基於長期感測雜訊變異量消息之無線感測網路能量效率化分散式估計

學生：黃千致　　　　　　　　　指導教授：李大嵩 博士

國立交通大學電信工程學系碩士班

摘要

在無線感測網路中因頻寬與能量的限制，每個感測器只能傳送有限個位元數至融合中心，此融合中心接收資料並且利用最佳線性不偏估計(best-linear-unbiased-estimator)融合規則估計未知參數。本論文中有兩種最佳能量分配策略被考慮：最小能量分散式估計及最小均方誤差(minimal mean square error)分散式估計。第一種策略為最小化總能量並滿足某程度的效能限制，而此效能限制為對量測雜訊變異量機率分佈取平均後的均方誤差。第二種策略為最小化平均均方誤差並滿足某程度的總能量限制。我們也考慮最小能量分散式估計在路徑衰減的拉瑞衰減通道中，而這個感測器與融合中心之間的通道用二進位對稱式通道(binary symmetric channel)來表示。現今有關能量分配的相關研究皆需已知瞬間雜訊變異量，而本論文提出的演算法只需知道長期雜訊變異量的統計特性。而這個問題被表示成凸型最佳化問題(convex optimization)，並且得到封閉式最佳解，最後從此演算法中可以發現一些與其他相關研究共同的特性。根據模擬結果，此演算法確實能顯著的改進能量使用效率比起均衡式能量分配。

Energy-efficient Decentralized Estimation in Wireless Sensor Network Based on Long-Term Noise Variance Knowledge

Student: Qian-Zhi Huang Advisor: Dr. Ta-Sung Lee

Department of Communication Engineering

National Chiao Tung University

Abstract

In a wireless sensor network, due to bandwidth and energy limitation, each sensor is only able to transmit a finite number of bits to the fusion center (FC) which combines the received bits to estimate the unknown parameter by the best-linear-unbiased-estimator (BLUE) fusion rule. In this thesis, the optimal power allocation strategies are considered for two cases: minimal energy decentralized estimation and minimal mean square error decentralized estimation. In the first case, the minimization of total energy is subject to a certain performance constraint in terms of mean square error (MSE) averaged over the noise variance distribution. In the second case, the minimization of the average MSE is subject to a certain energy constraint. We also consider the minimal energy decentralized estimation over rayleigh fading channels with path loss. The wireless links between sensors and the FC are characterized by the binary symmetric channels (BSCs). While most of the existing related works require the knowledge of instantaneous noise variance for energy allocation, the proposed approach instead relies on an associated model. The problems can be reformulated in the form of convex optimization and the closed-form optimal solutions are obtained. The proposed schemes share several attractive features of the existing designs and are seen to significantly improve energy efficiency against the uniform allocation schemes by the simulation results.

Acknowledgement

I would like to express my deepest gratitude to my advisor, Dr. Ta-Sung Lee, for his enthusiastic guidance and great patience. I learn a lot from his positive attitude in many areas. Heartfelt thanks are also offered to all members in the Communication System Design and Signal Processing (CSDSP) Lab for their constant encouragement. Finally, I would like to show my sincere thanks to my parents for their invaluable love.

Contents

List of Figures

Acronym Glossary

AWGN	additive white Gaussian noise
BER	bit error rate
BSC	binary symmetric channel
BPSK	binary phase shift keying
BLUE	best linear unbiased estimator
CSI	channel state information
DES	decentralized estimation schemes
FC	fusion center
IEEE	institute of electrical and electronics engineers
ISI	intersymbol interference
KKT	Karush-Kuhn-Tucker
LB	lower bound
LOS	line of sight
MMSE	minimum mean square error
MSE	mean square error
PES	percentage of energy saving
PHY	physical layer
PDF	probability density function
QAM	quadrature amplitude modulation
QPSK	quaternary phase shift keying
RX	receiver
SNR	signal-to-noise ratio
TX	transmitter
WSN	wireless sensor network

Chapter 1

Introduction

Wireless sensor networks (WSNs) are ideal for environmental monitoring applications because of their low implementation cost, agility, and robustness to sensor failures. A popular WSN architecture consists of a fusion center (FC) and a large number of spatially distributed sensors. The FC can be either a standard base station or a mobile access point such as an unmanned aerial vehicle hovering over the sensor field. Each sensor in a WSN is responsible for local data collection as well as occasional transmission of a summary of its observations to the FC via a wireless link. In a practical WSN, each sensor has only limited computation and communication capabilities due to various design considerations such as small size battery, bandwidth, and cost.

As a result, it is difficult for sensors to send their entire real-valued observations to the FC. Instead, a more practical decentralized estimation scheme is to let each sensor quantize its real-valued local measurement to an appropriate length and send the resulting discrete message (typically short) to the FC, while the latter combines all the received messages to produce a final estimate of the unknown parameter. Naturally, the message lengths are dictated by the power and bandwidth limitations, sensor noise characteristics, wireless channel condition as well as the desired final estimation accuracy.

Recently, several decentralized estimation schemes (DES) [1, 2, 3] have been proposed for parameter estimation in the presence of additive sensor noise. These DESs require each sensor to send only a few bits to the fusion center, with the message length determined by the sensor's local SNR. Performance of the resulting estimator is shown to be within a constant factor of the best linear unbiased estimator (BLUE) performance.

In a practical WSN, the wireless links from sensors to the FC may have different qualities, depending on the sensor locations relative to the FC. Intuitively, local message length should depend not only on the quality of sensor's observation (i.e., local SNR), but also on the quality of its wireless link to the FC. In particular, even if a sensor has a high quality observation, it should not perform any local quantization or transmission when its wireless link to the FC is weak, in order to conserve sensor energy. In general, minimizing the total sensor energy consumption for a decentralized estimation task is essential to ensure long lifespan of a WSN. Motivated by these considerations, the authors of [4, 5] proposed optimal coded and uncoded transmission strategies for sensor networks which can minimize the required energy per transmitted bit, although no consideration was given to the quantization effect and the accuracy of final estimation.

As energy efficiency is a critical concern for sensor network design [6, 7, 8], the decentralized estimation is formulated as optimal bit-loading problem. In the practical system the probability density function (pdf) of the observation noise is hard to characterize, especially for a large scale sensor network. The signal processing algorithms that do not require knowledge of the sensor noise pdf have been proposed [7, 8]. While most of the existing related works require the knowledge of instantaneous noise variances for energy allocation, the proposed approach instead relies on an associated statistical model. In order to improve the estimation performance against the variation of sensing conditions, repeated update of the noise profile would be needed. This comes inevitably at the cost of more training overhead and extra energy consumption. If the sensing environment is harsh, the sensing noise will change quickly. The proposed signal processing algorithm which relies on an associated sensing noise variance model is needed

This thesis is organized as follows. In Chapter 2, we introduce the system model of wireless sensor networks and decentralized estimation scheme. In Chapter 3, minimal energy decentralized estimation based on long-term noise variance knowledge is proposed. In Chapter 4, minimal mean square error decentralized estimation based on long-term noise variance knowledge is proposed. In Chapter 5, we consider minimal energy decentralized estimation with the noisy channel between each sensor and the FC by exploiting long term noise variance information. The main results are presented and the numerical performance of the proposed schemes are illustrated. Finally, we conclude this thesis and propose some potential future works in Chapter 6.

Chapter 2

Wireless Sensor Network Overview

Recent technological advances in Wireless Sensor Networks have led to the emergence of small, inexpensive, and low-power sensor devices with limited on-board processing and communication capabilities. When suitably programmed and deployed in large scale, such networked sensors can cooperate to accomplish various high-level tasks. Sensor networks of this type are well-suited for situation awareness applications such as environmental monitoring (air, water, and soil), smart factory instrumentation, military surveillance, precision agriculture, intelligent transportation and space exploration.

WSNs deploy geographically distributed sensor nodes to collect information of interest. The collected information is then aggregated via wireless transmissions at a fusion center to generate the final intelligence. A typical wireless sensor network consists of a fusion center and a number of sensors. The sensors typically have limited energy resources and communication capability. Each sensor in the network makes an observation of the quantity of interest, generates a local signal, and then sends it to the fusion center where the received sensor signals are combined to produce a final estimate of the observed quantity.

Since sensors have only small-size batteries whose replacement can be costly, sensor network operations must be energy efficient in order to maximize network lifespan. A main objective of current sensor network research is to design energy-efficient devices and algorithms to support all aspects of network operations.

2.1 System Model of Wireless Sensor Networks

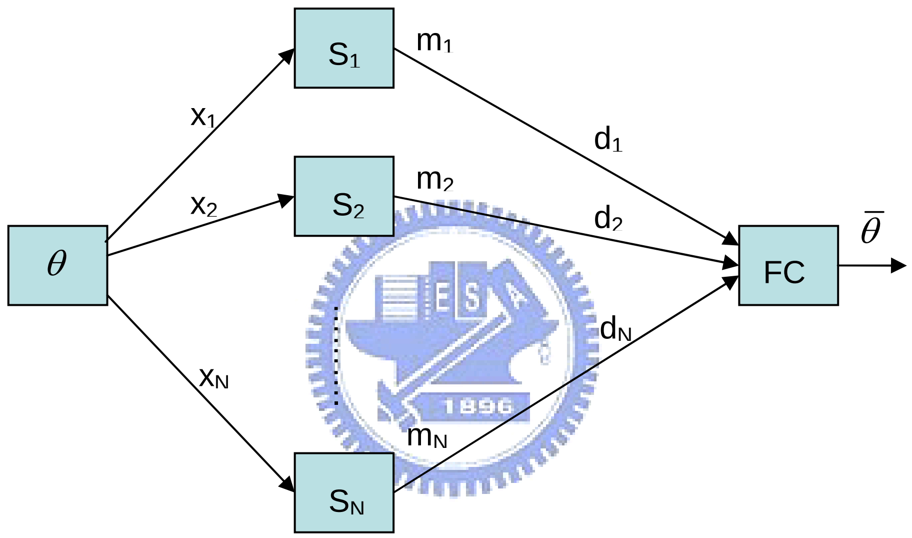

Figure 2.1 : System Model of Wireless Sensor Network

A common WSN architecture consists of a fusion center and a number of geographically distributed sensors. Such network architecture can be used to accomplish a joint signal processing task such as decentralized estimation and detection. In this chapter, we consider decentralized estimation of an unknown by a set of distributed sensor nodes and a fusion center. The sensors collect real-valued data, perform a local data compression and send the resulting discrete messages to the fusion center, while the latter combines the received messages to produce a final estimate of the observed signal.

The universal decentralized estimation schemes (DESs) let each sensor send to the fusion center a short discrete message whose length is determined by the local signal-to-noise ratio (SNR), while guaranteeing a mean squared estimation error (MSE) performance that is within a constant factor of that achieved by the centralized best linear unbiased estimator (BLUE). However, this chapter still assumes that the wireless channel between sensor and fusion center are ideal without any distortion.

2.1.1 Measurement and Quantization of Each Sensor

Consider a set of N distributed sensors, each making observations on deterministic source signal θ. The observations are corrupted by additive noise. The local observation at the *i*th node is

$$x_i = \theta + n_i,\ 1 \le i \le N, \tag{2.1}$$

where n_i is zero-mean measurement noise with variance σ_i^2. A commonly used statistical description for sensor noise variance is [6, 7]

$$\sigma_i^2 = \delta + \alpha z_i,\ 1 \le i \le N, \tag{2.2}$$

where δ models the network-wide noise variance threshold, α controls the underlying variation from the nominal minimum, and $z_i \sim \chi_1^2$ is a central Chi-Square distributed random variable with degrees-of-freedom equal to one[10, p-24]. Due to bandwidth and power limitations each sensor quantizes its observation into a b_i-bit message, and then transmits this locally processed data to the FC to generate a final estimate of θ. In this thesis the uniform quantization scheme with nearest-rounding [11, 12] is adopted.

The quantized message at the *i*th sensor can be modeled as

$$m_i = x_i + q_i,\ 1 \le i \le N, \tag{2.3}$$

where q_i is the quantization error which is uniformly distributed with zero mean and variance $\sigma_{q_i}^2 = R^2/12 \cdot 4^{b_i}$ [11]. $[-R/2, R/2]$ is the available signal amplitude range common to all sensors. With (2.1) and (2.3), the received data from all sensor output can be expressed in a vector form as

$$\mathbf{m} = \mathbf{1}\theta + \mathbf{n} + \mathbf{q}, \tag{2.4}$$

where $\mathbf{m} = [m_1.....m_N]^{\mathrm{T}}$, $\mathbf{1} = [1.....1]^{\mathrm{T}}$, $\mathbf{n} = [n_1.....n_N]^{\mathrm{T}}$, $\mathbf{q} = [q_1.....q_N]^{\mathrm{T}}$ and $(\bullet)^{\mathrm{T}}$ denotes the transpose.

2.1.2 Best Linear Unbiased Estimator (BLUE)

In order to generate a final estimate of θ, the Best Linear Unbiased Estimator (BLUE) [9] is used in the FC. This estimator can be determined with knowledge of only the first and second moments of the PDF. The BLUE is defined in (2.5).

$$\hat{\theta} = \sum_{n=1}^{N} a_n m[n]. \tag{2.5}$$

We observe the data set $\{m[1], m[2], ..., m[N]\}$ whose PDF $p(\mathbf{x};\theta)$ depends on an unknown parameter θ. The a_n's are constants to be determined. If we constrain this estimator to be unbiased $E(\hat{\theta}) = \theta$ and to minimize the variance $\mathrm{var}(\hat{\theta})$. Then the BLUE is given by

$$\hat{\theta} = \frac{\mathbf{1}^{\mathrm{T}}\mathbf{C}^{-1}\mathbf{m}}{\mathbf{1}^{\mathrm{T}}\mathbf{C}^{-1}\mathbf{1}}, \tag{2.6}$$

where $\mathbf{C}$ is covariance matrix. The minimum variance of the BLUE is (2.7).

$$\mathrm{var}(\hat{\theta}) = E\left(\left|\hat{\theta} - \theta\right|^2\right) = \frac{1}{\mathbf{1}^{\mathrm{T}}\mathbf{C}^{-1}\mathbf{1}}. \tag{2.7}$$

By assuming that the noise component $\{\mathbf{n}, \mathbf{q}\}$ in (2.4) are mutually independent with covariance matrices $\mathbf{C_n}$ and $\mathbf{C_q}$, then C is given by $\mathbf{C} = \mathbf{C}_n + \mathbf{C}_q$. By further assuming that the measurement noise n_i's are i.i.d, and the quantization noise q_i's are independent across all sensors, the mean square error (MSE) incurred by $\hat{\theta}$ can be immediately computed as [9]

$$E\left(\left|\hat{\theta}-\theta\right|^2\right)=\left(\sum_{i=1}^{N}\frac{1}{\sigma_i^2+R^2 4^{-b_i}/12}\right)^{-1}, \tag{2.8}$$

where σ_i^2 is defined in (2.2).

2.2 Decentralized Estimation Scheme (DES)

In this thesis, a star-like sensor network is considered. Each sensor in the network collects an observation, computes a local message and sends it to a fusion center. Sensor nodes do not communicate with each other. To reduce the communication requirement from sensors to the fusion center, local quantization/compression at each sensor site is needed. In fact, a central problem in sensor network research is to design discrete local message functions and the final fusion function in a way that minimizes the total bandwidth requirement while satisfying an overall system performance requirement. Clearly, optimal design of these functions will depend on the underlying sensor noise distributions. Unfortunately, characterizing the exact noise probability distributions for a large number of sensors is impractical, especially for applications in a dynamic sensing environment.

The decentralized estimation of a noise-corrupted deterministic parameter is considered. The sensor noises are assumed to be additive, zero mean, spatially uncorrelated, but otherwise unknown and possibly different across sensors due to varying sensor quality and inhomogeneous sensing environment. The classical BLUE linearly combines the real-valued sensor observations to minimize the MSE. Unfortunately, such a scheme cannot be implemented in a practical bandwidth-constrained sensor network due to its requirement to transmit real-valued messages. In paper [3], the authors construct a decentralized estimation scheme (DES) where each sensor compresses its observation to a small number of bits with length proportional to the local sensor signal-to-noise ratio (SNR). The resulting compressed bits from different sensors are then collected and combined by the fusion center to estimate the unknown parameter. It is shown that the MSE of the DES is within a constant factor of 25/8 to that achieved by the classical centralized BLUE estimator.

