

Distributed Estimation and Capacity Analysis in Wireless Sensor Networks

by

Ajib Setyo Arifin

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy

in

Engineering

in the

Graduate School of Science and Technology

of the

Keio University, Yagami

Committee in charge:

Professor Tomoaki Ohtsuki, Chair
Professor Iwao Sasase
Professor Hiroshi Shigeno
Professor Yukitoshi Sanada

March 2015

Distributed Estimation and Capacity Analysis in Wireless Sensor Networks

Copyright 2015

by

Ajib Setyo Arifin

Abstract

Distributed Estimation and Capacity Analysis in Wireless Sensor Networks

by

Ajib Setyo Arifin

Doctor of Philosophy in Engineering

Keio University, Yagami

Professor Tomoaki Ohtsuki, Chair

Wireless sensor networks (WSNs) have many potential application areas including environmental monitoring, health, security and surveillance, and robotic exploration. The goal of WSNs is often to deliver the sensing data from all sensors to a Fusion Center (FC). The received signals at FC are estimated. One of the concerns in WSNs is the power constraint, because the sensors have only small-size batteries whose replacement can be costly. In this dissertation, we focus on distributed estimation and capacity analysis in WSNs with power constraint.

In Chapter 1, we describe the background of WSNs consisting of source, sensors, and FC. We explain the fundamental theory related to the dissertation, i.e., mutual information and estimation theory.

In Chapter 2, we present survey on distributed estimation and capacity analysis in WSNs. We discuss the distributed estimation problem based on several factors including bandwidth constraint, energy constraint, Multiple Access Channel (MAC) models, and the presence/absence of Channel State Information (CSI). We review the capacity analysis following

the existing model including physical model and protocol model.

In Chapter 3, we consider distributed estimation in WSNs considering correlated data. We obtain a closed-form solution which follows water-filling strategy. We observe that when the data is more correlated, the distortion in terms of Mean Square Error (MSE) degrades. Simulation results show that the proposed method is better than equal power method in terms of MSE.

In Chapter 4, we investigate the properties of data collection in WSNs, in terms of both capacity and power allocation strategy. Based on the relationship between Mutual Information and Minimum Mean Square Error (I-MMSE), we derive the capacity of data collection in coherent and orthogonal MAC models. We obtain closed-form expressions of capacity under coherent and orthogonal MAC models. We show through simulation results that for both coherent and orthogonal MAC models, the capacity of the optimal power is larger than that of the equal power. We also show that the capacity of coherent MAC is larger than that of orthogonal MAC, particularly when the number of sensors is large and the total power is fixed.

In Chapter 5, we conclude with key points of this dissertation and future works, such as outage distortion probability

To my family

Contents

List of Figures	iv
List of Tables	vi
1 Introduction	1
1.1 Overview	1
1.2 Wireless Sensor Network (WSN)	4
1.2.1 Source	4
1.2.2 Sensor	4
1.2.3 Fusion Center	7
1.2.4 Application	7
1.2.5 Challenges	8
1.3 Fundamental of Estimation Theory	12
1.4 Fundamental of Mutual Information	14
1.5 Short Review on Distributed Estimation and Capacity Analysis in WSNs . .	17
1.6 Outline and Contributions of the Dissertation	18
2 Survey on Distributed Estimation and Capacity Analysis in WSNs	25
2.1 Distributed Inference	25
2.2 Distributed Estimation	28
2.2.1 Distributed Estimation in Bandwidth Constraint	29
2.2.2 Distributed Estimation in Coherent and Orthogonal MAC	31
2.2.3 Distributed Estimation with and without Channel State Information (CSI)	32
2.2.4 Distributed Estimation in Energy Constraint	34
2.3 Capacity Analysis	34
2.3.1 Capacity Analysis based Protocol Model	35
2.3.2 Capacity Analysis based Physical Model	40

3	Water-Filling Solution for Distributed Estimation of Correlated Data in WSN MIMO System	43
3.1	Introduction	43
3.2	Problem Formulation	45
3.2.1	Correlated Sources	48
3.2.2	Objective Function	49
3.3	Proposed Approach	50
3.3.1	Proposed Method	50
3.3.2	Equal Power Method	56
3.4	Simulation Results	56
3.5	Conclusion	62
4	Capacity of Data Collection in Wireless Sensor Networks based on Mutual Information and MMSE Estimation	63
4.1	Introduction	63
4.2	Problem Formulation	65
4.2.1	Capacity of the Gaussian Channel based I-MMSE approach	65
4.2.2	System Model	67
4.3	Capacity of Data Collection in Coherent MAC Model	68
4.3.1	Upper Bound on Capacity of Data Collection	68
4.3.2	Capacity of Data Collection for Equal Power Allocation	69
4.3.3	Capacity of Data Collection for Optimal Power Allocation	70
4.4	Capacity of Data Collection in Orthogonal MAC Model	73
4.4.1	Capacity of Data Collection for Equal Power Allocation	73
4.4.2	Capacity of Data Collection for Optimal Power Allocation	74
4.5	Simulation Results	75
4.6	Conclusion	79
5	Conclusion	82
5.1	Contributions	82
5.2	Future Work	83
	Appendix A	85
	Appendix B	87
	Bibliography	91
	Appendix C	104
C.1	Journals	104
C.2	International Conferences	104
C.3	Domestic Conferences	105

List of Figures

1.1	Sensor node diagram.	5
1.2	Position of each chapter in the existing research fields.	19
1.3	Configuration of this dissertation	21
3.1	Distributed estimation over a wireless sensor network.	45
3.2	The targets are assumed being in line with spacing distance, d	49
3.3	Distortion comparison between the proposed method, equal power method and the lower bound with parameters: $\rho = 0, L = 10, k = 8, N = 5, p = 5$. . .	57
3.4	Distortion comparison between difference correlation coefficients with parameters: $\rho = 0, 0.9, L = 10, k = 8, N = 5, p = 5$	58
3.5	Distortion versus k with parameters: $\rho = 0.7, L = 10, N = 5, p = 5, P = 5, 10, 20$ dB.	59
3.6	Distortion versus L with parameters: $\rho = 0.7, k = 8, N = 5, p = 5, P = 1, 2, 5, 8, 15, 20$ dB.	60
3.7	Distortion comparison of the proposed method with ideal sensor and ideal channel condition with parameters: $\rho = 0.5, k = 8, N = 5, p = 5$	61
4.1	System model of the capacity of data collection in WSNs.	67
4.2	Capacity of data collection for equal power method versus optimal power method in coherent MAC as P increases. Note that power P is taken relative to the channel noise power. Since we assume that the channel noise has unitary variance, thus we label the total transmit power in unit of dB.	76
4.3	Capacity of data collection for equal power method versus optimal power method in orthogonal MAC as P increases. Note that power P is taken relative to the channel noise power. Since we assume that the channel noise has unitary variance, thus we label the total transmit power in unit of dB. . .	77
4.4	Comparison between the capacity of data collection in coherent MAC and orthogonal MAC for optimal power allocation method.	78
4.5	Comparison between the capacity of data collection in coherent MAC and orthogonal MAC for equal power allocation method.	79

4.6	Capacity of data collection of equal power method versus optimal power method as L increases for either coherent and orthogonal MAC in a finite power budget, $P = 20$ dB.	80
4.7	Percentage of active sensors as P increases for optimal power method in orthogonal MAC.	81

List of Tables

1.1	Notation used in Chapter 2	22
1.2	Notation used in Chapter 3	23
1.3	Notation used in Chapter 4	24

Acknowledgments

Foremost, I own my gratitude to my advisor Prof. Tomoaki Ohtsuki for the continuous support of my study and research. His patience, enthusiasm, motivation and immense knowledge deeply impress me and will benefit me a lot in my whole life.

I would like to thank to Dr. Jihoon Hong, Fereidoun H. Panahi, and Masayoshi Ozawa for your help and nice discussions.

My sincere thanks also go to all the committee members for examining my thesis defense. I benefit tremendously from their personal insights and inspirational questions on my dissertation.

Many thanks go to Hitachi Scholarship Foundation for giving me support in my study, especially to Kazuyuki Miyanaga-san, Masaaki Kawamoto-san, Michio Sasamori-san, Tamammi Ono-san.

I am grateful to my whole family for their encouragement for my studying aboard, especially to my mother, my father, my brother : Indra, and my sister : Lestari.

Chapter 1

Introduction

One of the interesting topics in Wireless Sensor Networks (WSNs) is distributed estimation where data observations are collected by sensors and performed a final estimation at the fusion center (FC). Furthermore, to guarantee the data observations received by FC, study on the capacity analysis of data collection is important in WSNs. This dissertation is focused on modelling and performance evaluation related to distributed estimation and capacity analysis in WSNs.

1.1 Overview

Advanced technologies in electronics and wireless communications to make a small size sensor have created a new research area called WSNs. The sensor is designed for sensing, self-processing, and communication. The sensor is also designed to be low cost and low power consumption in order to deploy massively. WSNs consist of a large scale of sensors to observe phenomena in the area of interest. WSN can be used for many applications across different fields such as military, environmental monitoring, and health monitoring [1]. The sensors

cooperatively work making WSN attractive to perform advanced signal processing tasks such as distributed estimation, distributed detection, target localization and tracking [2].

The purpose of WSNs is often to deliver the sensing data from all sensors to the FC and then conduct further analysis at the FC. Thus, data collection is important in sensor network applications [3]. Theoretical measure that captures the limits of collection processing in sensor networks is the capacity of data collection. Capacity of data collection reflects how fast FC can collect sensing data from all sensors [4]. Understanding the capacity of the network is important for network designers in a feasibility of a large scale network deployment [5], particularly, to improve the performance of WSN [3].

An initial work on capacity of data collection in WSNs is investigated in [6], where a clustering architecture of WSNs is considered. An overall throughput bound per sensor node is derived. In [7], capacity of data collection with complex physical layer techniques is considered. A source/channel coding theorem is derived. The capacity that involves multiple selected sources and destination has been studied in [8]. The capacity of data collection of single and multi sinks are investigated [9]. In general, capacity of data collection is calculated based on protocol or physical models. The protocol model defines that a successful transmission occurs when a sensor falls inside the transmission range of its intended transmitter and falls outside the interference ranges of other non-intended transmitters. However, the protocol model is relatively inaccurate, when simultaneous transmissions are allowed in the network. The physical model is the Signal-to-Interference-plus-Noise Ratio (SINR) model. This model is based on practical transceiver designs of communication system which treats the interference as noise. However, capacity analysis considering multiple interference is difficult and an open problem.

Once the data is arrived at the FC, we further conduct the advance processing, such as distributed estimation. The FC estimates the target(s) observed by the sensors. In

practical scenario, if the targets or the sensors are close to each other, the data will be potentially correlated. An extensive analysis considering the correlated data is essential. Another important property of many WSNs is their stringent power constraint. In such networks, sensors have only small-size batteries whose replacement can be costly. Thus, sensor network operations must be energy efficient to maximize network lifetime.

FC in WSNs is designed to serve multiple users, i.e., sensors. Two types of Multiple Access Channel (MAC) models are considered in [10], [11], i.e., coherent and orthogonal MAC models. The coherent MAC is assumed that there is perfect synchronization between sensors and FC, so that the transmitted messages from local sensors can be coherently combined at FC. The orthogonal MAC model can be considered such as Frequency-Division Multiple-Access (FDMA). Motivation of using orthogonal MAC model is removal of the requirement of the carrier level synchronization among sensors.

Two issues addressed in this dissertation, i.e., distributed estimation and capacity of data collection, which can be studied reciprocal. Maximizing the capacity will minimize the distortion of estimation. Investigation on these are conducted under coherent and orthogonal MAC models. In coherent MAC model, assumption of perfect synchronization between sensors and the fusion center is necessary so that the transmitted messages from local sensors. With such an assumption, one key design consideration at local sensors and the fusion center is how to jointly process the sensed and received information. Moreover, the motivation for using orthogonal multiple access schemes is the removal of the requirement on the carrier level synchronization among sensors.

In the following section, a concept of WSN is explained. Two fundamental theories which are the basic knowledge of this dissertation including fundamental of estimation theory and fundamental of mutual information are discussed.

1.2 Wireless Sensor Network (WSN)

WSN is multidisciplinary area that involves, radio and networking, signal processing, artificial intelligence, database management, systems architectures for operator-friendly infrastructure administration, resource optimization, power management algorithms, and platform technology (hardware and software, such as operating systems) [12]. In general, WSN can be classified into three main parts, i.e. source, sensor, and fusion center. For further explanation is presented below.

1.2.1 Source

Source is any kind of the phenomena where WSNs observe. This phenomena can exist in physical world (e.g., wind, humidity, heat), biological system, or information technology framework [13]. This phenomena emits a certain value or magnitude that will be acquired by sensor or actuator for further processing in sensor node. The number of sources can be single or multiple. If the number of sources increases, the order of computation complexity increases in terms of signal processing at FC such as detection, estimation, or localization.

1.2.2 Sensor

Sensor is a device to sense or observe the phenomena. Sensor is equipped by technology for sensing of single or multiple objects. This sensing technology includes electric and magnetic sensors, radio wave sensors, biochemical sensors, location sensors, radars, and environmental parameter sensors [13]. Sensor can be designed for having multiple onboard sensing capabilities based on the application we want to observe. Sensor intended for WSN is able to conduct basic processing for extraction and manipulation of the received signal from the source. In general, sensor has four basic units including sensing unit, processing and storage

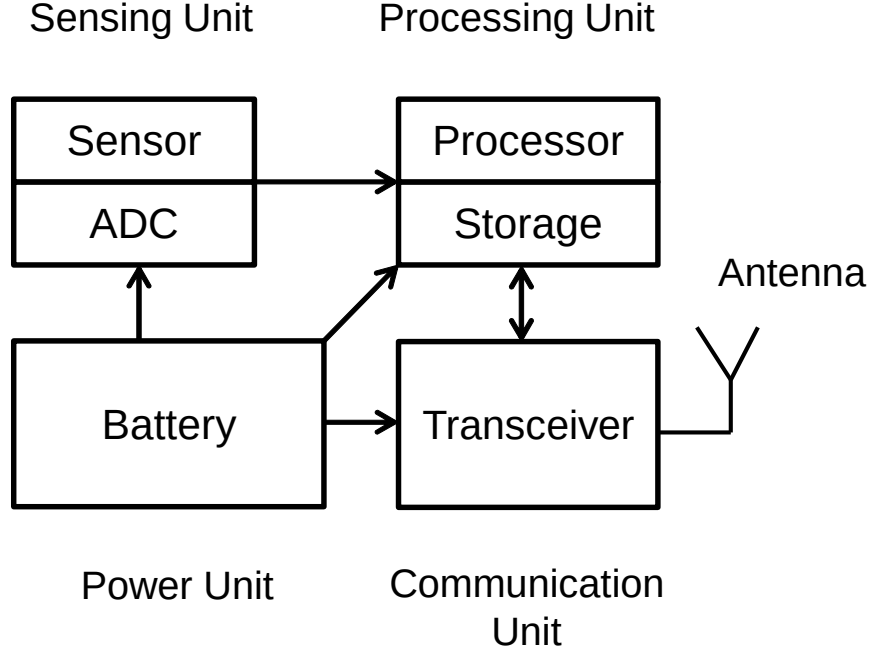


Figure 1.1: Sensor node diagram.

unit, communication unit, and power unit as depicted in Fig. 1.1. Sensing unit can be single or multiple sensors combined with Analog-to-Digital Converter (ADC). The sensed data is then processed before being fed to communication unit. In the communication unit, the data is modulated using certain carrier frequency. Communication unit is also responsible for frequency selection, carrier frequency generation, signal detection, and data encryption [1]. Sensor is designed to have single antenna, however, in certain scenarios which requires high data rate, multiple antennas are considered [14].

Since sensors will be deployed in a large scale, it should be designed at low cost. Every single sensor cost will affect the overall cost of WSN. Several considerations regarding to sensor include power consumption, operating system, protocol stack, and additional component such as location finding system, mobilizer, and actuator for special purpose of sensor node. Power consumption of sensor is important. Sensor may not be possible for changing

the battery or recharging the power. Hence, many researchers investigate of WSNs by considering power constraint. Sensor deployment has been investigated combined with power management [15]–[17]. Data collection in the presence of power constraint is proposed by maximizing the number of data gathering queries. To solve the problem, they present a generic cost model of energy consumption for data gathering queries [18]. Similar problems on data gathering also are presented in [19]–[22]. We also can see the other research topics which deal with power constraint including routing [23]–[25], data compression [26], [27], capacity [3], [28], [29], distributed detection [30]–[33], target tracking [34], control [35], MAC protocol [36].

To handle the sensor node operation, it is necessary to have Operating System (OS) embedded on the sensor node. This OS may be designed for special purpose operation depend on the intended application of WSNs. TinyOS is one such example of a standard. TinyOS's component library includes network protocols, distributed services, sensor drivers, and data acquisition tools [13]. In [37], software-defined radio transceiver is introduced. That work presents a through model of the backscatter radio link, the system architecture and a set of data extraction techniques for each sensor's information. The interesting work about OS includes Contiki [38], SOS [39], Mantis OS [40], Nano-RK [41], RETOS [42], and LiteOS [43].

Communication between sensor nodes and sink/fusion center needs a set of protocols. A generic protocol introduced in [44] includes physical layer, data link layer, network layer, transport layer, application layer, power management plane, mobility management plane, and task management plane. Physical layer consists of communication channel, sensing, actuation, and signal processing. Data link layer includes channel sharing, timing, multiplexing of data streams, data frame detection, medium access and error control. Network layer includes topology management and routing. Transport layer includes data dissemination

and accumulation, caching, and storage. Application layer consists of different application softwares. In addition, power, mobility, and task management planes monitor the power, movement, and task distribution among the sensor nodes.

1.2.3 Fusion Center

One that characterizes WSNs is the presence of FC [45]. When there is no FC, the network will be called ad hoc network. The FC collects data from sensors and produces a final decision or estimation. If the area of sensor deployment is very large and the coverage of FC is limited, in this scenario clustering can be considered. Each cluster has a cluster head which is responsible for collecting data from sensors in the cluster. The sensed data is, then, transmitted to the FC.

1.2.4 Application

WSNs' applications are spread in many areas of interest including homes, manufacture, military, and environmental. In general, these applications can be classified into indoor and outdoor. Indoor application includes, but not limited to, industrial control (asset management, process control, and energy management), manufacturing automation, home security, elderly monitoring, medical sensing and monitoring. Outdoor application includes, but not limited to, weather sensing, volcanic eruptions, vehicle tracking, traffic light sensors, toxing detection, environmental monitoring (land, air, sea), and battlefield monitoring.

In indoor application in [46], an effective indoor localization method of hybrid between Receive Signal Strength Index (RSSI) and Time Difference Of Arrival (TDOA) is proposed. The idea is taking benefits from both of them. RSSI method has lower computational complexity, but lower accuracy. While, TDOA has very good accuracy at the cost of extensive

computational complexity. The system measures RSSI of signals using a recursive filter. They also compare between polynomial fitting and log-normal shadowing. They show that Polynomial fitting is more accurate than log-normal shadowing. They also showed that the accuracy using those methods is approximate 0.5 meter.

Indoor monitoring of air quality is presented in [47]. The application is used to support peoples comfort, health, and safety. The sensor is designed for having very low current consumption when in idle position. In the context of energy efficiency, the sensors are expected to have a significant lifetime extension which is confirmed by an experimental results using 36 sensors.

In [48], they record and monitor redwood trees. Each sensor measures air temperature, relative humidity, and photosynthetically active solar radiation. Sensor nodes are deployed at different heights of the tree. These measurements are to monitor microclimate change in the area of interest.

WSNs is deployed to monitor seismic activity of oil and gas reservoir in [49]. Cabling and wiring are previously used in the conventional method, but those cause inefficiencies, large logistic, and weight costs. Wireless sensor networks are then expected to enable the future technology of seismic explorations. Some basic challenges are presented including strict sampling synchronization, high precision sensor localization, and high data rate.

1.2.5 Challenges

WSNs are often designed for specific applications. Because of different applications, WSNs may have different challenges. Each application has one or several important factors that at least should be fulfilled. For example, WSNs for patient monitoring, in here we take care of the personal information of the patients. As a result, we should concern about security

and privacy of each message delivering through WSNs. WSNs for environmental monitoring such as nuclear radiation monitoring, here we should be aware with false information. Information about the magnitude of nuclear radiation is the most important in this application, so that we should point out the confidence of the information. Designing this application should consider sensor failure, robust signal detection, and robust routing protocol. Moreover, WSNs for underwater scenario, we should consider propagation effect that may be different with general wireless communications. Because of limited propagation of radio for underwater, acoustic telemetry is other way for communications. