ldots, L\}$. Then the signal power terms in (A-15), (A-16), (A-14) and (A-17) simplifies after algebraic manipulations as follows -

$$\begin{aligned} S_0 &= P \\ S_{i,1} &= S_0(Ph_s)^1 \\ S_{i,2} &= S_{i,1}Ph_s = S_0(Ph_s)^2 \\ &\vdots \\ S_{i,l} &= S_{i,0}(Ph_s)^l \end{aligned} \quad \text{(A-20)}$$

Similarly, the noise power terms simplify as -

$$\begin{aligned} N_{i,1} &= Nh_nP \\ N_{i,2} &= Nh_nP(1 + (Ph_s)) \\ &\vdots \\ N_{i,l} &= Nh_nP(1 + Ph_s + (Ph_s)^2 + \ldots + (Ph_s)^{l-1}) \\ &= Nh_nP\left(\frac{1 - (Ph_s)^l}{1 - (Ph_s)}\right) \end{aligned} \quad \text{(A-21)}$$

From the generalized signal and noise power expressions above, we can write the SNR after l cooperative multihop stages, $\forall l \in \{1, \ldots, L\}$, as -

$$\begin{aligned} SNR_{i,l} &= \frac{S_0(Ph_s)^l(1 - (Ph_s))}{Nh_nP(1 - (Ph_s)^l)} \\ &= \frac{(Ph_s)^l(1 - (Ph_s))}{Nh_n(1 - (Ph_s)^l)} \end{aligned} \quad \text{(A-22)}$$

Note that (A-22) corresponds to the minimum achievable SNR at D in any stage l, $\forall\, l \in \{1, \ldots, L\}$ for chosen values of h_s and h_n. In order to obtain an upper bound

on the minimum number of hops that guarantees an SNR of at least ρ_0, we have -

$$\frac{(Ph_s)^L(1-(Ph_s))}{Nh_n(1-(Ph_s)^L)} \geq \rho_0 \Rightarrow L \leq \frac{\log\left(\frac{1-(Ph_s)}{Nh_n\rho_0}+1\right)}{\log\left(\frac{1}{Ph_s}\right)} \tag{A-23}$$

The bound on L gives the maximum number of hops permissible in the system in the worst case noise scenario beyond which the SNR drops below ρ_0 and the communication is regarded unsuccessful. It can be seen that the maximum minimum number of hops directly depends upon h_s and inversely on h_n. This implies that with greater amount of noise added to the system, only smaller number of hops can ensure that the SNR at the destination is greater than ρ_0. Thus, we finally have $L \leq \min\left\{\frac{\log\left(\frac{1-(Ph_s)}{Nh_n\rho_0}+1\right)}{\log\left(\frac{1}{Ph_s}\right)}, \frac{\ln n}{\alpha \ln np_n}\right\}$ and hence the result in (A-7). Setting the value of L in this manner ensures that vertex-disjoint paths are established between k source-destination pairs and the threshold SNR condition is satisfied at all Ds.

Probability of Outage or Unsuccessful Communication: Communication between any source-destination pair (s_i, d_i) is regarded unsuccessful if the SNR at D in any of the L cooperative multihop sequences falls below ρ_0. If $Pr\{fail_{i,l}\} = Pr\{SNR_l < \rho_0\},\ \forall\ l \in \{1,\ldots,L\}$, then the maximum probability of erroneous communication between is given as -

$$\begin{aligned} Pr\{Fail_i\} &= \bigcup_{l=1}^{L} Pr\{fail_{i,l}\} \\ &\leq \sum_{l=1}^{L} Pr\{fail_{i,l}\} \\ &= \sum_{l=1}^{L} Pr\left\{\frac{(Ph_s)^l(1-(Ph_s))}{Nh_n(1-(Ph_s)^l)} < \rho_0\right\}, \end{aligned}$$

where, the probability of erroneous communication in each stage is considered to be

independent of each other. Replacing N with N_l, we get -

$$\begin{aligned}
&Pr\{Fail_i\} \geq \\
&\quad \sum_{l=1}^{L} Pr\left\{\frac{(Ph_s)^l(1-Ph_s)}{N_l h_n(1-(Ph_s)^l)} < \rho_0\right\} \\
&\quad = \sum_{l=1}^{L} Pr\left\{I_l < \frac{(Ph_s)^l(1-Ph_s)}{h_n\rho_0(1-(Ph_s)^l)} - \sigma_{wn}^2\right\} \\
&\quad = \sum_{l=1}^{L} Pr\left\{k' < \left(\frac{(Ph_s)^l(1-Ph_s)}{P\mu\rho^2 h_n\mu\rho_0(1-(Ph_s)^l)} - \frac{\sigma_{wn}^2}{P\mu\rho^2}\right)\right\}
\end{aligned}$$

Hence, for successful communication, we obtain an upper bound for the maximum number of simultaneous transmissions (that determines the noise floor due to interference) in any stage $l, \forall l \in \{1, \ldots, L\}$ that depends on the orthogonality factor μ, ρ_0, P, h_s and h_n and this proves the result in (A-5). By setting k' to its maximum value $2(k-1)$ and $l = 1$, we readily obtain a supremum of the permissible value of ρ_0 as

$$\rho_0 \leq \frac{h_s}{h_n\left(2\mu\rho^2(k-1) + \frac{\sigma_{wn}^2}{P}\right)} \tag{A-24}$$

Equation (A-24) shows the tradeoff between connectivity and the throughput of the network. Setting high value for h_s and low value for h_n clearly enhances network throughput, as communication happens over good links resulting in relatively higher SNR at the destination nodes. However, such a configuration reduces the probability of link formation thereby resulting in a sparsely connected network.

A-5 Simulations and Discussion

We draw $|h_{i,j}|^2$ from an exponential probability density with zero mean and variance $\mu = .5$. We set $P = 100, \sigma_{wn}^2 = 1$ and perform simulations for the evaluation of throughput of the system.We first fix $p_n = \frac{5\ln n}{n}$, and set $k = \frac{n\ln np_n}{\ln n}, L = \frac{\ln n}{\ln np_n}$. We

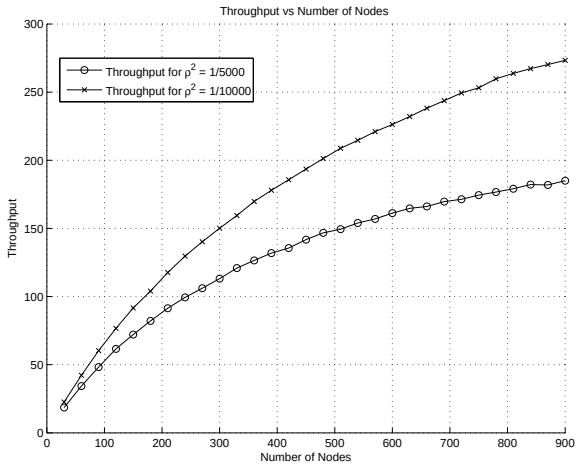


Figure A-3: Throughput of the network for different ρ^2 values.

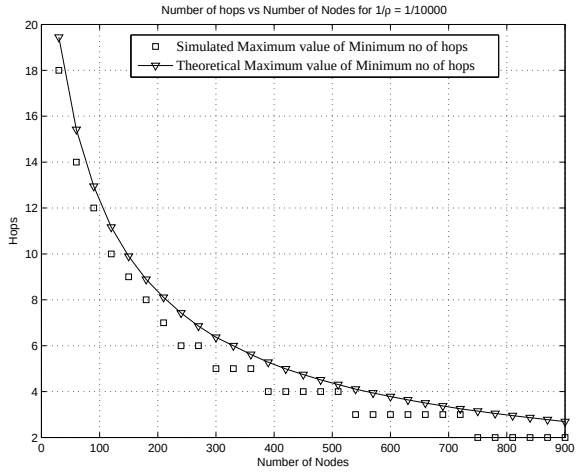


Figure A-4: Simulated and Theoretical values for maximum minimum number of hops in the network.

then find the optimal h_s and h_n values that maximize the throughput expression in (A-4). These values are used to perform monte carlo simulations to statistically evaluate the throughput of the system for $n = 30$ to 900 nodes and $\rho = 1/\sqrt{5000}, 1/\sqrt{10000}$. The results from simulation of the achievable throughput of the network are presented in figure A-3. We observe that the throughput of the system appears to scale asymptotically as $O(\log n)$. It may be noted that the throughput presented in the figure

cannot be claimed to be the maximum achievable value as the value of p_n has been set to its minimum permissible value, which may not be the optimal p_n that maximizes throughput. We verify the result pertaining to maximum minimum number of hops in equation (A-23) through simulations for $n = 30$ to 900 nodes in figure A-4. As predicted from the analytical result, we observe that the number of hops that help satisfy the SNR requirement at the destination decreases as noise increases in the system.

A-6 Summary

In this chapter, we characterized a cooperative amplify and forward multihop wireless network with random connections by studying parameters like SNR degradation, maximum permissible number of hops and source-destination pairs, outage probability, and finally the throughput achievable in the system. In order to realize the basic operation of the network, we demonstrated a scheme for node selection for transporting information over disjoint routes between any source-destination node pair in the system. Also presented are the existence condition and the characterization of disjoint routes between source and destination nodes in the network. Considering the channel strengths to be drawn from exponential density, the achievable throughput appears to scale asymptotically as $O(\log n)$. From simulations, we also see that the theoretical value of the maximum value of the minimum number of hops is exact.

References

[1] P. Varshney, *Distributed Detection And Data Fusion.* New York, NY: Springer-Verlag, Inc., 1997.

[2] S. M. Kay, *Fundamentals of Statistical Signal Processing.* Englewood Cliffs, NJ: Prentice-Hall, 1993.

[3] J. L. Speyer, "Computation and transmission requirements for a decentralized linear-quadratic-Gaussian control problem," *IEEE Trans. Autom. Control*, vol. AC-24, pp. 266 - 269, 1979.

[4] C. Y. Chong, "Hierarchical estimation," *Presented at the 2nd MIT/ONR Workshop on C3*, Monterey, CA, Jul. 1979.

[5] A. S.Willsky, M. Bello, D. A. Castanon, B. C. Levy, and G. Verghese, "Combining and updating of local estimates and regional maps along sets of one-dimensional tracks," *IEEE Trans. Autom. Control*, vol. AC-27, pp. 799 - 813, 1982.

[6] D. A. Castanon and D. Teneketzis, "Distributed estimation algorithms for nonlinear systems," *IEEE Trans. Autom. Control*, vol. AC-30, pp. 418 - 425, 1985.

[7] E. Ayanoglu, "On Optimal Quantization of Noisy Sources," *IEEE Transactions on Information Theory*, vol. 36, pp. 1450 - 1452, Nov. 1990.

[8] J. A. Gubner, "Distributed estimation and quantization," *IEEE Trans. Inf. Theory*, vol. 39, no. 4, pp. 1456 - 1459, Jul. 1993.

[9] W. M. Lam and A. R. Reibman, "Design of quantizers for decentralized systems with communication constraints," *IEEE Trans. Commun.*, vol. 41, pp. 1602 - 1605, Aug. 1993.

[10] V. Megalooikonomou and Y. Yesha, "Quantizer design for distributed estimation with communications constraints and unknown observation statistics," *IEEE Trans. Commun.*, vol. 48, no. 2, pp. 181 - 184, Feb. 2000.

[11] H. Papadopoulos, G. Wornell, and A. Oppenheim, "Sequential signal encoding from noisy measurements using quantizers with dynamic bias control," *IEEE Transactions on Information Theory*, vol. 47, pp. 978 - 1002, 2001.

[12] S. S. Pradhan, J. Kusuma, and K. Ramchandran, "Distributed compression in a dense microsensor network," *IEEE Signal Processing Magazine*, vol. 19, pp. 51 - 60, March 2002.

[13] A. Ribeiro and G. Giannakis, "Bandwidth-constrained distributed estimation for wireless sensor networks, part I: Gaussian case", *IEEE Trans. Signal Process.*, vol. 54, no. 3, pp. 1131 - 1143, Mar. 2006.

[14] Z.Q. Luo, "An Isotropic Universal Decentralized Estimation Scheme for a Bandwidth Constrained Ad Hoc Sensor Network", *IEEE J. Select. Areas Commun.*, vol. 23, pp. 735 - 744, Apr. 2005.

[15] J. J. Xiao, S. Cui, Z. Q. Luo, and A. J. Goldsmith, "Joint estimation in sensor networks under energy constraints", *IEEE 1st conference on Sensor and Ad Hoc Communications and Networks,* Santa Clara, CA, USA, Oct. 2004.

[16] V. Aravinthan, S. K. Jayaweera, and K. Al Tarazi, Distributed estimation in a

power constrained sensor networks, *Proc. VTC 2006- Spring,* vol. 3, pp. 1048 - 1052, 2006.

[17] Y. Mostofi, T.H. Chung, R.M. Murray, J.W. Burdick, "Communication and sensing trade-offs in decentralized mobile sensor networks: a cross-layer design approach", *Fourth International Symposium on Information Processing in Sensor Networks, IPSN 2005,* pp. 118 - 125, Apr. 2005.

[18] Y. Mostofi, T.H. Chung, R.M. Murray, "Effect of time-varying fading channels on control performance of a mobile sensor," *2004 First Annual IEEE Communications Society Conference on Sensor and Ad Hoc Communications and Networks*, pp. 317 - 324, Oct. 2004.

[19] J. Li and G. AlRegib, "Rate-constrained distributed estimation in Wireless Sensor Networks, *IEEE Transactions on Signal Processing,* vol. 55, no. 5, pp. 1634 - 1643, 2007.

[20] J. Li and G. AlRegib, "Energy-constrained distributed estimation in Wireless Sensor Networks, *Military Communications Conference, 2007 - MILCOM 2007, IEEE*, pp. 1 - 7, Oct. 2007 .

[21] J. Li and G. AlRegib, "Function-based Network Lifetime for Estimation in Wireless Sensor Networks, *IEEE Signal Processing Letters,* vol. 15, pp. 533 - 536, 2008.

[22] M. Chiang, S.H. Low, A.R. Calderbank and J.C. Doyle, "Layering as Optimization Decomposition: A Mathematical Theory of Network Architectures," *Proceedings of the IEEE,* pp. 255 - 312, Jan. 2007.

[23] X. Luo and G. B. Giannakis, "Energy-Constrained Optimal Quantization for Wireless Sensor Networks," *EURASIP Journal on Advances in Signal Processing* vol. 2008, Issue 2, no. 41, Jan. 2008.

[24] R. Horst and N. V. Thoai, "DC Programming: Overview," *Journal of Optimization Theory And Applications,* Vol. 103, No. 1, pp. 1 - 43, Oct. 1999.

[25] R. Krishnan and B. Natarajan, "Joint Power and Quantization Optimization for Target Tracking in Wireless Sensor Networks", *IEEE Globecom 2008,* New Orleans, Lousiana, Dec. 2008.

[26] R. Krishnan and B. Natarajan, "On the Error Bounds for Distributed BLUE in Wireless Sensor Networks", Under Peer Review - *IEEE Communication Letters 2009.*

[27] R. Krishnan and B. Natarajan, "Bounds on the Distributed BLUE Variance in Wireless Sensor Networks", Under Peer Review - *IEEE Signal Processing Letters 2009.*

[28] R. Krishnan and B. Natarajan, "Joint Power and Quantization Bits Optimization for Distributed Estimation in Wireless Sensor Networks", To be Submitted to *IEEE Transaction on Signal Processing Letters 2009.*

[29] R. Krishnan, B. Natarajan, and S. Warren, "Two-Stage Approach for Detection and Reduction of Motion Artifacts in Photoplethysmographic Data", To be Submitted to *IEEE Transactions on Biomedical Engineering.*

[30] R. Krishnan, B. Natarajan, and S. Warren, "Analysis and Detection of Motion Artifacts in Photoplethysmographic Data using Higher Order Statistics", *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP) 2008,* Las Vegas, Nevada, pp. 613 - 616, Apr. 2008.

[31] R. Krishnan, B. Natarajan, and S. Warren, "Motion Artifact Reduction in Photoplethysmography using Magnitude-based Frequency Domain Independent Component Analysis", *17TH International Conference on Computer Communications and Network 2008*, St. Thomas, Virgin Islands, pp. 1 – 5, Aug. 2008.

[32] Z. Chair and P. Varshney, "Optimal data fusion in multiple sensor detection systems", *IEEE Transactions on Aerospace Electronic Systems, AES,* vol. 1, Issue 1, pp. 98 - 101, Jan 1986.

[33] C. C. Lee and J. J. Chao, "Optimum local decision space partitioning for distributed detection, *IEEE Transaction on Aerospace Electronic Systems, AES,* vol. 25, Issue 7, pp. 536 – 544, Jul. 1989.

[34] S. Lee, B. L. Ibey, W. Xu, M. A. Wilson, M. N. Ericson, and G. L. Cote, "Processing of pulse oximeter data using discrete wavelet analysis", *IEEE Transactions on Biomedical Engineering*, vol. 52, Issue 7, pp. 1350 – 1352, Jul. 2005.

[35] J. M. Graybeal, and M. T. Petterson. "Adaptive filtering and alternative calculations revolutionizes pulse oximetry sensitivity and specificity during motion and low perfusion", *26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2004*, San Fransisco, vol. 2, pp. 5363- 5366, Sept. 2004.

[36] J. Y. A. Foo, "Comparison of wavelet transformation and adaptive filtering in restoring artifact-induced time-related measurement", *Biomedical Signal Processing and Control*, vol. 1, Issue 1, pp. 93 – 98, 2006.

[37] J. E. Scharf, S. Athan, and D. Cain. "Pulse oximetry through spectral analysis", *Proceedings of the* 12^{th} *Southern Biomedical Engineering Conference, 1993*, New Orleans, LA, pp. 227 - 229, Apr. 1993.

[38] B. S. Kim, and S. K. Yoo, "Motion artifact reduction in photoplethysmography using independent component analysis", *IEEE Transactions on Biomedical Engineering*, vol. 53, Issue 3, pp. 566 - 568, Mar 2006.

[39] J. Weng, Z. Ye, and J. Weng. "An Improved Pre-processing Approach for Pho-

toplethysmographic Signal", *27th Annual International Conference of the Engineering in Medicine and Biology Society, 2005*, pp. 41 - 44, Sep. 2005.

[40] N. Mammone, and F.C Morabito. "Independent component analysis and high-order statistics for automatic artifact rejection", *Proceedings of the IEEE International Joint Conference on Neural Networks, IJCNN '05*, Montreal, Quebec, vol. 4, pp. 2447 - 2452, Aug. 2005.

[41] Jianchu Yao, and S. Warren. "A Short Study to Assess the Potential of Independent Component Analysis for Motion Artifact Separation in Wearable Pulse Oximeter Signals", *27th Annual International Conference of the Engineering in Medicine and Biology Society, 2005*, pp. 3585 - 3588, 2005.

[42] B. Pilgram, P. Castiglioni, G. Parati, and M. Di Rienzo. "Dynamic detection of rhythmic oscillations in heart-rate tracings: A state-space approach based on fourth-order cumulants", *Biol. Cyber.,* vol. 76, pp. 299 - 309, 1997.

[43] J. M. Mendel. "Tutorial on Higher-Order Statistics (Spectra) in Signal Processing and System Theory: Theoretical Results and Some Applications", *Proc. IEEE*, vol. 79, pp. 278 - 305, Jan. 1991.

[44] C. L. Nikias and A. P. Petropulu. *Higher-Order Spectra Analysis: A Nonlinear Signal Processing Framework,* Prentice Hall Inc., 1993.

[45] A Swami, J. M. Mendel and C. L. Nikias. *Higher-Order Spectral Analysis Toolbox For Use with MATLAB*, User's Guide Version 2.

[46] C. L. Nikias, and M.R. Raghuveer. "Bispectrum Estimation: A digital signal processing framework", *Proc. IEEE,* vol. 75, no.7, pp. 867 - 891, 1987.

[47] A. Mansour, and C. Jutten."What should we say about the kurtosis?", *Signal Processing Letters, IEEE*, vol. 6, Issue 12, pp. 321 – 322, Dec 1999

[48] R. F. Dwyer, "Use of the Kurtosis Statistic in the Frequency Domain as an Aid in Detecting Random Signals", *IEEE Journal of Oceanic Engineering*, vol. 9, no.2, pp. 85 - 92, Apr. 1984.

[49] J. Muthuswamy, D. L. Sherman, and N. V. Thakor. "Higher-Order Spectral Analysis of Burst Patterns in EEG", *IEEE Transactions On Biomedical Engineering,*, vol. 46, no.1, Jan. 1999.

[50] P. Smaragdis. "Blind Separation of Convolved Mixtures in the Frequency Domain", *International Workshop on Independence and Artificial Neural Networks* University of La Laguna, Tenerife, Spain, Feb. 1998.

[51] A. Hyvrinen. "Fast and robust fixed-point algorithms for independent component analysis", *IEEE Transactions on Neural Networks*, vol. 10, Issue 3, pp. 626 - 634, Mar. 1999.

[52] H. M. Park, Y. Jung, T.W. Lee, and S.Y. Lee. "On Subband-Based Blind Signal Separation for Noisy Speech Recognition", *Electronic Letters*, vol. 35, pp. 2011 - 2012, 1999.

[53] E. Bingham, and A. Hyvrinen. "A fast fixed-point algorithm for independent component analysis of complex-valued signals", *International Journal of Neural Systems*, pp. 1 - 8, 2000.

[54] F. Kelly, A. Maulloo, and D. Tan. "Rate Control in Communication Networks: Shadow Prices, Proporational Fairness and Stability", *Journal of Operational Research Society* no. 49, pp. 237 - 252, 1998.

[55] R. Krishnan, and B. Natarajan, "On the Achievability in Cooperative Amplify and Forward Wireless Relay Network", Submitted to *Globecom 2009.*

[56] P. Gupta and P. R. Kumar, "The capacity of wireless networks," *IEEE Transactions on Information Theory,* vol. 46, no. 2, pp. 388 - 404, Mar. 2002.

[57] M. Gastpar and M. Vetterli, "On the capacity of wireless networks: the relay case," IEEE INFOCOM 2002 - Twenty-First Annual Joint Conference of the IEEE Computer and Communications Societies Proceedings, vol. 3, pp. 1577 - 1586, 2002.

[58] R. Gowaikar, B. Hochwald and B. Hassibi, "Communication over a wireless network with random connections," *IEEE Transactions on Information Theory*, vol. 52, no. 7, pp. 2857 - 2871, Jul. 2006.

[59] M. Franceschetti, O. Dousse, D. Tse, and P. Tiran, "Closing the Gap in the Capacity of Wireless Networks Via Percolation Theory," *IEEE Transactions on Information Theory,* vol. 53, no. 3, pp. 1009 - 1018, Mar. 2007.

[60] L. Xie and P. R. Kumar, "A Network Information Theory for Wireless Communication: Scaling Laws and Optimal Operation", *IEEE Transactions on Information Theory*, vol. 50, no. 5, May 2004.

[61] B. Sirkeci-Mergen, A. Scaglione, "A Continuum Approach to Dense Wireless Networks with Cooperation", *Proceedings of the IEEE INFOCOM,* Miami, FL, Mar. 2005.

[62] J. N. Laneman, D. N. C. Tse and G. W. Wornell, "Cooperative Diversity in Wireless Networks: Efficient Protocols and Outage Behavior", *IEEE Transactions on Information Theory*, vol. 50, no. 12, Dec. 2004

[63] H. Blcskei, R. U. Nabar, . Oyman and A. J. Paulraj, "Capacity scaling laws in MIMO relay networks," *IEEE Transactions on Wireless Communication,* vol. 5, no. 6, pp. 1433 - 1444, Jun. 2006.

[64] V.I. Morgenshtern and H. Bolcskei, "Crystallization in Large Wireless Networks", *IEEE Transactions on Information Theory,* vol. 53, no. 10, pp. 3319 - 3349, Oct. 2007.

[65] S. Borade, Z. Lizhong and R. Gallager, "Amplify-and-Forward in Wireless Relay Networks: Rate, Diversity, and Network Size", *IEEE Transactions on Information Theory,* vol. 53, no. 10, pp. 3302 - 3318, Oct. 2007.

[66] S. Yeh, O. Leveque, "Asymptotic capacity of multi-level amplify-and-forward relay networks," *Proceedings of IEEE International Symposium on Information Theory,* Nice, France, Jun. 2007.

[67] R. U. Nabar, H. Blcskei, and F. W. Kneubhler, "Fading relay channels: Performance limits and space-time signal design," *IEEE Journal on Selected Areas of Communication,* vol. 22, no. 6, pp. 1099 - 1109, Aug. 2004.

[68] A. Z. Broder, A. M Frieze, S. Suen and E. Upfal, "An Efficient Algorithm for the Vertex-Disjoint Paths Problem in Random Graphs," *Proc 7th Symp. Discrete Algorithms,* Atlanta, pp. 261 - 268, 1996.