2.3 Mean Square Error (MSE) of Decentralized Estimation

That the sensor messages $\{m_i : k = 1, 2, ..., K\}$ are perfectly received by the FC with no errors is assumed. According to (2.1) and (2.3), m_i can be represented as

$$m_i = \theta + n_i + q_i,\ 1 \le i \le N, \tag{2.9}$$

$$E(m_i) = \theta, \tag{2.10}$$

$$var(m_i) = \sigma_i^2 + R^2 4^{-b_i} / 12, \tag{2.11}$$

where n_i is the sensor measurement noise and q_i is quantization noise. Therefore the final estimator is

$$\bar{\theta} = \left(\sum_{i=1}^{N} \frac{1}{var(m_i)} \right)^{-1} \sum_{i=1}^{N} \frac{m_i}{var(m_i)}. \tag{2.12}$$

Notice that $\bar{\theta}$ is an unbiased estimator of θ since every m_i is an unbiased quantization of x_i. It has an MSE :

$$
\begin{aligned}
\bar{D} &= E\left(\left|\bar{\theta}-\theta\right|^2\right) \\
&= E\left[\left(\left(\sum_{i=1}^{N}\frac{1}{\mathrm{var}(m_i)}\right)^{-1}\sum_{i=1}^{N}\frac{m_i-\theta}{\mathrm{var}(m_i)}\right)^2\right] \\
&= \left(\sum_{i=1}^{N}\frac{1}{\mathrm{var}(m_i)}\right)^{-2}\sum_{i=1}^{N}\frac{E\left(\left|m_i-\theta\right|^2\right)}{(\mathrm{var}(m_i))^2} \\
&= \left(\sum_{i=1}^{N}\frac{1}{\mathrm{var}(m_i)}\right)^{-1}. \qquad (2.13)
\end{aligned}
$$

When each m_i is transmitted to the FC through a nonperfect channel with finite power, bit error occurs. It will impact on the estimation t accuracy at the FC. The links between each sensor and the FC are modeled as a memoryless binary symmetric channel. Suppose the probability of bit error achieved by sensor i is p_b^i and m_i' is the decoded version of m_i at the receiver. Let D' denote the MSE achieved by the estimator (2.12) based on the received message $\{m_1', m_2', \ldots, m_N'\}$.According to [6], if $\{p_b^i\}$ satisfy (for some $p_0 > 0$)

$$
p_0 \geq \frac{4R}{\sigma_i}\sqrt{\frac{Np_b^i}{3}},\ 1 \leq i \leq N, \qquad (2.14)
$$

then

$$
\begin{aligned}
D' &\leq \left(1+p_0\right)^2\left(\sum_{i=1}^{N}\frac{1}{\mathrm{var}(m_i)}\right)^{-1} \\
&= \left(1+p_0\right)^2\bar{D}.
\end{aligned} \qquad (2.15)
$$

It shows that the actual achieved MSE is at most a constant factor away from what is achievable with perfect sensor channels, provided that each sensor's bit error rate (BER) is bounded above (2.14). Because the perfect MSE $\bar{D}$ is easier to derived, the upper bound of actual achieved MSE D' in (2.15) will be used to formulate the optimization problem later.

Chapter 3

Minimal Energy Decentralized Estimation Based on Sensor Noise Variance Statistics

This chapter studies minimal-energy decentralized estimation in sensor network under BLUE fusion rule. While most of the existing related works [6, 7, 8] require the knowledge of instantaneous noise variances for energy allocation, the proposed approach instead relies on an associated statistical model. Subject to severe energy and bandwidth limitation, each sensor in this scenario is allowed to transmit only a quantized version of its raw measurement to the FC to generate a final parameter estimate. While quantized message with longer bit length provide improved data fidelity, the consumed transmission energy is however proportional to the bit loads. As energy efficiency is a critical concern for sensor network design, the minimal-energy decentralized estimation problem which formulated in an optimal bit-loading setup has been recently considered.

One key feature common to the existing related works is that the energy allocated to each sensor must be determined via instantaneous local sensor noise characteristics (the noise variance), if the fusion rule follows the BLUE principle. In order to improve the estimation performance against the variation of sensing conditions, repeated update of the noise profile would be needed. This comes inevitably at the cost of more training overhead and extra energy consumption. One typical approach to resolving such a drawback is to exploit the partial (or long-term) information of the noise characteristics.

This chapter attempts to provide a solution to minimal-energy decentralized estimation by exploiting long term noise variance information. A commonly used statistical model [6, 7] for noise variance is used and the estimation performance is assessed through an MSE based metric average with respect to the considered distribution. A closed-form expression of the overall MSE requirement is derived. The analysis of the energy-minimization problem is formulated in the form of convex optimization. The problem is then analytically solved.

The proposed optimal scheme shares several interesting aspects pertaining to those based on the instantaneous noise variance information. Sensors with bad channel quality (specified via the path distance to FC) are shut off to conserve energy, and for those active nodes the allocated energy is proportional to the individual channel gain. Simulation results show that the proposed optimal solution yields significant energy saving against the equal-bit allocation policy.

3.1 Average Mean Square Error of Decentralized Estimation

For a fixed set of noise variances σ_i^2's, the energy minimization problem subject to an allowable parameter distortion γ (in terms of MSE) can be formulated as

$$\text{Min} \sum_{i=1}^{N} E_i, \text{ subject to } \left(\sum_{i=1}^{N} \frac{1}{\sigma_i^2 + \left(R^2/12\right) 4^{-b_i}} \right)^{-1} \leq \gamma, \tag{3.1}$$

where E_i is the consumed energy for transmitting the message m_i. (3.1) is equivalent to

$$\text{Min} \sum_{i=1}^{N} E_i, \text{ subject to } \sum_{i=1}^{N} \frac{1}{\sigma_i^2 + \left(R^2/12\right) 4^{-b_i}} \geq \gamma^{-1}. \tag{3.2}$$

In order to obtain universal solution with averaged measurement noise conditions, the following optimization problem is considered:

$$\text{Min} \sum_{i=1}^{N} E_i, \text{ subject to } \int_{\mathbf{z}} \sum_{i=1}^{N} \frac{1}{\delta + \alpha z_i + \left(R^2/12\right) 4^{-b_i}} p(\mathbf{z}) d\mathbf{z} \geq \gamma^{-1}, \tag{3.3}$$

where $\mathbf{z} = [z_1, z_2, ..., z_N]^{\mathrm{T}}$ with $p(\mathbf{z})$ denoting the associated distribution. In the optimization problem (3.3), the equivalent MSE performance metric in (3.2) is averaged with respect to the noise variance statistic characterized in (2.2).

To solve (3.3), a crucial step is to derive an analytic expression of the average MSE performance measure. Since $z_i \sim \chi_1^2$ is a central i.i.d. Chi-Square distributed random variable with degrees-of-freedom equal to one[10, p-24]

$$p_{\chi_1^2}(z) = \begin{cases} \dfrac{1}{\sqrt{2\pi z}} \exp(-z/2), & z \geq 0, \\ 0, & z < 0. \end{cases} \tag{3.4}$$

The average MSE performance can be derived as

$$\begin{aligned} &\int_{\mathbf{z}} \sum_{i=1}^{N} \frac{1}{\delta + \alpha z_i + \left(R^2/12\right) 4^{-b_i}} p(\mathbf{z}) d\mathbf{z} \\ &= \sum_{i=1}^{N} \int_0^{\infty} \frac{1}{\alpha z_i + \beta_i} \cdot \frac{e^{-z_i/2}}{\sqrt{2\pi z_i}} dz_i \\ &= \frac{1}{\sqrt{2\pi}} \sum_{i=1}^{N} \int_0^{\infty} \frac{e^{-z_i/2}}{(\alpha z_i + \beta_i)\sqrt{z_i}} dz_i, \end{aligned} \tag{3.5}$$

where $\beta_i = \delta + R^2 4^{-b_i} / 12$. The following lemma, with proof given in Appendix A, provides a closed-form expression of the integral involved in the summation in (3.5).

Lemma 3.1 : With $\alpha > 0$ and $\beta_i > 0$ as defined in (3.5), we have

$$\int_0^\infty \frac{e^{-z_i/2}}{(\alpha z_i + \beta_i)\sqrt{z_i}} dz_i = \frac{2\pi \cdot e^{\beta_i/2\alpha} \cdot \mathrm{Q}(\sqrt{\beta_i/\alpha})}{\sqrt{\alpha\beta_i}}, \tag{3.6}$$

where $\mathrm{Q}(x) = \int_x^\infty \frac{e^{-t^2/2}}{\sqrt{2\pi}} dt$ is the Gaussian tail function.

With (3.5) and Lemma 3.1, the optimization problem (3.3) can be equivalently rewritten as

$$\mathrm{Min} \sum_{i=1}^N E_i, \text{ subject to } \sqrt{\frac{2\pi}{\alpha}} \sum_{i=1}^N \frac{e^{\beta_i/2\alpha}\, \mathrm{Q}(\sqrt{\beta_i/\alpha})}{\sqrt{\beta_i}} \geq \gamma^{-1} . \tag{3.7}$$

Exact solution to problem (3.7) appear intractable since the target MSE is highly nonlinear in b_i. We will thus seek for the suboptimal alternatives which can otherwise admit simple analytic expression. The underlying approach toward this end is to derive an easy-to-tackle lower bound on the target MSE metric, and then replace the MSE constraint in (3.7) by one which forces the lower bound to be above γ^{-1}. Such a procedure will considerably simplify the analysis without incurring any loss in the desired MSE performance. This is done with the aid of the next lemma with proof given in Appendix B.

Lemma 3.2 : The following inequality holds:

$$\sqrt{\frac{2\pi}{\alpha}} \sum_{i=1}^N \frac{e^{\beta_i/2\alpha}\, \mathrm{Q}(\sqrt{\beta_i/\alpha})}{\sqrt{\beta_i}} \geq cN\, \mathrm{Q}\left(\frac{1}{N} \sum_{i=1}^N \left(\sqrt{\delta/\alpha} + R 2^{-b_i} / \sqrt{12\alpha} \right) \right), \tag{3.8}$$

which c is a constant defined by $c=\sqrt{\frac{2\pi}{\alpha}}\cdot\frac{e^{\delta/2\alpha}}{\sqrt{\delta+R^2/12}}$.

Lemma 3.2 suggests that we can replace the MSE constraint in (3.7) by the following one without incurring any loss in the target MSE:

$$cN\,\mathrm{Q}\left(\frac{1}{N}\sum_{i=1}^{N}\left(\sqrt{\frac{\delta}{\alpha}}+\frac{R2^{-b_i}}{\sqrt{12\alpha}}\right)\right)\geq\gamma^{-1}, \tag{3.9}$$

or equivalently

$$\frac{1}{N}\sum_{i=1}^{N}\left(\sqrt{\frac{\delta}{\alpha}}+\frac{R2^{-b_i}}{\sqrt{12\alpha}}\right)\leq\mathrm{Q}^{-1}\left(\frac{1}{cN\gamma}\right). \tag{3.10}$$

Since $\mathrm{Q}(\bullet)$ is one-to-one and monotone decreasing, we will thus instead focus on the optimization problem with a modified MSE performance constraint:

$$\mathrm{Min}\sum_{i=1}^{N}E_i,\ \text{subject to}\ \frac{R}{\sqrt{12\alpha}N}\sum_{i=1}^{N}2^{-b_i}\leq\mathrm{Q}^{-1}\left(\frac{1}{cN\gamma}\right)-\sqrt{\frac{\delta}{\alpha}}. \tag{3.11}$$

This optimization problem will lead to a simple closed-form solution.

3.2 Energy Density Factor of Sensor Nodes

We assume that each sensor sends messages to FC using a separate channel. This can be achieved by using a multiple access technique such as TDMA or FDMA. Each channel is corrupted by additive white Gaussian noise (AWGN) with power spectral density $N_0/2$:

$$\hat{m}_i=d_i^{-\kappa/2}m_i+v_i, \tag{3.12}$$

where $\hat{m}_i$ is the received message at FC and v_i is the AWGN. The signal power received at the FC is assumed to be inversely proportional to d_i^{κ} where d_i is the distance between sensor i and the FC, and κ is the path loss exponent common to all sensor-to-FC links. Suppose that message m_i has length b_i bit.

We will assume that energy E_i required for transmission of m_i is proportional to the number of bits in the message. If M-QAM is used, the consumed energy at the ith sensor is defined as

$$E_i = w_i b_i, \; 1 \le i \le N, \tag{3.13}$$

where energy density factor w_i is defined as [4, 5, 7]

$$w_i = \rho d_i^{\kappa} \cdot \frac{\left(2^s - 1\right)}{s} \cdot \ln\left(\frac{4\left(1 - 2^{-s}\right)}{sP_b}\right), \tag{3.14}$$

in which ρ is a constant depending on the noise profile, s is the number of bits per QAM symbol, and P_b is the target bit error rate. With (3.13), the specification of the energy allocated to the ith sensor amounts to determining the number of quantization bits b_i.

For a fixed set of noise variances σ_i^2's, the energy minimization problem subject to an allowable parameter distortion level γ (in terms of MSE) can be formulated as

$$\text{Min} \sum_{i=1}^{N} w_i b_i, \text{ subject to } \frac{R}{\sqrt{12\alpha}N} \sum_{i=1}^{N} 2^{-b_i} \le \mathrm{Q}^{-1}\left(\frac{1}{cN\gamma}\right) - \sqrt{\frac{\delta}{\alpha}}. \tag{3.15}$$

In (3.15), the cost function is linear and the constrain is convex. It is thus a convex optimization problem and will moreover lead to a simple closed-form solution as shown below.

3.3 Problem Formulation and Optimal Closed-form Solution

The final optimization problem is as follows

$$
\begin{aligned}
&\text{Min} \sum_{i=1}^{N} w_i b_i, \\
&\text{subject to } \frac{R}{\sqrt{12\alpha}N}\sum_{i=1}^{N} 2^{-b_i} \le \mathrm{Q}^{-1}\left(\frac{1}{cN\gamma}\right)-\sqrt{\frac{\delta}{\alpha}},\ b_i \ge 0,\ 1\le i\le N.
\end{aligned} \tag{3.16}
$$

In order to solve problem(3.16), let us form the Lagrangian function as

$$
\begin{aligned}
&\mathrm{L}\left(b_1,\ldots,b_N,\lambda,\mu_1,\ldots,\mu_N\right) \\
&=\sum_{i=1}^{N} w_i b_i+\lambda\left(\frac{R}{\sqrt{12\alpha}N}\sum_{i=1}^{N}2^{-b_i}-\mathrm{Q}^{-1}\left(\frac{1}{cN\gamma}\right)+\sqrt{\frac{\delta}{\alpha}}\right)-\sum_{i=1}^{N}\mu_i b_i.
\end{aligned} \tag{3.17}
$$

The associated set of Karush-Kuhn-Tucker (KKT) [14] conditions is as followed:

$$
w_i+\lambda\cdot\frac{(-\ln 2)R2^{-b_i}}{\sqrt{12\alpha}N}-\mu_i=0,\ 1\le i\le N, \tag{3.18}
$$

$$
\lambda\left(\frac{R}{\sqrt{12\alpha}N}\sum_{i=1}^{N}2^{-b_i}-\mathrm{Q}^{-1}\left(\frac{1}{cN\gamma}\right)+\sqrt{\frac{\delta}{\alpha}}\right)=0, \tag{3.19}
$$

$$
\lambda\ge 0,\ \mu_i\ge 0,\ \mu_i b_i=0,\ b_i\ge 0,\ 1\le i\le N. \tag{3.20}
$$

If $\lambda=0$, equation (3.18) implies $\mu_i=w_i>0$ for all $1\le i\le N$, and hence $b_i=0$, $1\le i\le N$. This case should be precluded since otherwise all sensors will remain silent.

We must have $\lambda>0$. It means that the MSE constraint in (3.16) is active so that

$$
\frac{R}{\sqrt{12\alpha}N}\sum_{i=1}^{N}2^{-b_i}=\mathrm{Q}^{-1}\left(\frac{1}{cN\gamma}\right)-\sqrt{\frac{\delta}{\alpha}}. \tag{3.21}
$$

Solving (3.18) and (3.21) leads to

$$
b_i=\log_2\left\{\frac{R\bar{\lambda}}{\sqrt{12\alpha}N\left(w_i-\mu_i\right)}\right\}, \tag{3.22}
$$

where

$$
\bar{\lambda}=\lambda\ln 2=\frac{\sum_{i=1}^{N}\left(w_i-\mu_i\right)}{\mathrm{Q}^{-1}\left(1/(cN\gamma)\right)-\sqrt{\delta/\alpha}}. \tag{3.23}
$$

By taking into account the constraint $b_i\ge 0$, the optimal pair $\left(b_i^{opt},\bar{\lambda}^{opt}\right)$ is given by the next lemma with proof given in Appendix C.

Lemma 3.3: Assume $w_1 \geq w_2 \geq \cdots \geq w_N$ without loss of generality, and define the function:

$$f(K) = \frac{\sum_{i=N-K+1}^{N} w_i}{N \cdot w_K}. \tag{3.24}$$

Let $1 \leq K_1 \leq N$ be such that $f(K_1 - 1) < 1$ and $f(K_1) \geq 1$. Then we have

$$b_i^{opt} = \begin{cases} 0, & 1 \leq i \leq N - K_1, \\ \log_2 \left\{ \dfrac{R\bar{\lambda}^{opt}}{\sqrt{12\alpha} N w_i} \right\}, & N - K_1 + 1 \leq i \leq N, \end{cases} \tag{3.25}$$

where

$$\bar{\lambda}^{opt} = \frac{\sum_{j=N-K_1+1}^{N} w_j}{\mathrm{Q}^{-1}(1/(cN\gamma)) - \sqrt{\delta/\alpha}}. \tag{3.26}$$

3.4 Discussions of Optimal Solution

1. The target distortion level γ cannot be set unlimitedly small. It is otherwise bounded by the MSE attained by the benchmark estimate based on un-quantized real-valued sensor measurements (i.e., the case when $b_i = \infty$, $1 \leq i \leq N$). By setting $b_i = \infty$ in the average MSE formula specified in (3.7), the minimal allowable γ can be immediately determined as

$$\gamma \geq \left[N e^{\delta/2\alpha} \mathrm{Q}\left(\sqrt{\frac{\delta}{\alpha}} \right) \sqrt{\frac{2\pi}{\alpha\delta}} \right]^{-1}. \tag{3.27}$$

2. Since $0 \leq b_i < \infty$, a necessary condition for validating the MSE constraint in (3.15) is therefore

$$Q^{-1}\left(\frac{1}{cN\gamma}\right)-\sqrt{\frac{\delta}{\alpha}}\geq 0. \tag{3.28}$$

By definition of the constant c in Lemma 3.2 and with (3.28), the MSE attainable by the proposed method is lower bounded by

$$\gamma \geq \left[Ne^{\delta/2\alpha}\, Q\left(\sqrt{\frac{\delta}{\alpha}}\right)\sqrt{\frac{2\pi}{\alpha\left(\delta+R^2/12\right)}}\right]^{-1}. \tag{3.29}$$

The lower bound (3.29) is indeed larger than the lower bound (3.27).

3. In (3.14), the energy density factor w_i is proportional to the path loss d_i^{κ}, if the same bit error rate is assumed throughout all the links. The large values of w_i correspond to the sensors deployed far away from the FC. They are usually with poor background channel gain. In this point, the proposed optimal solution (3.25) is intuitively attractive. The sensors with large w_i are turned off to conserve energy. A similar energy conservation strategy via shutting off the sensors with poor channel links is found in [6], in which a scenario with instantaneous noise variance available to the FC is considered.
4. From(3.25), the assigned message length is inversely proportional to w_i for those active sensors. This is intuitively reasonable since sensors with better link conditions should be allocated with more bits to realize desired performance.
5. Based on the inequality constraint for average MSE in (3.15), the equal-bit schemes maintaining the desired MSE can be obtained by solving

$$\frac{R2^{-\tilde{b}}}{\sqrt{12\alpha}}=Q^{-1}\left(\frac{1}{cN\gamma}\right)-\sqrt{\frac{\delta}{\alpha}}. \tag{3.30}$$

It leads to

$$\tilde{b} = \log_2 \left\{ \frac{R}{\sqrt{12\alpha}\left[\mathrm{Q}^{-1}\left(1/(cN\gamma)\right) - \sqrt{\delta/\alpha} \right]} \right\}. \tag{3.31}$$

Numerical simulations in the next section show that the proposed optimal scheme (3.25) yields significant energy saving when compared with equal-bit scheme (3.31).

3.5 Numerical Simulation

For a fixed set of energy density factors w_i, $1 \le i \le N$, the performance is measured via the percentage of energy saving (PES) [6, 7]:

$$PES = \frac{\tilde{b}\sum_{i=1}^{N} w_i - \sum_{i=1}^{N} w_i b_i^{opt}}{\tilde{b}\sum_{i=1}^{N} w_i} \times 100, \tag{3.32}$$

where b_i^{opt} and $\tilde{b}$ are defined respectively in (3.25) and (3.31). We simply set $w_i = d_i^{\kappa}$ where $\kappa = 3.5$ and $d_i = 10 + 10Z_i$ with $Z_i \sim \chi_1^2$ being i.i.d. Chi-Square distributed random variable. The results are averaged over 50000 independent trials. The total number of sensors is $N=1500$ under $\gamma = 0.005$.

The Figure 3.1(a) shows the PES for $0.1 \le \alpha \le 1.6$ and Figure 3.1(b) depicts the computed $\tilde{b}$ in (3.31) with fixed $\delta = 0.8$. That the PES exhibits two “jumps” can be observed. This accounts for the two level change of $\tilde{b}$ as α varies. Within each duration of constant $\tilde{b}$, energy efficiency of the optimal solution improves as α increases (a large α corresponds to a more inhomogeneous sensing environment). We note that a similar phenomenon has been observed in the existing works relying on instantaneous noise variance knowledge [6, 7]. When the sensing condition becomes more inhomogeneous, it is more likely that a large fraction of sensors suffers from poor measurement quality and will be shut off. It leads to improved energy efficiency. Since the proposed solution (3.25) based on statistical noise variance description would reflect the long-term characteristic of the schemes [6, 7], this consistency is expected.

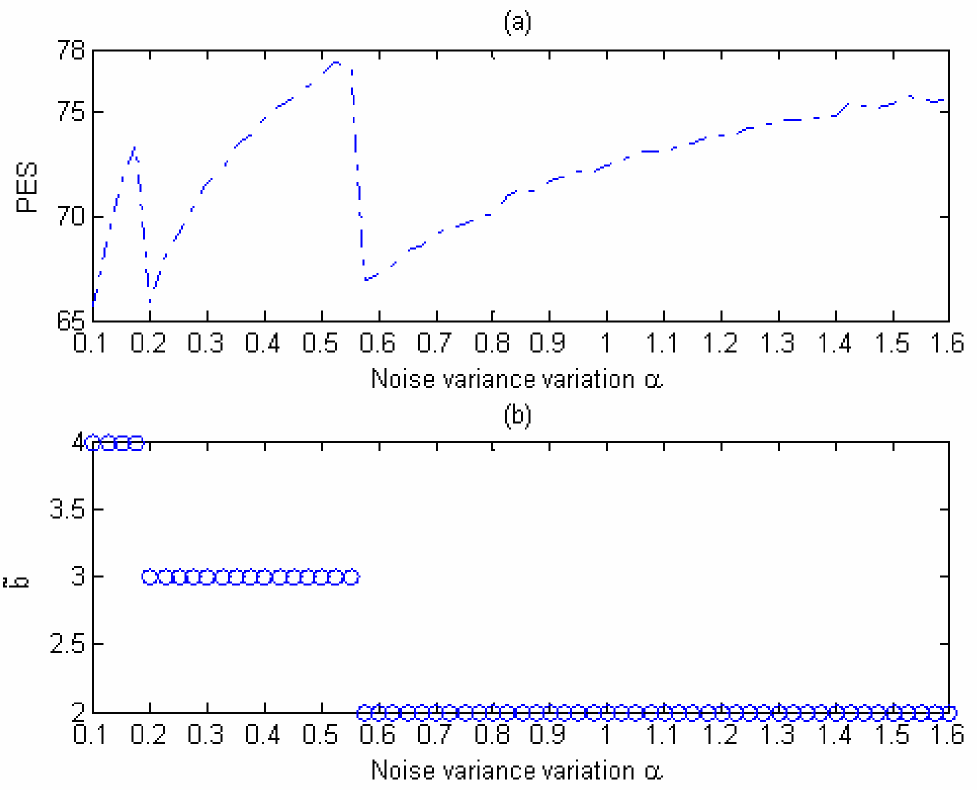

Figure 3.1 : PES for fixed minimal noise variance threshold ($\delta = 0.8$)

We repeat the experiment by fixing $\alpha = 0.4$ and varying the minimal threshold δ. The results are shown in Figure 3.2. Obviously, the PES exhibits a counter tendency as compared to Figure 3.1. For each duration of constant $\tilde{b}$, the energy saving achieved by proposed optimal solution is lower as δ increases. This is reasonable because the large minimal noise variance threshold results in severe noise corruption in all sensor measurement. More sensor nodes should be turned on to provide a sufficient amount of information for MSE reduction.

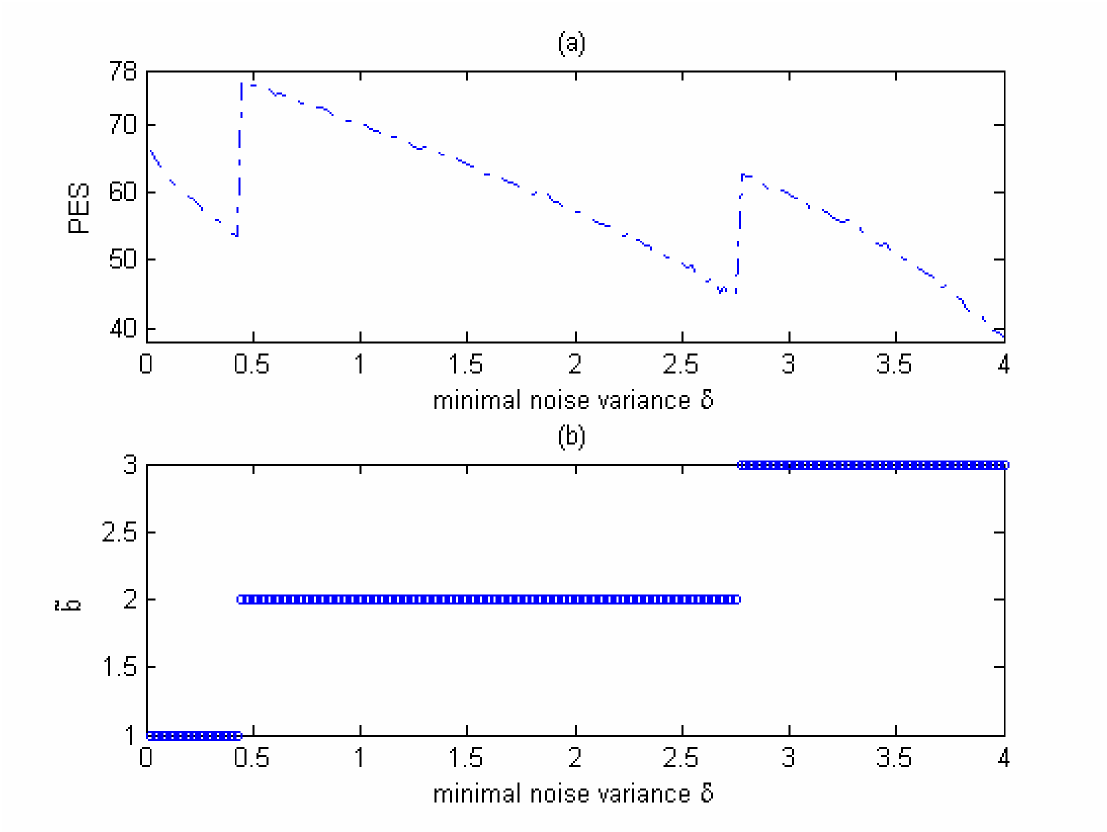

Figure 3.2 : PES for fixed noise variance variation ($\alpha = 0.4$)

3.6 Summary

This chapter provides a solution to the minimal-energy decentralized estimation problem by exploiting a statistical noise variance model. Based on a closed-form expression of the MSE performance measure averaged over the noise variance distribution, energy minimization is reformulated as convex optimization problem. The proposed solution simply allocates energies to sensors with large channel gain and shut off those suffering from poor link quality. Numerical simulation shows that the proposed optimal solution is capable of reducing about 80% energy consumption when compared with the uniform-allocation scheme. The energy saving efficiency is particularly significant when the minimal measurement noise variance threshold is small or the variation factor is large.

Chapter 4

Minimal Mean Square Error Decentralized Estimation Based on Sensor Noise Variance Statistics

Relying on partial noise variance knowledge in the form of the background distribution, the problem of minimizing total transmission energy under an allowable average distortion level is recently considered in [15]. This chapter considers the counterpart problem: how to find the optimal bit load which minimizes the average distortion under a fixed total energy budget. The main contribution of the current work can be summarized as follows:

i. While the design metric, the reciprocal of the average MSE is shown in [15] to be highly nonlinear in the sensor bit load. Several analytic approximation relations are used to derive an associated tractable low bound.

ii. By maximizing this lower bound, the problem can be further formulated in the form of convex optimization which yields a closed-form solution.

The analytic results reveal that under limited energy budget, sensors with bad link quality should be shut off toward utmost estimation accuracy, and energy allocated to those active nodes should be proportional to the individual channel gain. A similar energy conservation policy is also found in the previous work [6, 7, 15]. Numerical simulations show the effectiveness of the proposed scheme which outperforms the uniform allocation strategy under an energy-limited environment.

4.1 Average Mean Square Error of Decentralized Estimation

The MMSE decentralized estimation which is counterpart problem of (3.1) can be formulated as

$$\text{Min}\left(\sum_{i=1}^{N}\frac{1}{\sigma_i^2+\beta 4^{-b_i}}\right)^{-1}, \text{ subject to } \sum_{i=1}^{N} E_i \le E_T, \tag{4.1}$$

or equivalently,

$$\text{Max}\sum_{i=1}^{N}\frac{1}{\sigma_i^2+\beta 4^{-b_i}}, \text{ subject to } \sum_{i=1}^{N} E_i \le E_T, \tag{4.2}$$

where $\beta = R^2/12$ and E_T is allowable energy level. The equivalent MSE cost function is averaged with respect to the noise variance statistic characterized in (2.2):

$$\text{Max}\int_{\mathbf{z}}\sum_{i=1}^{N}\frac{1}{\delta+\alpha z_i+\beta 4^{-b_i}}p(\mathbf{z})\,d\mathbf{z}, \text{ subject to } \sum_{i=1}^{N} E_i \le E_T, \tag{4.3}$$

where $\mathbf{z} = [z_1, z_2, \ldots., z_N]^{\mathrm{T}}$ with $p(\mathbf{z})$ denoting the associated distribution. In order to solve problem (4.3), the first step is to find an analytic expression of the equivalent average MSE metric.

By equation (3.5) and lemma 3.1, problem (4.3) can be equivalently rewritten as

$$\text{Max } \sqrt{2\pi}\cdot\sum_{i=1}^{N}\frac{e^{\left(\delta+\beta 4^{-b_i}\right)/2\alpha}\cdot \mathrm{Q}\left(\sqrt{\left(\delta+\beta 4^{-b_i}\right)/\alpha}\right)}{\sqrt{\alpha\left(\delta+\beta 4^{-b_i}\right)}}, \tag{4.4}$$

$$\text{subject to } \sum_{i=1}^{N} E_i \le E_T.$$

Exact solution to the considered optimization (4.4) appears formidable to tackle because the cost function is highly nonlinear in b_i. An alternative formulation which is more tractable is proposed and an analytic solution can be obtained. By the following approximation to $Q(.)$ function [16, p115]

$$\mathrm{Q}(x)\approx\frac{1}{\sqrt{2\pi}}\left[\frac{e^{-x^2/2}}{\left(1-\pi^{-1}\right)x+\pi^{-1}\sqrt{x^2+2\pi}}\right], \tag{4.5}$$

and some straightforward manipulations, the cost function can be approximated by

$$\begin{aligned}&\sqrt{2\pi}\cdot\sum_{i=1}^{N}\frac{e^{\left(\delta+\beta 4^{-b_i}\right)/2\alpha}\cdot \mathrm{Q}\left(\sqrt{\left(\delta+\beta 4^{-b_i}\right)/\alpha}\right)}{\sqrt{\alpha\left(\delta+\beta 4^{-b_i}\right)}}\\ &\approx\sum_{i=1}^{N}\frac{1}{\left(1-\pi^{-1}\right)\left(\delta+\beta 4^{-b_i}\right)+\pi^{-1}\sqrt{\left(\delta+\beta 4^{-b_i}\right)^2+2\pi\alpha\left(\delta+\beta 4^{-b_i}\right)}}.\end{aligned} \tag{4.6}$$

The main advantage of (4.6) is that it can lead to an associated lower bound in a more tractable form. Thought maximizing this lower bound we can eventually obtain a closed-form optimal solution. By the inequality equation:

$$\sqrt{\left(\delta+\beta 4^{-b_i}\right)^2+2\pi\alpha\left(\delta+\beta 4^{-b_i}\right)}\le\left(\delta+\beta 4^{-b_i}\right)+\pi\alpha, \tag{4.7}$$

the approximated cost function in (4.6) can be lower bounded by

$$
\begin{aligned}
&\sum_{i=1}^{N} \frac{1}{\left(1-\pi^{-1}\right)\left(\delta+\beta 4^{-b_i}\right)+\pi^{-1}\sqrt{\left(\delta+\beta 4^{-b_i}\right)^2+2\pi\alpha\left(\delta+\beta 4^{-b_i}\right)}} \\
&\geq \sum_{i=1}^{N} \frac{1}{\left(1-\pi^{-1}\right)\left(\delta+\beta 4^{-b_i}\right)+\pi^{-1}\left[\left(\delta+\beta 4^{-b_i}\right)+\pi\alpha\right]} \\
&= \sum_{i=1}^{N} \frac{1}{\left(\delta+\beta 4^{-b_i}\right)+\alpha} \\
&= \sum_{i=1}^{N} \frac{4^{b_i}}{\beta+\left(\alpha+\delta\right)4^{b_i}}.
\end{aligned} \tag{4.8}
$$

We will thus focus on maximizing the lower bound:

$$\text{Max} \sum_{i=1}^{N} \frac{4^{b_i}}{\beta+\left(\alpha+\delta\right)4^{b_i}}, \text{ subject to } \sum_{i=1}^{N} E_i \leq E_T. \tag{4.9}$$

The cost function is simple in (4.9). It can lead to an analytic solution of the optimization problem.

4.2 Energy Density Factor of Sensor Nodes

We assume that each sensor sends messages to FC using a separate channel. This can be achieved by using a multiple access technique such as TDMA or FDMA. Each channel is corrupted by additive white Gaussian noise (AWGN) with power spectral density $N_0/2$:

$$\hat{m}_i = d_i^{-\kappa/2} m_i + v_i, \tag{4.10}$$

where $\hat{m}_i$ is the received message at FC and v_i is the AWGN. The signal power received at the FC is assumed to be inversely proportional to d_i^{κ} where d_i is the distance between sensor i and the FC, and κ is the path loss exponent common to all sensor-to-FC links.

We assume that the ith sensor sends the b_i-bit message m_i by using quadratic amplitude modulation with a constellation size 2^{bi}. The consumed energy is [4, 5, 6, 17]

$$E_i = w_i\left(2^{b_i} - 1\right),\ 1 \le i \le N. \tag{4.11}$$

The energy density factor w_i is defined as

$$w_i = \rho d_i^{\kappa} \cdot \ln\left(\frac{2}{P_b}\right), \tag{4.12}$$

in which ρ is a constant depending on the noise profile, and P_b is the target bit error rate assumed common to all sensor-to-FC links.

With (4.11), the specification of the energy allocated to the ith sensor amounts to determining the number of quantization bits b_i. For a fixed set of noise variances σ_i^2's, the MSE minimization problem subject to an allowable energy level E_T can be formulated as

$$\text{Max} \sum_{i=1}^{N} \frac{4^{b_i}}{\beta + (\alpha + \delta) 4^{b_i}}, \text{ subject to } \sum_{i=1}^{N} w_i\left(2^{b_i} - 1\right) \le E_T. \tag{4.13}$$

Since $b_i \ge 0$, it follows $\sum_{i=1}^{N} w_i\left(2^{b_i} - 1\right) \le \sum_{i=1}^{N} w_i\left(4^{b_i} - 1\right)$. This implies that we can replace the total energy constraint in (4.13) by the following one without violating the overall energy budget requirement:

$$\sum_{i=1}^{N} w_i\left(4^{b_i} - 1\right) \le E_T. \tag{4.14}$$

With the aid of (4.14) and by performing a change of variable with $B_i = 4^{b_i} - 1$, the optimization problem then becomes

$$\text{Max} \sum_{i=1}^{N} \frac{B_i + 1}{(\alpha + \beta + \delta) + (\alpha + \delta) B_i}, \text{ subject to } \sum_{i=1}^{N} w_i B_i \le E_T. \tag{4.15}$$

In (4.15), the intermediate variable B_i is relaxed to be a nonnegative real number so as to render the problem tractable. While the optimal real-valued B_i is computed, the associated bit loads can be obtain through upper integer rounding. The major advantage of the alternative problem formulation is that it admits the form of convex optimization and can lead to a simple closed-form solution. It is shown in the next section.

4.3 Problem Formulation and Optimal Closed-form Solution

The finial optimization problem is as followed

$$\text{Max} \sum_{i=1}^{N} \frac{B_i + 1}{(\alpha+\beta+\delta)+(\alpha+\delta)B_i},$$
$$\text{subject to} \sum_{i=1}^{N} w_i B_i \le E_T,\ B_i \ge 0,\ 1 \le i \le N. \tag{4.16}$$

In order to solve problem (4.16), let us form the Lagrangian as

$$L(b_1,\ldots,b_N,\lambda,\mu_1,\ldots,\mu_N)$$
$$= \sum_{i=1}^{N} \frac{B_i + 1}{(\alpha+\beta+\delta)+(\alpha+\delta)B_i} - \lambda\left(\sum_{i=1}^{N} w_i B_i - E_T\right) + \sum_{i=1}^{N} \mu_i B_i. \tag{4.17}$$

The associated set of KKT conditions [14] is as followed:

$$\frac{\beta}{\left[(\alpha+\beta+\delta)+(\alpha+\delta)B_i\right]^2} - \lambda w_i + \mu_i = 0,\ 1 \le i \le N, \tag{4.18}$$

$$\lambda\left(\sum_{i=1}^{N} w_i B_i - E_T\right) = 0, \tag{4.19}$$

$$\lambda \ge 0,\ \mu_\mathrm{i} \ge 0,\ \mu_\mathrm{i} b_i = 0,\ \mathrm{b_i} \ge 0,\ 1 \le \mathrm{i} \le \mathrm{N}. \tag{4.20}$$

The condition (4.18) leads to

$$B_i = \frac{1}{\alpha+\delta}\sqrt{\frac{\beta}{\lambda w_i - \mu_i}} - \left(1 + \frac{\beta}{\alpha+\delta}\right). \tag{4.21}$$

If $\lambda = 0$, equation (4.18) implies $\mu_i > 0$ for all $1 \le i \le N$ and hence

$b_i = 0,\ \ 1 \le \mathrm{i} \le \mathrm{N}$. This case should be precluded since all sensors are turned off. From (4.19) and (4.21), $\sqrt{\lambda}$ can be obtained:

$$\sqrt{\lambda} = \frac{\sqrt{\beta}}{\alpha+\delta}\left(\sum_{i=1}^{N}\sqrt{w_i}\right)\left(E_T + \left(1+\frac{\beta}{\alpha+\delta}\right)\sum_{i=1}^{N} w_i\right)^{-1}. \tag{4.22}$$

In (4.21), λ and μ_i 's should be determined to fulfill the desired constraints.

$w_1 \ge w_2 \ge \cdots\cdots \ge w_N$ is assumed without loss of generality and we define the function

$$f\left(K\right) = \frac{E_T\left(1+\dfrac{\beta}{\alpha+\delta}\right)^{-1} + \sum\limits_{i=K}^{N} w_i}{\sqrt{w_K}\sum\limits_{i=K}^{N}\sqrt{w_i}},\ 1 \le K \le N. \tag{4.23}$$

Let $1 \le K_1 \le N$ be the unique integer such that $f\left(K_1 - 1\right) < 1$ and $f\left(K\right) \ge 1$. If $f\left(K\right) \ge 1$ for all $1 \le K \le N$, then set $K_1 = 1$. The existence and uniqueness of such K_1 is shown in Lemma 4.1 with proof given in Appendix D.

Lemma 4.1 : $f\left(K\right)$ defined in (4.23) is monotone increasing and $f\left(N\right) > 1$.

If K_1 such that $f\left(K_1\right) \ge 1$ exists, then K_1+1 will lead to $f\left(K_1+1\right) \ge 1$.

The optimal solution pair $\left(B_i^{opt}, \lambda^{opt}\right)$ is given by

$$B_i^{opt} = \begin{cases} 0, & 1 \le i \le K_1 \text{ -1}, \\ \dfrac{1}{\alpha+\delta}\sqrt{\dfrac{\beta}{\lambda^{opt} w_i}} - \left(1+\dfrac{\beta}{\alpha+\delta}\right), & K_1 \le i \le N, \end{cases} \tag{4.24}$$

where

$$\sqrt{\lambda^{opt}} = \frac{\sqrt{\beta}}{\alpha+\delta}\left(\sum_{i=K_1}^{N}\sqrt{w_i}\right)\left(E_T + \left(1+\frac{\beta}{\alpha+\delta}\right)\sum_{i=K_1}^{N} w_i\right)^{-1}. \tag{4.25}$$

Since $B_i = 4^{b_i} - 1$ and with (4.24) (4.25), the optimal bit load is

$$b_i^{opt} = \begin{cases} 0, & 1 \le i \le K_1 - 1, \\ \frac{1}{2}\log_2\left\{\left[\sqrt{w_i}\sum_{j=K_1}^{N}\sqrt{w_j}\right]^{-1}\left[E_T + \left(1+\frac{\beta}{\alpha+\delta}\right)\sum_{j=K_1}^{N} w_j\right] - \frac{\beta}{\alpha+\delta}\right\}, & \mathrm{K}_1 \le i \le N. \end{cases}$$

(4.26)

The resultant average distortion level then equals

$$\overline{MSE} = \left(\sqrt{2\pi}\cdot\sum_{i=K_1}^{N}\frac{e^{\left(\delta+\beta 4^{-b_i^{opt}}\right)/2\alpha} \mathrm{Q}\left(\sqrt{\left(\delta+\beta 4^{-b_i^{opt}}\right)/\alpha}\right)}{\sqrt{\alpha\left(\delta+\beta 4^{-b_i^{opt}}\right)}} \right)^{-1}. \tag{4.27}$$

4.4 Discussions of Optimal Solution

1. The minimal average MSE is attained when all the raw sensor measurements with infinite-precision (i.e., $b_i = 0,\ 1 \le i \le N$) are available to the FC. Hence, by setting $b_i = \infty$ in the mean MSE formula specified (4.4), we have the following performance bound

$$MSE_{\min} = \left[N e^{\delta/2\alpha}\,\mathrm{Q}\left(\sqrt{\delta/\alpha}\right)\sqrt{\frac{2\pi}{\alpha\delta}} \right]^{-1}. \tag{4.28}$$

   Formula (4.28) reveals the impacts of the noise model parameters α and δ on the estimation performance. It is easy to see from (4.28) that the minimal MSE increases with α. This implies the estimation accuracy degrades as the sensing environment becomes more and more inhomogeneous (corresponding to large α). Furthermore it can be checked that $MSE_{\min}$ also increases with the minimal noise power threshold δ. This is reasonable since large δ implies poor measurement quality of all sensor data and a less accurate parameter estimate. Although these facts are inferred based on the idealized distortion measure (4.28), similar tendency is also observed for $\overline{MSE}$ in (4.27) attained with sensor data quantization.

2. The energy density factor w_i is proportional to the path loss d_i^{κ} (assuming all links with the same κ). Large value of w_i correspond to sensors deployed far away from the FC, usually with poor background channel gains. In this point, the proposed optimal solution is intuitively attractive. The sensors associated with the *(K_1-1)*th largest w_i's are turned off to conserve energy. A similar energy conservation strategy via shutting off sensors with poor channel links is also found in [6, 7, 15]. From (4.26), the assigned message length for those active nodes is inversely proportional to $\sqrt{w_i}$. This is intuitively reasonable since sensors with better link conditions should be allocated with more bits (energy) to improve the estimation accuracy.

3. In order to prevent sensors from exhausting energy quickly, one natural way is to impose an additional peak energy constraint:

$$w_i\left(2^{b_i}-1\right)\le E_P,\ 1\le i\le N. \tag{4.29}$$

In optimization problem (4.16), with extra inequality requirement (4.29), there does not seem to exist a closed-form optimal solution. As a simple suboptimal alternative, we can first identify the infeasible node index set $\Gamma=\left\{i\middle|w_i\left(2^{b_i^{opt}}-1\right)>E_P,\ \ K_1\le i\le N\right\}$ from (4.26) and then instead fix the energy associated with each of these nodes to be E_P. The resultant solution is thus

$$\tilde{b}_i=\begin{cases}0, & 1\le i\le \mathrm{K}_1-1,\\ \dfrac{1}{2}\log_2\left\{\left[\sqrt{w_i}\displaystyle\sum_{j=K_i}^{N}\sqrt{w_j}\right]^{-1}\left[E_T+\left(1+\dfrac{\beta}{\alpha+\delta}\right)\displaystyle\sum_{j=K_1}^{N}w_i\right]-\dfrac{\beta}{\alpha+\delta}\right\}, & \mathrm{K}_1\le i\le N \text{ and } i\notin\Gamma,\\ \log_2\left(1+E_P/w_i\right), & \mathrm{K}_1\le i\le N \text{ and } i\in\Gamma.\end{cases}$$

(4.30)

The actual solution can be obtained by using the iterative procedures reported in [18] with (4.30) as an initialization point. The algorithm to derive the optimal analytical solution is followed.

(1) Solve the problem without individual power constraints (4.16) to obtain the solution (4.26).

Set the index set $\Gamma = \left\{ i \middle| w_i \left(2^{b_i^{opt}} - 1 \right) > E_P, \;\; K_1 \le i \le N \right\}$.

(2) Set $b_i^{opt} = \log_2 \left(1 + E_P / w_i \right)$ for $i \in \Gamma$.

Set $E_T = E_T - \sum_{i \in \Gamma} w_i \left(2^{b_i} - 1 \right)$.

Remove b_i for $i \in \Gamma$ from the design variable space.

(3) Repeat the first and second steps until Γ is empty in the first step.

To prove that the algorithm leads to the global optimum, we need only to prove that in the second step we do not lose optimality of b_i^{opt} for $i \in \Gamma$ when we set $b_i^{opt} = \log_2 \left(1 + E_P / w_i \right)$ for $i \in \Gamma$.

4.5 Numerical Simulation

We compare the simulated performance of proposed optimal solution (4.26) against the uniform energy allocation scheme with bit load determined through

$$w_i \left(2^{b_i} - 1 \right) = E_T / N, \; 1 \le i \le N. \tag{4.31}$$

In (4.31), b_i is computed via lower integer rounding so that the resultant total energy can be kept below E_T. It leads to

$$\tilde{b} = \log_2 \left(\frac{E_T}{N w_i} + 1 \right). \tag{4.32}$$

In each independent run we simply choose $w_i = d_i^{\kappa}$, where $\kappa = 2$ and $d_i = 0.5 + 0.3Z_i$ with $Z_i \sim \chi_1^2(z)$ being i.i.d.. The total number of trial is 50000. In the following experiments we set the number of sensors to be *N=200*, and consider three different levels of total energy. $E_T = \gamma \sum_{i=1}^{N} w_i$ with $\gamma = 0.25,\ 1,\ 3$ respectively correspond to the low, medium, and high energy cases.

With fixed $\delta = 2$, Figure 4.1 shows the computed average MSE as α varies from 0.5 to 8. With fixed $\alpha = 2$, Figure 4.2 shows the average MSE as δ varies from 0.5 to 8. Both figures show that the estimation accuracy improves as E_T increases. It is expected. The proposed solution (4.26) outperforms uniform energy allocation (4.32), especially when E_T is small. It is more effective in an energy-limited environment. The simulated average MSE increases with both α and β.

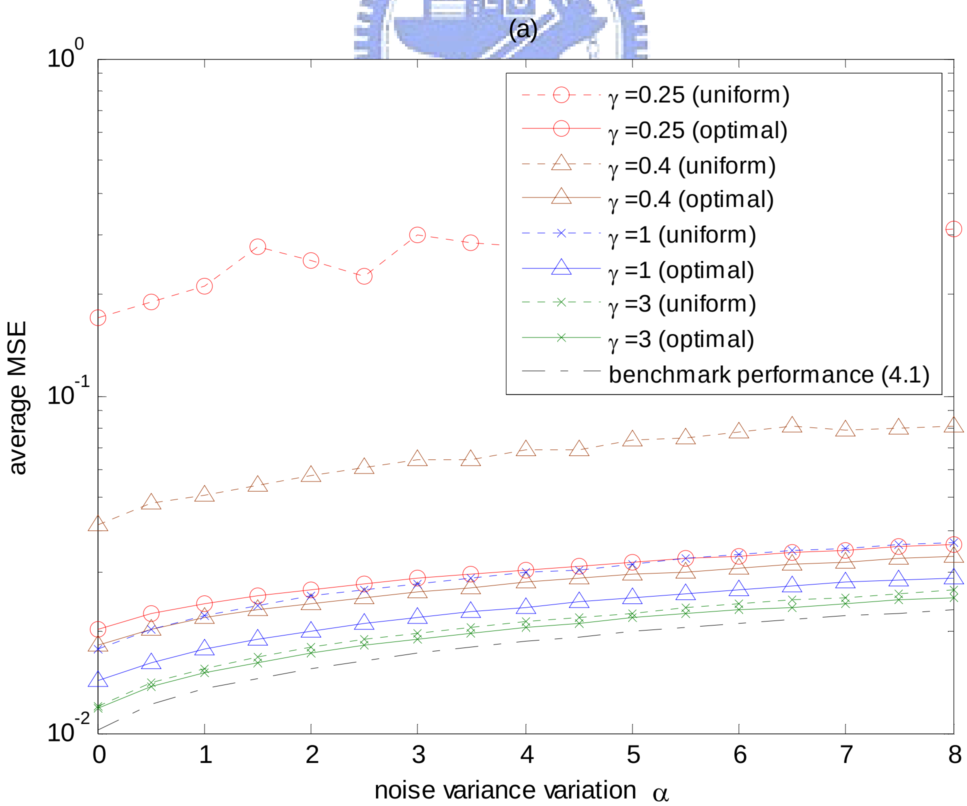

Figure 4.1 : Average MSE for fixed minimal noise variance threshold ($\delta = 2$)

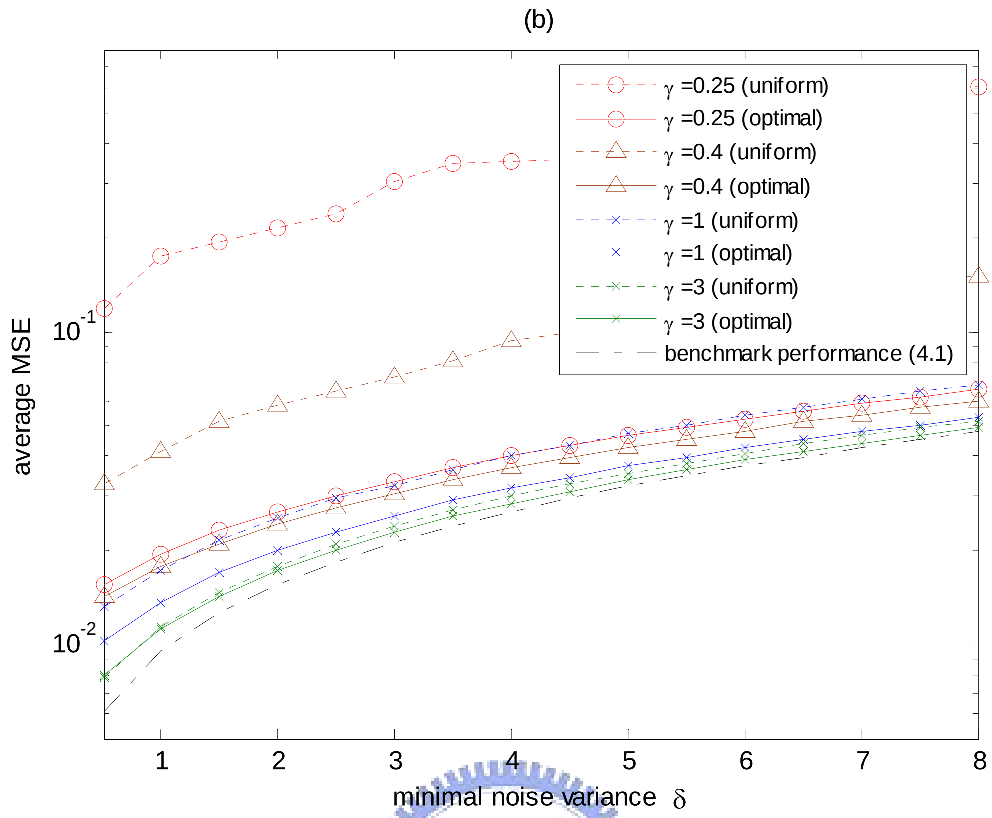

Figure 4.2 : Average MSE for fixed noise variance variation ($\alpha = 2$)

4.6 Summary

This chapter provides a solution to the minimal-MSE decentralized estimation problem by exploiting a statistical noise variance model. Based on a closed-form expression of the MSE performance measure averaged over the noise variance distribution, MSE minimization is reformulated as convex optimization problem. The analytic closed-form solution reveals the energy saving policy. The proposed solution simply allocates energies to sensors with a large channel gain and shut off those suffering from poor link quality. Numerical simulation shows that the estimation accuracy improves as total energy increases. The proposed solution outperforms uniform energy allocation especially when the total used energy is small, and thus is more effective in an energy-limited environment.

Chapter 5

Minimal Energy Decentralized Estimation over Rayleigh Fading Channel Based on Sensor Noise Variance Statistics

As energy efficiency is a critical concern for sensor network design, the minimal-energy decentralized estimation problem which is formulated in an optimal bit-loading setup has been recently considered. In order to improve the estimation performance against the variation of sensing conditions, repeated update of the noise profile would be needed. This comes inevitably at the cost of more training overhead and extra energy consumption. One typical approach to resolving such a drawback is to exploit the partial (or long-term) information of the noise characteristics.

Another key feature common to the existing related works [1, 6, 7] is that they all assume error-free transmission. They consider the sensors experiencing the perfect wireless channel. There is no bit error in the wireless channels between sensors and the FC. The work in [6] uses the upper bound (2.15) to show that the actual achieved MSE is at most a constant factor away from what is achievable with perfect sensor channels. It use the MSE constraint with perfect channel to formulate the convex optimization problem which derive an optimal bit loading scheme. Chapter 3 and Chapter 4 of my thesis also use the MSE constraint with perfect channel to formulate the optimization problems instead by the long-term information of the noise characteristics. The work in [19] considers the noisy channel between each sensor and the FC by modeling it as a binary symmetric channel (BSC) model with crossover probability which is controlled by the transmitted bit energy, but it uses the instantaneous local sensor noise characteristics to formulate the optimization problem.

This chapter attempts to provide a solution to the minimal-energy decentralized estimation with the noisy channel between each sensor and the FC by exploiting long term noise variance information. A commonly used statistical model [6, 7] for noise variance is used and the estimation performance is assessed through an MSE based metric average with respect to the considered distribution. The BSC models [19] are used to characterize the wireless multi-path fading channels with path loss. A closed-form expression for the overall MSE requirement is derived. The analysis of the energy-minimization problem is formulated in the form of convex optimization. The problem is then analytically solved.

The proposed suboptimal scheme shares several interesting aspects pertaining to those based on the instantaneous noise variance information. Sensors with bad channel quality (specified via the path distance to FC) are shut off to conserve energy, and for those active nodes the allocated energy is proportional to the individual channel gain. Simulation results show that the proposed optimal solution yields energy saving against the equal-bit allocation policy.

5.1 System Model

There are N spatially deployed sensors which cooperate with a FC for estimating an unknown deterministic parameter θ where $\theta \in [0,\ 1]$. In order to simplify the following analysis, we set $\theta \in [0,\ 1]$ which is a special case for general case $\theta \in [-R/2,\ R/2]$ where R is the parameter range. The following analytic results for the general case and special case are different in a constant factor.

The local observation at the ith node is

$$x_i = \theta + n_i,\ \ 1 \le i \le N, \tag{5.1}$$

where n_i is a zero-mean measurement noise with variance [6, 7]

$$\sigma_i^2 = \delta + \alpha z_i. \tag{5.2}$$

In (5.2), δ models the network-wide noise variance threshold, α controls the underlying variation from the nominal minimum, and $z_i \sim \chi_1^2$ is a central Chi-Square distributed random variable with degrees-of-freedom equal to one. Due to bandwidth and power limitations each sensor quantizes its observation into a b_i-bit message, and then transmits this locally processed data to the FC to generate a final estimate of θ.

The uniform quantization scheme with nearest-rounding is adopted. The quantized message at the ith sensor can be modeled as

$$m_i = x_i + q_i,\ 1 \le i \le N, \tag{5.3}$$

where q_i is the quantization error which is uniformly distributed with zero mean and variance $\sigma_{q_i}^2 = 1\big/\left(12 \cdot 4^{b_i}\right)$, and [0, 1] is the available signal amplitude range common to all sensors. The quantized value m_i can be modeled as

$$m_i = \sum_{k=1}^{b_i} a_k^{(i)} 2^{-k}, \tag{5.4}$$

where $a_k^{(i)}$ is the kth quantization bit of the ith sensor in which the quantization bit length is b_i. The quantization bit $a_k^{(i)}$ is transmitted through the wireless channel to the FC and is received as $\hat{a}_k^{(i)}$. The messages received in the FC from the ith sensor can be modeled as

$$y_i = \sum_{k=1}^{b_i} \hat{a}_k^{(i)} 2^{-k}, \tag{5.5}$$

where $\hat{a}_k^{(i)}$ is the kth quantization bit received in the FC from the ith sensor.

For simplicity, we consider only uncoded transmissions and channels that are memoryless with different bits experiencing independent fading effects. Under these conditions, we can model the wireless air-interface between the ith sensor and the FC as a binary symmetric channel (BSC) with crossover probability ε_i. The BSC model shown in Figure 5.1 can be used to characterize a more general class of channels including multi-path fading.

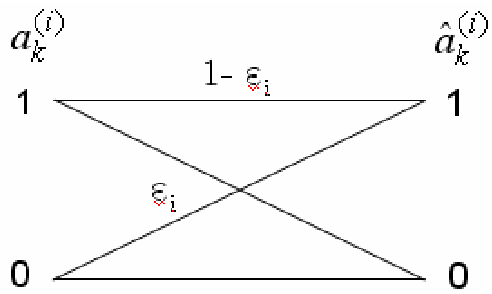

Figure 5.1 : Binary symmetric channel

The received message y_i in the FC from the ith sensor can thus be model as

$$y_i = m_i + c_i, \tag{5.6}$$

where c_i is the wireless channel error induced by the BSC with crossover probability

ε_i. The received data in the FC can be expressed in a vector form as

$$\mathbf{y} = \mathbf{1}\theta + \mathbf{n} + \mathbf{q} + \mathbf{c}, \tag{5.7}$$

where $\mathbf{y} = [y_1, y_2, ..., y_N]^{\mathrm{T}}$, $\mathbf{1} = [1, 1, ..., 1]^{\mathrm{T}}$, $\mathbf{n} = [n_1, n_2, ..., n_N]^{\mathrm{T}}$, $\mathbf{q} = [q_1, q_2, ..., q_N]^{\mathrm{T}}$, $\mathbf{c} = [c_1, c_2, ..., c_N]^{\mathrm{T}}$, and $(\bullet)^{\mathrm{T}}$ denotes the transpose.

We focus on linear fusion rules for parameter recovery. By assuming that the noise component $\{\mathbf{n}, \mathbf{q}, \mathbf{c}\}$ in (5.7) are mutually independent with covariance matrices $\mathbf{C_n}$, $\mathbf{C_q}$ and $\mathbf{C_c}$, the parameter θ is retrieved by the BLUE estimator via

$$\hat{\theta} = \frac{\mathbf{1}^{\mathrm{T}}\mathbf{C}^{-1}\mathbf{m}}{\mathbf{1}^{\mathrm{T}}\mathbf{C}^{-1}\mathbf{1}}, \quad \text{where } \mathbf{C} = \mathbf{C_n} + \mathbf{C_q} + \mathbf{C_c}. \tag{5.8}$$

We further assume that the measurement noise n_i's are i.i.d., and the quantization noise q_i's and wireless channel noise c_i's are independent across all sensors. The MSE incurred by $\hat{\theta}$ can be immediately computed as

$$E\left[\left|\hat{\theta} - \theta\right|^2\right] = \left(\mathbf{1}^{\mathrm{T}}\mathbf{C}^{-1}\mathbf{1}\right)^{-1} = \left(\sum_{i=1}^{N} \frac{1}{\operatorname{var}(n_i) + \operatorname{var}(q_i) + \operatorname{var}(c_i)}\right)^{-1}. \tag{5.9}$$

5.2 Variance of Distortion in Binary Symmetric Channel (BSC)

We assume that measurement value x_i's are uniform distributed within [0, 1] in all sensors. Then the quantization bit $a_k^{(i)}$ is equal prior probability at 1 or 0. Some useful relations between $a_k^{(i)}$ and $\hat{a}_k^{(i)}$ can be obtained:

$$E\left[\left|\hat{a}_k^{(i)} - a_k^{(i)}\right|\right] = \varepsilon_i, \tag{5.10}$$

$$E\left[\left|\hat{a}_k^{(i)} - a_k^{(i)}\right|^2\right] = \varepsilon_i. \tag{5.11}$$

The mean of the wireless channel error c_i can be derived as

$$E\left[c_i\right] = E\left[y_i - m_i\right] = \sum_{k=1}^{b_i} E\left[\hat{a}_k^{(i)} - a_k^{(i)}\right] 2^{-k}. \tag{5.12}$$

Because the quantization bit $a_k^{(i)}$ is equal prior probability at 1 or 0, we can derive

$$E\left[\hat{a}_k^{(i)} - a_k^{(i)}\right] = 0. \tag{5.13}$$

By (5.12) and (5.13), the wireless channel error c_i is zero mean. The upper bound of the wireless channel error variance can be derived as

$$\begin{aligned} E\left[{c_i}^2\right] &= E\left[\left|y_i - m_i\right|^2\right] = E\left[\left|\sum_{k=1}^{b_i}\left(\hat{a}_k^{(i)} - a_k^{(i)}\right)2^{-k}\right|^2\right] \le E\left[\left(\sum_{k=1}^{b_i}\left|\hat{a}_k^{(i)} - a_k^{(i)}\right|2^{-k}\right)^2\right] \\ &= E\left[\sum_{k=1}^{b_i}\left|\hat{a}_k^{(i)} - a_k^{(i)}\right|^2 2^{-2k} + \sum_{k=1}^{b_i}\sum_{\substack{h=1 \\ h\neq k}}^{b_i}\left|\hat{a}_k^{(i)} - a_k^{(i)}\right|\left|\hat{a}_h^{(i)} - a_h^{(i)}\right|2^{-k}2^{-h}\right]. \end{aligned} \tag{5.14}$$

By (5.10), (5.11), (5.14) and the following Lemma with proof given in Appendix E, the final formulation of the upper bound is

$$\begin{aligned} E\left[{c_i}^2\right] &\le \varepsilon_i \sum_{k=1}^{b_i} 2^{-2k} + \varepsilon_i^2 \sum_{k=1}^{b_i}\sum_{\substack{h=1 \\ h\neq k}}^{b_i} 2^{-k}2^{-h} \\ &= \left(\frac{\varepsilon_i + 2\varepsilon_i^2}{3}\right) - 2\varepsilon_i^2 \cdot 2^{-b_i} + \left(\frac{4\varepsilon_i^2 - \varepsilon_i}{3}\right)\cdot 4^{-b_i} \\ &\le \left(\sqrt{\frac{\varepsilon_i + 2\varepsilon_i^2}{3}} - \sqrt{\frac{3\varepsilon_i^3}{2\varepsilon_i + 1}}\cdot 2^{-b_i}\right)^2. \end{aligned} \tag{5.15}$$

Lemma 5.1 : If $0 \le \varepsilon_i \le 0.5$, we can obtain the following inequality:

$$\left(\frac{\varepsilon_i + 2\varepsilon_i^2}{3}\right) - 2\varepsilon_i^2 \cdot 2^{-b_i} + \left(\frac{4\varepsilon_i^2 - \varepsilon_i}{3}\right)\cdot 4^{-b_i} \le \left(\sqrt{\frac{\varepsilon_i + 2\varepsilon_i^2}{3}} - \sqrt{\frac{3\varepsilon_i^3}{2\varepsilon_i + 1}}\cdot 2^{-b_i}\right)^2. \tag{5.16}$$

5.3 Average Bit Error Rate (BER) in BSC

Mode over Rayleigh Fading Channel with Path Loss

We consider that the links between sensors and the FC are Rayleigh fading channels with path loss. The average bit error probability for BPSK in Rayleigh fading channel is [20]

$$\varepsilon = \frac{1}{2}\left[1 - \sqrt{\frac{r_b}{1+r_b}}\right], \tag{5.17}$$

where r_b is the average SNR per bit which is defined as

$$r_b = \frac{P_r}{N_{total}}, \tag{5.18}$$

where P_r is the received power and N_{total} is the power of the noise introduced by the receiver front-end.

With the effect of path loss, the received power can be expressed as [20]

$$P_r = P_t \cdot G \cdot d^{-\kappa}, \tag{5.19}$$

where P_t is the transmission power, d is the distance between a sensor and the FC, and G is the gain factor at $d=1$(m). The energy per bit is defined as

$$w = P_t \cdot T_b, \tag{5.20}$$

where T_b is the bit duration.

Considering the individual sensor and with (5.17), (5.18), (5.19), and (5.20), we have

$$\varepsilon_i = \frac{1}{2}\left[1 - \sqrt{\frac{w_i}{G_0 d_i^{\kappa} + w_i}}\right]. \tag{5.21}$$

where ε_i is the crossover probability of the BSC between the ith sensor and the FC (because the quantization bits at sensors are all equal prior probability at 1 or 0), d_i is the distance between the ith sensor and the FC, w_i is the transmission energy per bit in the ith sensor, and $G_0 = \left(T_b \cdot N_{total}\right)/G$ is a constant depending on the noise profile and path loss gain factor.

5.4 Average Mean Square Error of Decentralized Estimation

We assume that the consumed energy for transmitting one bit at each sensor is the same. Then the total consumed energy for transmitting the message m_i at the ith sensor is proportional to number of bits b_i. That is

$$E_i = wb_i \quad \text{for } 1 \le i \le N. \tag{5.22}$$

With (5.22), the specification of the energy allocated to the ith sensor thus amounts to determining the number of quantization bits b_i. For a fixed set of measurement noise variances σ_i 's and distances d_i's between sensors and the FC, the energy minimization problem subject to an allowable parameter distortion level γ (in term of MSE) can be formulated as

$$\begin{aligned}
&\text{Min } \sum_{i=1}^{N} b_i, \\
&\text{subject to } \left(\sum_{i=1}^{N} \frac{1}{\alpha z_i + \beta_i + f\left(\varepsilon_i, b_i\right)} \right)^{-1} \le \gamma,\ b_i \ge 0, \text{ and } 1 \le i \le N,
\end{aligned} \tag{5.23}$$

where $\beta_i = \delta + \dfrac{4^{-b_i}}{12}$ and $f\left(\varepsilon_i, b_i\right) = \left(\sqrt{\dfrac{\varepsilon_i + 2\varepsilon_i^2}{3}} - \sqrt{\dfrac{3\varepsilon_i^3}{2\varepsilon_i + 1}} \cdot 2^{-b_i} \right)^2$, or equivalently,

$$\begin{aligned}
&\text{Min } \sum_{i=1}^{N} b_i, \\
&\text{subject to } \sum_{i=1}^{N} \frac{1}{\alpha z_i + \beta_i + f\left(\varepsilon_i, b_i\right)} \ge \gamma^{-1},\ b_i \ge 0, \text{ and } 1 \le i \le N.
\end{aligned} \tag{5.24}$$

We will consider the following optimization problem, in which the equivalent MSE performance metric in (5.24) is instead averaged with respect to the noise variance statistic characterized in (5.2):

$$\begin{aligned}&\text{Min} \sum_{i=1}^{N} b_i,\\&\text{subject to} \int_{\mathbf{z}} \sum_{i=1}^{N} \frac{1}{\alpha z_i + \beta_i + f(\varepsilon_i, b_i)} p_{\chi_1^2}(\mathbf{z})\, d\mathbf{z} \ge \gamma^{-1},\ b_i \ge 0,\ \text{and } 1 \le i \le N,\end{aligned} \tag{5.25}$$

$$p_{\chi_1^2}(z) = \begin{cases} \dfrac{1}{\sqrt{2\pi z}} \exp\left(\dfrac{-z}{2}\right), & z \ge 0, \\ 0, & z < 0, \end{cases} \tag{5.26}$$

where $\mathbf{z} = [z_1, z_2, ..., z_N]$ with $p_{\chi_1^2}(\mathbf{z})$ denoting the associated distribution. In (5.25), the constraint that all b_i are nonnegative integers are relaxed to be $b_i \ge 0$ so as to render the problem tractable. The suboptimal b_i's can be obtained through upper integer rounding. The solution to the problem (5.25) is discussed next.

To solve (5.25), a crucial step is to derive an analytic expression of the average MSE performance measure. We have

$$\begin{aligned}&\int_{\mathbf{z}} \sum_{i=1}^{N} \frac{1}{\alpha z_i + \beta_i + f(\varepsilon_i, b_i)} p(\mathbf{z}) d\mathbf{z}\\&= \sum_{i=1}^{N} \int_0^{\infty} \frac{1}{\alpha z_i + \beta_i + f(\varepsilon_i, b_i)} \cdot \frac{e^{-z_i/2}}{\sqrt{2\pi z_i}} dz_i\\&= \frac{1}{\sqrt{2\pi}} \sum_{i=1}^{N} \int_0^{\infty} \frac{e^{-z_i/2}}{\left(\alpha z_i + \beta_i + f(\varepsilon_i, b_i)\right)\sqrt{z_i}} dz_i.\end{aligned} \tag{5.27}$$

The following lemma, with proof given in Appendix F, provides a closed-form expression for (5.27).

Lemma 5.2 : With $\alpha > 0$ and $x_i > 0$, we have

$$\int_0^{\infty} \frac{e^{-z_i/2}}{(\alpha z_i + x_i)\sqrt{z_i}} dz_i = \frac{2\pi \cdot e^{x_i/2\alpha} \cdot \mathrm{Q}(\sqrt{x_i/\alpha})}{\sqrt{\alpha x_i}}. \tag{5.28}$$

where $\mathrm{Q}(x) = \int_x^{\infty} \frac{e^{-t^2/2}}{\sqrt{2\pi}} dt$ is the Gaussian tail function.

By lemma 5.2 and change of the variable $x_i = \beta_i + f\left(\varepsilon_i, b_i\right)$, we have

$$\int_0^\infty \frac{e^{-z_i/2}}{(\alpha z_i + \beta_i + f\left(\varepsilon_i, b_i\right))\sqrt{z_i}} dz_i = \frac{2\pi \cdot e^{\left(\beta_i + f\left(\varepsilon_i, b_i\right)\right)/2\alpha} \cdot \mathrm{Q}(\sqrt{\left(\beta_i + f\left(\varepsilon_i, b_i\right)\right)/\alpha})}{\sqrt{\alpha\left(\beta_i + f\left(\varepsilon_i, b_i\right)\right)}}. \tag{5.29}$$

With (5.27) and (5.29), the optimization problem (5.25) can be equivalently rewritten as

$$\begin{aligned} &\text{Min } \sum_{i=1}^{N} b_i, \\ &\text{subject to } \sqrt{\frac{2\pi}{\alpha}} \sum_{i=1}^{N} \frac{e^{\left(\beta_i + f\left(\varepsilon_i, b_i\right)\right)/2\beta}\, \mathrm{Q}\left(\left(\beta_i + f\left(\varepsilon_i, b_i\right)\right)/\alpha\right)}{\sqrt{\beta_i + f\left(\varepsilon_i, b_i\right)}} \geq \gamma^{-1}, \\ &\qquad\qquad b_i \geq 0, \text{ and } 1 \leq i \leq N. \end{aligned} \tag{5.30}$$

Exact solutions to problem (5.30) appear intractable since the design constraint is highly nonlinear in b_i. We will thus seek for suboptimal alternatives which can otherwise admit simple analytic expressions. The underlying approach toward this end is to derive an easy-to-tackle lower bound on the target MSE metric. Then we replace the MSE constraint in (5.30) by one which forces the lower bound to be above γ^{-1}. Such a procedure will considerably simplify the analysis without incurring any loss in the desired MSE performance. This is done with the aid of the next lemma with proof given in Appendix G.

Lemma 5.3 : The following inequality holds:

$$\begin{aligned} &\sqrt{\frac{2\pi}{\alpha}} \sum_{i=1}^{N} \frac{e^{\left(\beta_i + f\left(\varepsilon_i, b_i\right)\right)/2\beta}\, \mathrm{Q}\left(\left(\beta_i + f\left(\varepsilon_i, b_i\right)\right)/\alpha\right)}{\sqrt{\beta_i + f\left(\varepsilon_i, b_i\right)}} \\ &\geq d \cdot N \cdot \mathrm{Q}\left(\sqrt{\frac{\delta}{\alpha}} + \frac{1}{N} \sum_{i=1}^{N} \left(\sqrt{\frac{4^{-b_i}}{12\alpha}} + \sqrt{\frac{f\left(\varepsilon_i + b_i\right)}{\alpha}}\right)\right). \end{aligned} \tag{5.31}$$

where $d=\sqrt{\frac{2\pi}{\alpha}}\frac{\exp\left(\frac{\delta+4^{-b_{\max}}/12}{2\alpha}\right)}{\sqrt{\delta+1/12+f\left(\varepsilon_{\max},b_{\max}\right)}}$, $\varepsilon_{\max}$ is the maximum crossover probability of all links between sensors and the FC, and $b_{\max}$ is the allowable maximum bits length of all sensors.

Lemma 5.3 suggests that we can replace the MSE constraint in (5.30) by the following one without incurring any loss in the target MSE:

$$d\cdot N\cdot \mathrm{Q}\left(\sqrt{\frac{\delta}{\alpha}}+\frac{1}{N}\sum_{i=1}^{N}\left(\sqrt{\frac{4^{-b_i}}{12\alpha}}+\sqrt{\frac{f\left(\varepsilon_i+b_i\right)}{\alpha}}\right)\right)\geq\gamma^{-1}, \tag{5.32}$$

or equivalently,

$$\frac{1}{N\sqrt{\alpha}}\left(\sum_{i=1}^{N}\frac{2^{-b_i}}{\sqrt{12}}+\sum_{i=1}^{N}\sqrt{f\left(\varepsilon_i,b_i\right)}\right)\leq \mathrm{Q}^{-1}\left(\frac{1}{dN\gamma}\right)-\sqrt{\frac{\delta}{\alpha}}, \tag{5.33}$$

since $\mathrm{Q}(\bullet)$ is one-to-one and monotone decreasing.

We will thus instead focus on the optimization problem with a modified MSE performance constraint:

$$\begin{aligned}&\mathrm{Min}\ \sum_{i=1}^{N}b_i,\\&\text{subject to }\ \frac{1}{\sqrt{12}}\sum_{\mathrm{i}=1}^{\mathrm{N}}2^{-\mathrm{b_i}}+\sum_{\mathrm{i}=1}^{\mathrm{N}}\sqrt{f\left(\varepsilon_i,b_i\right)}\leq \mathrm{N}\sqrt{\alpha}\left(\mathrm{Q}^{-1}\left(\frac{1}{dN\gamma}\right)-\sqrt{\frac{\delta}{\alpha}}\right),\\&\qquad b_i\geq 0,\ \text{and } 1\leq i\leq N,\end{aligned} \tag{5.34}$$

or equivalently,

$$\begin{aligned}&\mathrm{Min}\ \sum_{i=1}^{N}b_i,\\&\text{subject to }\ \frac{1}{\sqrt{12}}\sum_{i=1}^{N}2^{-b_i}-\sum_{i=1}^{N}\sqrt{\frac{3\varepsilon_i^3}{2\varepsilon_i+1}}2^{-b_i}\leq NA,\\&\qquad b_i\geq 0,\ \text{and } 1\leq i\leq N,\end{aligned} \tag{5.35}$$

where $A=\sqrt{\alpha}\left(\mathrm{Q}^{-1}\left(\frac{1}{dN\gamma}\right)-\sqrt{\frac{\delta}{\alpha}}\right)-\frac{1}{N}\sum_{i=1}^{N}\sqrt{\frac{\varepsilon_i+2\varepsilon_i^2}{3}}$. The main advantage of the alternative design formulation in (5.35) is that the cost function is linear and the constraints are convex. It is thus a convex optimization problem and will lead to a simple closed-form solution as shown below.

5.5 Problem Formulation and Suboptimal Closed-form Solution

The finial optimization problem as follows:

$$\begin{aligned}&\text{Min}\sum_{i=1}^{N} b_i,\\ &\text{subject to } \frac{1}{\sqrt{12}}\sum_{i=1}^{N}2^{-b_i}-\sum_{i=1}^{N}\sqrt{\frac{3\varepsilon_i^3}{2\varepsilon_i+1}}2^{-b_i}\le NA,\ b_i\ge 0,\ \text{and } 1\le i\le N.\end{aligned} \tag{5.36}$$

To solve problem (5.36), let us form the Lagrangian as:

$$\begin{aligned}&L\left(b_1,\ldots,b_N,\lambda,\mu_1,\ldots,\mu_N\right)\\ &=\sum_{i=1}^{N}b_i+\lambda\left(\frac{1}{\sqrt{12}}\sum_{i=1}^{N}2^{-b_i}-\sum_{i=1}^{N}\sqrt{\frac{3\varepsilon_i^3}{2\varepsilon_i+1}}2^{-b_i}-NA\right)-\sum_{i=1}^{N}\mu_i b_i.\end{aligned} \tag{5.37}$$

The associated set of KKT conditions [14] is followed:

$$1+\lambda\left(\frac{1}{\sqrt{12}}-\sqrt{\frac{3\varepsilon_i^3}{2\varepsilon_i+1}}\right)(-\ln 2)2^{-b_i}-\mu_i=0,\ 1\le i\le N, \tag{5.38}$$

$$\lambda\left(\frac{1}{\sqrt{12}}\sum_{i=1}^{N}2^{-b_i}-\sum_{i=1}^{N}\sqrt{\frac{3\varepsilon_i^3}{2\varepsilon_i+1}}2^{-b_i}-NA\right)=0, \tag{5.39}$$

$$\lambda\ge 0,\ \mu_i\ge 0,\ b_i\ge 0,\ \mu_i b_i=0,\ 1\le i\le N. \tag{5.40}$$

We first observe that, if $\lambda=0$, equation (5.38) implies $\mu_i=1>0$ for all $1\le i\le N$, and hence $b_i=0$ for all $1\le i\le N$. This case should be precluded since all sensors will remain silent. Accordingly, we must have $\lambda>0$ which means that the

MSE constraint in (5.35) is active so that

$$\frac{R}{\sqrt{12}}\sum_{i=1}^{N}2^{b_i}-\sum_{i=1}^{N}\sqrt{\frac{3\varepsilon_i^3}{2\varepsilon_i+1}}2^{-b_i}=NA. \tag{5.41}$$

Solving (5.38) and (5.41) leads to

$$b_i=\log_2\left[\frac{\left(1/12-\sqrt{3\varepsilon_i^3/(2\varepsilon_i+1)}\right)\cdot\ln 2}{1-\mu_i}\cdot\lambda\right], \tag{5.42}$$

where

$$\lambda=\frac{N-\sum_{i=1}^{N}\mu_i}{N\cdot A\cdot\ln 2}. \tag{5.43}$$

The b_i can be obtained finally:

$$b_i=\log_2\left[\frac{\left(1/12-\sqrt{3\varepsilon_i^3/(2\varepsilon_i+1)}\right)}{A}\cdot\frac{N-\sum_{i=1}^{N}\mu_i}{N-N\mu_i}\right]. \tag{5.44}$$

By taking into account the constraint $b_i\geq 0$, the suboptimal bit length b_i^{subopt} is given by the next Lemma with proof given in Appendix H.

Lemma 5.4 : Assume $\varepsilon_1\leq\varepsilon_2\leq....\leq\varepsilon_N$ without loss of generality, and define the function:

$$Y(i)=\frac{\left(1/12-\sqrt{3\varepsilon_i^3/(2\varepsilon_i+1)}\right)}{A},\quad 1\leq i\leq N. \tag{5.45}$$

Find the maximum *K1* such that $Y(K1)\geq 1$. Then we define the function:

$$Z(i)=Y(i)\cdot\frac{K1}{N},\quad 1\leq i\leq N. \tag{5.46}$$

Find the maximum *K2* such that $Z(K2)\geq 1$. Then we have

$$b_i^{subopt} = \begin{cases} \log 2\left[Z(i)\right], & 1 \le i \le K2. \\ 0\ , & K2+1 \le i \le N. \end{cases} \tag{5.47}$$

5.6 Discussions of Suboptimal Solution

1. We note that the target distortion level γ cannot be set unlimitedly small. It is lower bounded by the MSE attained by the benchmark estimate based on un-quantized real-valued sensor measurements. It is the case for $b_i = \infty,\ 1 \le i \le N$. By setting $b_i = \infty$ in the average MSE formula specified in (5.32), the minimal allowable γ can be immediately determined as

$$\gamma \ge \left[d \cdot N \cdot \mathrm{Q}\left(\sqrt{\frac{\delta}{\alpha}} + \frac{1}{\sqrt{\alpha}} \sum_{i=1}^{N} \sqrt{\frac{\varepsilon_i + 2\varepsilon_i^2}{3}} \right) \right]^{-1}. \tag{5.48}$$

2. If this bit-loading optimization problem has the proposed suboptimal solution, there must be some index i leading to $b_i \ge 0$. From (5.42), the constraint $b_i \ge 0$ also implies

$$\lambda^{-1}\left(1-\mu_i\right) \le \ln 2 \cdot \left(\frac{1}{12} - \sqrt{\frac{3\varepsilon_i^3}{2\varepsilon_i + 1}} \right). \tag{5.49}$$

From (5.43), we have

$$\lambda^{-1}\left(N - \sum_{i=1}^{N} \mu_i \right) = N \cdot A \cdot \ln 2. \tag{5.50}$$

By (5.49) and (5.50), we can derive

$$N \cdot A \le \frac{N}{12} - \sum_{i=1}^{N} \sqrt{\frac{3\varepsilon_i^3}{2\varepsilon_i + 1}}, \tag{5.51}$$

or equivalently

$$\gamma \le \left[d \cdot N \cdot \mathrm{Q}\left(\frac{e}{\sqrt{\alpha}} + \sqrt{\frac{\delta}{\alpha}} \right) \right]^{-1}. \tag{5.52}$$

where $e=\frac{1}{12}+\frac{1}{N}\sum_{i=1}^{N}\left(\sqrt{\frac{\varepsilon_i+2\varepsilon_i^2}{3}}-\sqrt{\frac{3\varepsilon_i^3}{2\varepsilon_i+1}}\right)$. It is the upper bound of the designed average MSE.

3. Recall from (5.21) that the bit error rate ε_i is proportional to the path gain d_i^{κ} (if the same transmission energy is assumed throughout all sensors). Large values of ε_i correspond to sensors deployed far away from the FC. They are usually with poor background channel gain. By this point the proposed suboptimal solution is intuitively attractive. The sensors associated with the $(N-K2-1)$th largest ε_i's are turned off to conserve energy. We note that a similar energy conservation strategy via shutting off sensors alone poor channel links is also found in [19], in which the scenario with instantaneous noise variances available to the FC is considered.
4. From (5.47), we further note that the assigned message length is inversely proportional to ε_i for those active nodes. This is intuitively reasonable since sensors with better link conditions should be allocated with more bits to realize the desired MSE performance.
5. Based on the inequality constraint for average MSE in (5.35), the equal-bit scheme maintaining the desired MSE can be obtained by solving

$$\left(\frac{N}{\sqrt{12}}-\sum_{i=1}^{N}\sqrt{\frac{3\varepsilon_i^3}{2\varepsilon_i+1}}\right)2^{-\tilde{b}}=N\cdot A, \tag{5.53}$$

leading to

$$\tilde{b}=\log_2\left\{\left(\frac{N}{\sqrt{12}}-\sum_{i=1}^{N}\sqrt{\frac{3\varepsilon_i^3}{2\varepsilon_i+1}}\right)\Big/\left(N\cdot A\right)\right\}. \tag{5.54}$$

Simulation results in the next section show that the proposed suboptimal scheme (5.47) yields energy saving when compared with (5.54).

5.7 Numerical Simulation

For a fixed set of distances between sensors and the FC (d_i, $1 \le i \le N$), the performance is measured via the percentage of energy saving (PES) [6, 7]:

$$PES = \frac{\tilde{b}\sum_{i=1}^{N} w_i - \sum_{i=1}^{N} w_i b_i^{subopt}}{\tilde{b}\sum_{i=1}^{N} w_i} \times 100, \tag{5.55}$$

where b_i^{subopt} and $\tilde{b}$ are defined respectively in (5.47) and (5.54). We assume that the transmission energy per bit is 1(mW) throughout all sensors. A fixed set of distances d_i for $1 \le i \le N$ induces a fixed set of bit error rate ε_i for $1 \le i \le N$ by (5.21) with $\kappa = 3.5$ and $G_0 = 3.4489 \times 10^{-5}$ [20]. We simply set $d_i = 5 + 20Z_i$ (m) with $Z_i \sim \chi_1^2$ being i.i.d. Chi-Square distributed random variable. The results are averaged over 50000 independent trials. The total number of sensors is *N=150* under $\gamma = 0.02$.

Figure 5.2(a) shows the PES by fixing $\delta = 0.85$ for $0.7 \le \alpha \le 1.4$ and Figure 5.2(b) depicts the average active sensors. It shows that energy efficiency of the suboptimal solution improves as α increases (a large α corresponds to a more inhomogeneous sensing environment). We note that a similar phenomenon has been observed in the existing works [6, 7] relying on instantaneous noise variance knowledge and considering the perfect wireless channel. When the sensing condition becomes more inhomogeneous, it is more likely that a large fraction of sensors suffers from poor measurement quality and will be shut off. It leads to improved energy efficiency. Since the proposed solution (5.47) based on statistical noise variance description would reflect the long-term characteristic of the schemes [6, 7], this consistency is expected.

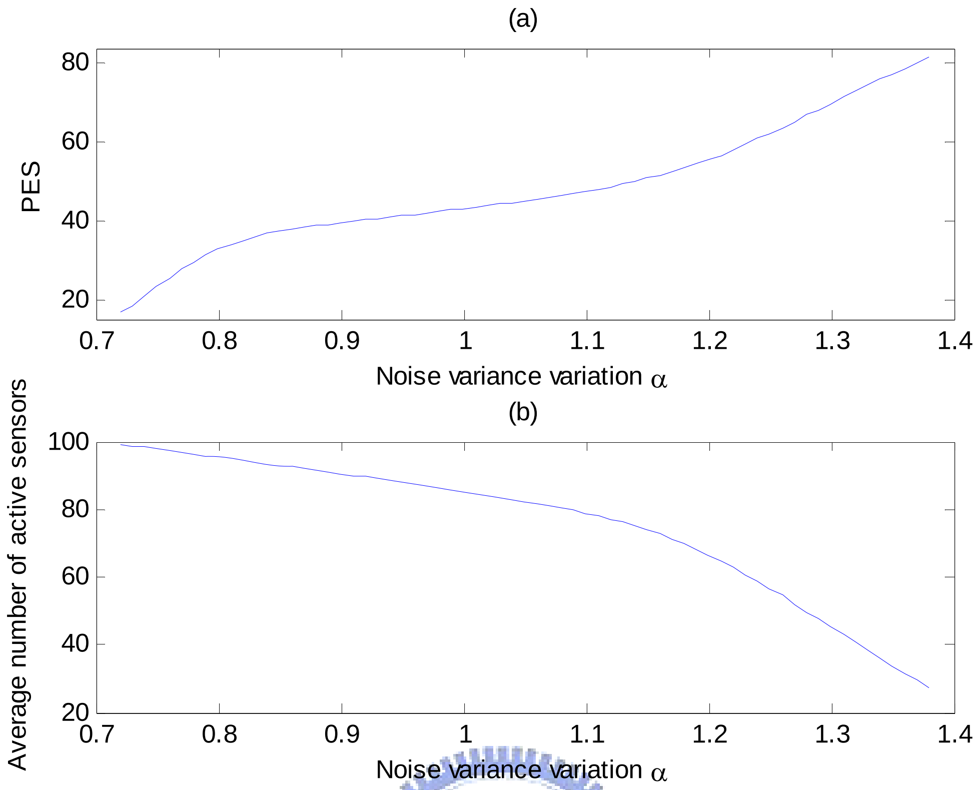

Figure 5.2 : PES for fixed minimal noise variance threshold ($\delta = 0.85$)

We repeat the experiment by fixing $\alpha = 1.45$ and varying the minimal threshold δ. The results are shown in Figure 5.3. Obviously, the PES exhibits a counter tendency as compared to Figure 5.2. It shows that the energy saving achieved by proposed suboptimal solution is lower as δ increases. This is reasonable because the large minimal noise variance threshold results in severe noise corruption in all sensor measurement. More sensor nodes should be turned on to provide a sufficient amount of information for MSE reduction.

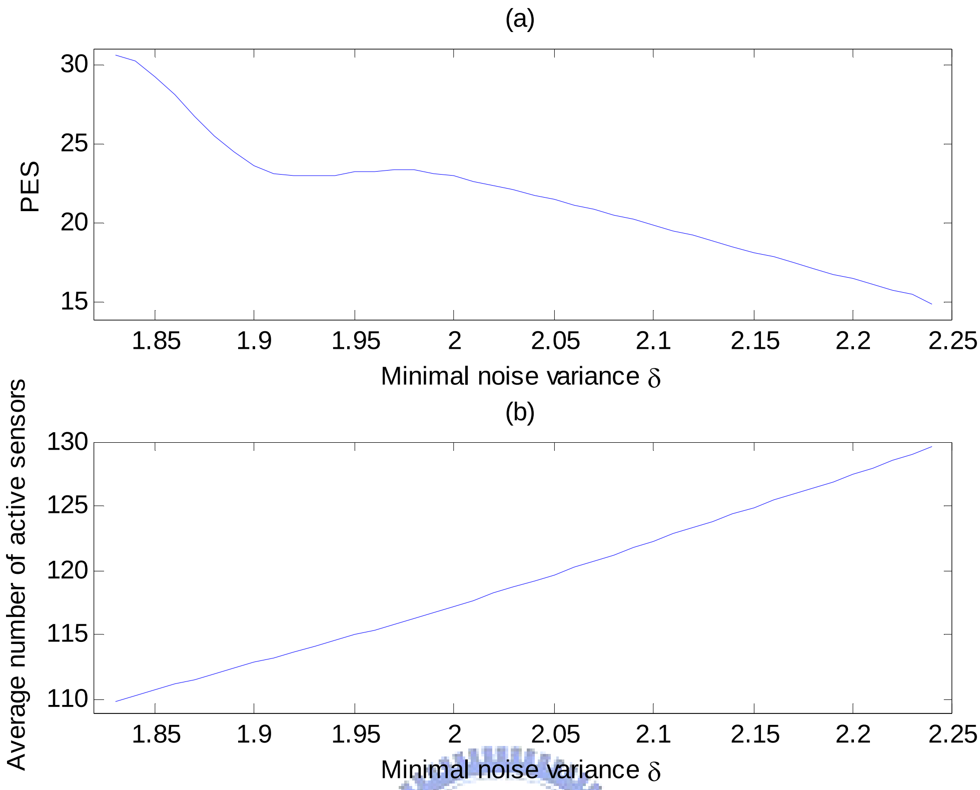

Figure 5.3 : PES for fixed noise variance variation ($\alpha = 1.45$)

5.8 Summary

This chapter provides a solution to the minimal-energy decentralized estimation problem by exploiting a statistical noise variance model and considering the nonperfect wireless channel between sensors and the FC. The wireless channel is Rayleigh fading channel with path loss. We use BSC model to characterize this wireless channel. Based on a closed-form expression of the MSE performance measure averaged over the noise variance distribution, energy minimization is reformulated as convex optimization problem. The analytic closed-form solution reveals the energy saving policy. The proposed solution simply allocates energies to sensors with large channel gain and shut off those suffering from poor link quality. We compare the proposed suboptimal solution with the uniform-allocation scheme.

Numerical simulation shows that the energy saving efficiency is particularly significant when the minimal measurement noise variance threshold is small or the variation factor is large.

Chapter 6

Conclusion

In this thesis, we consider the wireless sensor network (WSN) which is used for environmental monitoring. A popular WSN architecture consists of a fusion center and a large number of spatially distributed sensors. Each sensor in a WSN is responsible for local data collection and occasional transmission of a summary of its observations to the FC via a wireless link. In a practical WSN, each sensor has only limited computation and communication capability due to various design consideration such as small size battery, bandwidth and cost. As a result, it is difficult for sensor to send their entire real-valued observation to the FC. Instead, a more practical decentralized estimation scheme is to let each sensor quantize its real-value local measurement to an appropriate length and send the resulting discrete message to the FC. The FC combines all the received messages to produce a final estimate of the unknown parameter. Naturally, the message lengths are determined by the power and bandwidth limitation, sensor noise characteristics, wireless channel conditions, and the desired final estimation accuracy.

As energy efficiency is a critical concern for sensor network design [6, 7, 8], the decentralized estimation is formulated as optimal bit-loading problem. In the practical system the probability density function (pdf) of the observation noise is hard to characterize, especially for a large scale sensor network. The signal processing algorithms that do not require knowledge of the sensor noise pdf have been proposed [7, 8].

While most of the existing related works require the knowledge of instantaneous noise variances for energy allocation, the proposed approach instead relies on long-term noise variance knowledge. In order to improve the estimation performance against the variation of sensing conditions, repeated update of the noise profile would be needed. This comes inevitably at the cost of more training overhead and extra energy consumption. If the sensing environment is harsh, the sensing noise will change quickly. The proposed signal processing algorithm which relies on an associated sensing noise variance model is needed. Especially when the sensing environment is harsher or the instantaneous noise variance is hard to know in the FC, the proposed signal processing algorithm is useful.

Chapter 3 of this thesis attempts to provide a solution to minimal-energy decentralized estimation by exploiting long-term noise variance information. A commonly used statistical model [6, 7] for noise variance is used and the estimation performance is assessed through an MSE based metric average with respect to the considered distribution. A closed-form expression of the overall MSE requirement is derived. The analysis of the energy-minimization problem is formulated in the form of convex optimization with the average MSE constraint and then the problem is analytically solved. Chapter 4 of this thesis considers the counterpart problem: how to find the optimal bit load which minimizes the average MSE distortion under a fixed total energy budget.

Another key feature common to Chapter 3 and Chapter 4 of this thesis is that we all assume error-free transmission. We consider the sensors experiencing the perfect wireless channel. There is no bit error in the wireless channels between sensors and the FC. Chapter 5 of this thesis considers the noisy channel between each sensor and the FC by modeling it as a binary symmetric channel (BSC) model with crossover probability which is controlled by the transmitted bit energy and it use the long-term noise variance knowledge to formulate the optimization problem. The BSC models are used to characterize the wireless multi-path fading channels with path loss. A closed-form expression of the overall MSE requirement is derived and the optimization problem is then analytically solved.

The proposed signal processing algorithms share several interesting aspects pertaining to those based on the instantaneous noise variance information. Sensors with bad channel quality (specified via the path distance to FC) are shut off to conserve energy, and for those active nodes the allocated energy is proportional to the individual channel gain. The simulation results show that the proposed schemes yield energy saving against the equal-bit allocation policy.

Furthermore, if we design the problem with correlated sensor measurement noise, the results may be more suited for practical systems. In general environment, the sensor measurement noises of the adjacent sensors are highly correlated. We can also consider the wireless time-varying channel between sensors and the FC. The results may be useful for mobile sensor network. However, it is not easy for us to derive the closed-form formula of the average MSE.

Appendix

Appendix A : Proof of Lemma 3.1

By change of variable $u=\alpha z_i+\beta_i$, and hence $z_i=(u-\beta_i)/\alpha$, we have

$$\int_0^\infty \frac{e^{-z_i/2}}{(\alpha z_i+\beta_i)\sqrt{z_i}}dz_i=\int_{\beta_i}^\infty \frac{e^{(-u+\beta_i)/2\alpha}}{u\sqrt{(u-\beta_i)/\alpha}}\cdot\frac{1}{\alpha}du=\frac{e^{\beta_i/2\alpha}}{\sqrt{\alpha}}\int_{\beta_i}^\infty \frac{e^{-u/2\alpha}}{u\sqrt{u-\beta_i}}du. \quad \text{(A.1)}$$

It thus suffices to check

$$\int_{\beta_i}^\infty \frac{e^{-u/2\alpha}}{u\sqrt{u-\beta_i}}du=\frac{2\pi}{\sqrt{\beta_i}}\mathrm{Q}(\sqrt{\beta_i/\alpha}). \quad \text{(A.2)}$$

Let us define $u=\beta_i\csc^2\theta$, and hence $du=-2\beta_i\csc^2\theta\cot\theta d\theta$. We then have

$$\begin{aligned}\int_{\beta_i}^\infty \frac{e^{-u/2\alpha}}{u\sqrt{u-\beta_i}}du&=\int_{\pi/2}^0 \frac{e^{-\beta_i\csc^2\theta/2\alpha}}{\beta_i^2\csc^2\theta\cdot\sqrt{\beta_i}\cot\theta}\left(-2\beta_i^2\csc^2\theta\cot\theta\right)d\theta\\ &=\frac{2}{\sqrt{\beta_i}}\int_0^{\pi/2}e^{-\beta_i/2\alpha\sin^2\theta}d\theta. \end{aligned} \quad \text{(A.3)}$$

We note that the $\mathrm{Q}(\bullet)$ function admits the following alternative expression [13, p-71]:

$$\mathrm{Q}(x)=\frac{1}{\pi}\int_0^{\pi/2}e^{-x^2/2\sin^2\theta}d\theta. \quad \text{(A.4)}$$

The assertion (A.2) follows immediately from (A.3) and (A.4).

Appendix B : Proof of Lemma 3.2

We first observe that, since $\beta_i=\delta+R^2 4^{-b_i}/12$ and $0\le b_i\le\infty$, we have $e^{\beta_i/2\alpha}\ge e^{\delta/2\alpha}$ and $\sqrt{\beta_i}\le\sqrt{\delta+R^2/12}$, leading to

$$\sqrt{\frac{2\pi}{\alpha}}\cdot\frac{e^{\beta_i/2\alpha}\mathrm{Q}\left(\sqrt{\beta_i/\alpha}\right)}{\sqrt{\beta_i}}\ge\sqrt{\frac{2\pi}{\alpha}}\cdot\frac{e^{\delta/2\alpha}}{\sqrt{\delta+R^2/12}}\cdot\mathrm{Q}\left(\sqrt{\beta_i/\alpha}\right). \quad \text{(A.5)}$$

Also, as $\sqrt{\frac{\beta_i}{\alpha}} = \sqrt{\frac{\delta + R^2 4^{-b_i}/12}{\alpha}} \le \sqrt{\frac{\delta}{\alpha}} + \sqrt{\frac{R^2 4^{-b_i}}{12\alpha}} = \sqrt{\frac{\delta}{\alpha}} + \frac{R2^{-b_i}}{\sqrt{12\alpha}}$ and $\mathrm{Q}(\bullet)$ is one-to-one and monotone decreasing, we have

$$\mathrm{Q}\left(\sqrt{\frac{\beta_i}{\alpha}}\right) \ge \mathrm{Q}\left(\sqrt{\frac{\delta}{\alpha}} + \frac{R2^{-b_i}}{\sqrt{12\alpha}}\right). \tag{A.6}$$

Inequalities (A.5) and (A.6) then imply

$$\sqrt{\frac{2\pi}{\alpha}} \sum_{i=1}^{N} \frac{e^{\beta_i/2\alpha}\,\mathrm{Q}\left(\sqrt{\beta_i/\alpha}\right)}{\sqrt{\beta_i}} \ge c\sum_{i=1}^{N} \mathrm{Q}\left(\sqrt{\frac{\delta}{\alpha}} + \frac{R2^{-b_i}}{\sqrt{12\alpha}}\right). \tag{A.7}$$

Further, since $\mathrm{Q}(t)$ is convex for $t > 0$, it follows

$$\frac{1}{N}\sum_{i=1}^{N} \mathrm{Q}\left(\sqrt{\frac{\delta}{\alpha}} + \frac{R2^{-b_i}}{\sqrt{12\alpha}}\right) \ge \mathrm{Q}\left(\frac{1}{N}\sum_{i=1}^{N}\left(\sqrt{\frac{\delta}{\alpha}} + \frac{R2^{-b_i}}{\sqrt{12\alpha}}\right)\right), \tag{A.8}$$

and hence

$$c\sum_{i=1}^{N} \mathrm{Q}\left(\sqrt{\frac{\delta}{\alpha}} + \frac{R2^{-b_i}}{\sqrt{12\alpha}}\right) \ge cN\,\mathrm{Q}\left(\frac{1}{N}\sum_{i=1}^{N}\left(\sqrt{\frac{\delta}{\alpha}} + \frac{R2^{-b_i}}{\sqrt{12\alpha}}\right)\right), \tag{A.9}$$

and the result thus follows.

Appendix C : Proof of Lemma 3.3

By substituting $\bar{\lambda}$ in (3.23) into (3.22), it is straightforward to see that the constraint $b_i \ge 0$ is equivalent to

$$\frac{R\left(\sum_{i=1}^{N} w_i - \mu_i\right)}{\sqrt{12\alpha}\,N\left(\mathrm{Q}^{-1}\left(1/(cN\gamma)\right) - \sqrt{\delta/\alpha}\right)\left(w_i - \mu_i\right)} \ge 1. \tag{A.10}$$

Hence $\mu_i \ge 0$ must be properly chosen to simultaneously meet (A.10) and the equality constraint (3.21). The equation (3.23) can be rewritten as

$$\bar{\lambda}^{-1}\sum_{i=1}^{N}\left(w_i - \mu_i\right) = \mathrm{Q}^{-1}\left(1/(cN\gamma)\right) - \sqrt{\delta/\alpha}. \tag{A.11}$$

We observe (3.22). The constraint $b_i \ge 0$ also implies

$$\bar{\lambda}^{-1}\left(w_i - \mu_i\right) \le R \Big/ \left(\sqrt{12\alpha}N\right). \tag{A.12}$$

From (A.11) and (A.12), we have $\mathrm{Q}^{-1}\left(1/cN\gamma\right) - \sqrt{\delta/\alpha} \le R\big/\sqrt{12\alpha}$ or

$$\frac{R}{\sqrt{12\alpha}\left(\mathrm{Q}^{-1}\left(1/(cN\gamma)\right) - \sqrt{\delta/\alpha}\right)} \ge 1. \tag{A.13}$$

Note that constraint (A.13) is equivalent to

$$\gamma \le \left(Ne^{\delta/2\alpha}\,\mathrm{Q}\left(\sqrt{\frac{\delta}{\alpha}} + \frac{R}{\sqrt{12\alpha}}\right)\sqrt{\frac{2\pi}{\alpha\left(\delta + R^2/12\right)}}\right)^{-1}. \tag{A.14}$$

Since this upper bound is feasible, we may without loss of generality chose γ to be within this range so that (A.13) holds. Then we must solve μ_i such that

$$\frac{\sum_{i=1}^{N}\left(w_i - \mu_i\right)}{N\cdot\left(w_i - \mu_i\right)} \ge 1. \tag{A.15}$$

If the integer K_1 exits, it is straightforward to show that $\mu_i = w_i$ for $1 \le i \le N - K_1$ and $\mu_i = 0$ for $N - K_1 + 1 \le i \le N$. Form (3.24) and (A.13), the set of μ_i fulfill (A.10). The existence of K_1 is indeed guaranteed by the construction of $f\left(K\right)$ in (3.24): $f\left(1\right) = 1/N$, $f\left(K\right)$ is monotone increasing with K, and

$$f\left(N\right) = \left(\sum_{i=1}^{N} w_i\right)\Big/\left(N\cdot w_K\right) \ge N/N = 1.$$

Appendix D : Proof of Lemma 4.1

$$f\left(K_1+1\right)=\frac{E_T\left(1+\frac{\beta}{\alpha+\delta}\right)^{-1}+\sum_{i=K_1+1}^{N} w_i}{\sqrt{w_{K_1+1}}\sum_{i=K_1+1}^{N}\sqrt{w_i}}$$

$$=\frac{E_T\left(1+\frac{\beta}{\alpha+\delta}\right)^{-1}+\sum_{i=K_1}^{N} w_i-w_{K_1+1}}{\sqrt{w_{K_1+1}}\sum_{i=K_1}^{N}\sqrt{w_i}-w_{K_1+1}}. \tag{A.16}$$

Because of $w_{K_1} \geq w_{K_1+1}$, equation (A.16) will lead to

$$f\left(K_1+1\right)\geq\frac{E_T\left(1+\frac{\beta}{\alpha+\delta}\right)^{-1}+\sum_{i=K_1}^{N} w_i-w_{K_1+1}}{\sqrt{w_{K_1}}\sum_{i=K_1}^{N}\sqrt{w_i}-w_{K_1+1}}. \tag{A.17}$$

Let us define

$$\eta=\frac{w_{K_1+1}}{\sqrt{w_{K_1}}\sum_{i=K_1}^{N}\sqrt{w_i}}, \tag{A.18}$$

and then we have $0<\eta<1$. From (A.17) and (A.18), the lower bound of $f\left(K_1+1\right)$ can be obtained:

$$f\left(K_1+1\right)\geq\frac{f\left(K_1\right)-\eta}{1-\eta}\geq\frac{1-\eta}{1-\eta}=1. \tag{A.19}$$

The result of (A.19) is from the assumption $f\left(K_1\right)\geq 1$.

Appendix E : Proof of Lemma 5.1

$$\left(\frac{\varepsilon_i+2\varepsilon_i^2}{3}\right)-2\varepsilon_i^2\cdot 2^{-b_i}+\left(\frac{4\varepsilon_i^2-\varepsilon_i}{3}\right)\cdot 4^{-b_i}$$

$$=\left(\frac{\varepsilon_i+2\varepsilon_i^2}{3}\right)-2\left(\sqrt{\frac{\varepsilon_i+2\varepsilon_i^2}{3}}\right)\left(\sqrt{\frac{3}{\varepsilon_i+2\varepsilon_i^2}}\cdot\varepsilon_i^2\right)\cdot 2^{-b_i}+\left(\frac{3\varepsilon_i^4}{\varepsilon_i+2\varepsilon_i^2}\right)\cdot 4^{-b_i}+\left(\frac{4\varepsilon_i^2-\varepsilon_i}{3}-\frac{3\varepsilon_i^4}{\varepsilon_i+2\varepsilon_i^2}\right)\cdot 4^{-b_i}$$

$$=\left(\sqrt{\frac{\varepsilon_i+2\varepsilon_i^2}{3}}-\sqrt{\frac{3\varepsilon_i^3}{2\varepsilon+1}}\cdot 2^{-b_i}\right)^2+\left(\frac{4\varepsilon_i^2-\varepsilon_i}{3}-\frac{3\varepsilon_i^4}{\varepsilon_i+2\varepsilon_i^2}\right)\cdot 4^{-b_i}. \qquad \text{(A.20)}$$

If $0<\varepsilon_i<0.5$, then the useful inequality can be derived:

$$\because(\varepsilon_i-2)^2\geq 0$$

$$\Rightarrow \varepsilon_i\left(\varepsilon_i^2-4\varepsilon_i+4\right)+(2\varepsilon_i+3)\geq 0$$

$$\Rightarrow \varepsilon_i^4-4\varepsilon_i^3+6\varepsilon_i^2+3\varepsilon_i\geq 0$$

$$\Rightarrow \left(4\varepsilon_i^2-3\right)\left(2\varepsilon_i^2+\varepsilon_i\right)\leq 9\varepsilon_i^4$$

$$\Rightarrow \frac{4\varepsilon_i^2-3}{3}\leq\frac{3\varepsilon_i^4}{2\varepsilon_i^2+\varepsilon_i}. \qquad \text{(A.21)}$$

The inequality (5.16) can be proofed by (A.20) and (A.21).

Appendix F : Proof of Lemma 5.2

By change of variable $u=\alpha z_i+x_i$, and hence $z_i=(u-x_i)/\alpha$, we have

$$\int_0^\infty\frac{e^{-z_i/2}}{(\alpha z_i+x_i)\sqrt{z_i}}dz_i=\int_{x_i}^\infty\frac{e^{(-u+x_i)/2\alpha}}{u\sqrt{(u-x_i)/\alpha}}\cdot\frac{1}{\alpha}du=\frac{e^{x_i/2\alpha}}{\sqrt{\alpha}}\int_{x_i}^\infty\frac{e^{-u/2\alpha}}{u\sqrt{u-x_i}}du. \qquad \text{(A.22)}$$

It thus suffices to check

$$\int_{x_i}^\infty\frac{e^{-u/2\alpha}}{u\sqrt{u-x_i}}du=\frac{2\pi}{\sqrt{x_i}}\mathrm{Q}(\sqrt{x_i/\alpha}). \qquad \text{(A.23)}$$

Let us define $u=x_i\csc^2\theta$, and hence $du=-2x_i\csc^2\theta\cot\theta d\theta$. We then have

$$\int_{x_i}^{\infty} \frac{e^{-u/2\alpha}}{u\sqrt{u-x_i}} du = \int_{\pi/2}^{0} \frac{e^{-x_i \csc^2\theta/2\alpha}}{x_i^2 \csc^2\theta \cdot \sqrt{x_i}\cot\theta} \left(-2x_i^2 \csc^2\theta \cot\theta\right) d\theta$$

$$= \frac{2}{\sqrt{x_i}} \int_{0}^{\pi/2} e^{-x_i/2\alpha\sin^2\theta} d\theta. \tag{A.24}$$

We note that the $\mathrm{Q}(\bullet)$ function admits the following alternative expression

$$\mathrm{Q}(x) = \frac{1}{\pi} \int_{0}^{\pi/2} e^{-x^2/2\sin^2\theta} d\theta. \tag{A.25}$$

The assertion (A.23) follows immediately from (A.24) and (A.25).

Appendix G : Proof of Lemma 5.3

By the bits length constraint $0 \le b_i \le b_{\max}$, we have

$$\delta + \frac{4^{-b_{\max}}}{12} \le \beta_i \le \delta + \frac{1}{12}, \tag{A.26}$$

$$0 \le f\left(\varepsilon_i, b_i\right) \le f\left(\varepsilon_{\max}, b_{\max}\right). \tag{A.27}$$

By (A.26) and (A.27), the following inequality holds:

$$\begin{aligned} &\sqrt{\frac{2\pi}{\alpha}} \sum_{i=1}^{N} \frac{e^{\left(\beta_i + f\left(\varepsilon_i, b_i\right)\right)/2\beta} \mathrm{Q}\left(\left(\beta_i + f\left(\varepsilon_i, b_i\right)\right)/\alpha\right)}{\sqrt{\beta_i + f\left(\varepsilon_i, b_i\right)}} \\ &\ge d \cdot \sum_{i=1}^{N} \mathrm{Q}\left(\sqrt{\frac{\beta_i + f\left(\varepsilon_i, b_i\right)}{\alpha}}\right). \end{aligned} \tag{A.28}$$

Also as $\sqrt{\frac{\beta_i + f\left(\varepsilon_i, b_i\right)}{\alpha}} = \sqrt{\frac{\delta + 4^{-b_i}/12 + f\left(\varepsilon_i, b_i\right)}{\alpha}} \le \sqrt{\frac{\delta}{\alpha}} + \sqrt{\frac{4^{-b_i}}{12\alpha}} + \sqrt{\frac{f\left(\varepsilon_i, b_i\right)}{\alpha}}$ and

$\mathrm{Q}(\bullet)$ is one-to-one and monotone decreasing, we have

$$\mathrm{Q}\left(\sqrt{\frac{\beta_i + f\left(\varepsilon_i, b_i\right)}{\alpha}}\right) \ge \mathrm{Q}\left(\sqrt{\frac{\delta}{\alpha}} + \sqrt{\frac{4^{-b_i}}{12\alpha}} + \sqrt{\frac{f\left(\varepsilon_i, b_i\right)}{\alpha}}\right). \tag{A.29}$$

Inequalities (A.28) and (A.29) then imply

$$\begin{aligned} &\sqrt{\frac{2\pi}{\alpha}} \sum_{i=1}^{N} \frac{e^{\left(\beta_i + f\left(\varepsilon_i, b_i\right)\right)/2\beta} \mathrm{Q}\left(\left(\beta_i + f\left(\varepsilon_i, b_i\right)\right)/\alpha\right)}{\sqrt{\beta_i + f\left(\varepsilon_i, b_i\right)}} \\ &\ge d \cdot \sum_{i=1}^{N} \mathrm{Q}\left(\sqrt{\frac{\delta}{\alpha}} + \sqrt{\frac{4^{-b_i}}{12\alpha}} + \sqrt{\frac{f\left(\varepsilon_i, b_i\right)}{\alpha}}\right). \end{aligned} \tag{A.30}$$

Further, since $\mathrm{Q}(t)$ is convex for $t > 0$, it follows

$$\begin{aligned}&\frac{1}{N}\sum_{i=1}^{N}\mathrm{Q}\left(\sqrt{\frac{\delta}{\alpha}}+\sqrt{\frac{4^{-b_i}}{12\alpha}}+\sqrt{\frac{f\left(\varepsilon_i,b_i\right)}{\alpha}}\right)\\ &\geq \mathrm{Q}\left(\frac{1}{N}\sum_{i=1}^{N}\left(\sqrt{\frac{\delta}{\alpha}}+\sqrt{\frac{4^{-b_i}}{12\alpha}}+\sqrt{\frac{f\left(\varepsilon_i,b_i\right)}{\alpha}}\right)\right),\end{aligned} \tag{A.31}$$

and hence

$$\begin{aligned}&d\cdot\sum_{i=1}^{N}\mathrm{Q}\left(\sqrt{\frac{\delta}{\alpha}}+\sqrt{\frac{4^{-b_i}}{12\alpha}}+\sqrt{\frac{f\left(\varepsilon_i,b_i\right)}{\alpha}}\right)\\ &\geq d\cdot N\cdot\mathrm{Q}\left(\frac{1}{N}\sum_{i=1}^{N}\left(\sqrt{\frac{\delta}{\alpha}}+\sqrt{\frac{4^{-b_i}}{12\alpha}}+\sqrt{\frac{f\left(\varepsilon_i,b_i\right)}{\alpha}}\right)\right),\end{aligned} \tag{A.32}$$

and the result thus follows.

Appendix H : Proof of Lemma 5.4

It is straightforward to see that the constraint $b_i \geq 0$ is equivalent to

$$\frac{\left(1/12-\sqrt{3\varepsilon_i^3/\left(2\varepsilon_i+1\right)}\right)}{A}\cdot\frac{N-\sum_{i=1}^{N}\mu_i}{N-N\mu_i}\geq 1. \tag{A.33}$$

If we choose γ within the range which we discuss in section III-D, the suboptimal solution exists. It means that inequality (A.33) holds for some index *i*. Because of $\varepsilon_1 \leq \varepsilon_2 \leq \ldots. \leq \varepsilon_N$, we find the first *K1*th sensors leading to

$$\frac{\left(1/12-\sqrt{3\varepsilon_i^3/\left(2\varepsilon_i+1\right)}\right)}{A}\geq 1,\quad 1\leq i\leq K1. \tag{A.34}$$

Then we set $\mu_i = 0$ for $1 \leq i \leq K1$ and $\mu_i = 1$ for $K1+1 \leq i \leq N$. By setting μ_i, the inequality (A.33) becomes

$$\frac{\left(1/12-\sqrt{3\varepsilon_i^3/\left(2\varepsilon_i+1\right)}\right)}{A}\cdot\frac{K1}{N}\geq 1. \tag{A.35}$$

Finally we find the first *K2*th sensor such that the inequality (A.35) holds and the closed-form suboptimal solution is shown in (5.47).

Bibliography

[1] Z.-Q. Luo, "Universal decentralized estimation in a bandwidth constrained sensor network," *IEEE Trans. Information Theory*, vol. 51 no. 6 pp. 2210-2219 June 2005

[2] Z.-Q. Luo, "An isotropic universal decentralized estimation scheme for a bandwidth constrained ad hoc sensor network," *IEEE J. Select. Areas in Communication*, vol. 23, no. 4, pp. 735-744, April 2005

[3] Z.-Q. Luo, and J.-J. Xiao "Universal decentralized estimation in an inhomogeneous environment," *IEEE Trans. Information Theory*, vol. 51, no. 10, pp. 3564-3575, October 2005

[4] S. Cui, A. J. Goldsmith, A. Bahai, "Energy-Constrained Modulation Optimization," *IEEE Trans. Wireless Communications*, vol. 4, no. 5, pp. 2349-2360, September 2005

[5] S. Cui, A. J. Goldsmith, A. Bahai, "Joint modulation and multiple access optimization under energy constraints," *IEEE Global Telecommunications Conference*, pp 151-155, November-December 2004

[6] J. Xiao, S. Cui, Z.-Q. Luo and A. J. Goldsmith, "Power Scheduling of Universal Decentralized Estimation in Sensor Networks," *IEEE Trans. Signal Processing*, Vol. 54, No. 2, pp. 413-422, Feb., 2006

[7] A. Krasnopeev, J. J. Xiao, Z.Q. Luo, "Minimum Energy Decentralized Estimation in a Wireless Sensor Network with Correlated Sensor Noises," *EURASIP Journal on Wireless Communications and Networking*, pp 473-482, April 2005

[8] P. Venkitasubramaniam, G. Mergen, L. Tong, and A. Swami, "Quantization for distributed estimation in large scale sensor networks," Proc. ICISIP, pp. 121-127, 2005

[9] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, Prentice-Hall PTR, 1993

[10] S. M. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory*, Prentice-Hall PTR, 1998

[11] S. J. Orfanidis, *Introduction to Signal Processing*, Prentice-Hall, Inc., 1996

[12] A.V. Oppenheim and R. W. Schafer, *Discrete-Time Signal Processing*, Prentice-Hall, Inc.,1989

[13] M.K. Simon and M.S. Alouini, *Digital Communication over Fading Channel: A Unified Approach to Performance Analysis*, John Wiley &Sons, Inc., 2000

[14] Edwin K.P. Chong and S. H. Zak, *An Introduction to Optimization*, John Wiley & Sons, Inc, 2001

[15] J. Y. Wu, Q. Z. Huang, and T. S. Lee, "Minimal Energy Decentralized Estimation Based on Sensor Noise Variance Statistic," *Proc. ICASSP 2007*, vol 2, pp II-1001~II-1004, April 2007.

[16] A. Leon-Garcia, *Probability and Random Process for Electrical Engineering*, 2nd edition, Addison-Wesley, 1994

[17] J.J. Xiao, A. Ribeiro, Z. Q. Luo, and G. B. Giannakis, "Distributed Compress-Estimation Using Wireless Sensor Network," *IEEE Signal Processing Magazine*, vol. 23, no. 4, pp. 27-41, July 2006

[18] S. Cui, J. Xiao, A. Goldsmith, Z. Q. Luo, and V. Poor, "Estimation Diversity and Energy Efficiency in Distributed Sensing," *IEEE Trans Signal Processing 2007*, to be appeared

[19] X. Luo and G. B. Giannakis, "Energy-Constrained Optimal Quantization for Wireless Sensor Networks," *2004 First Annual IEEE Communications Society Conference on Sensor and Ad Hoc Communications and Networks.*

[20] A. Goldsmith, *Wireless Communications*, Cambridge University Press 2005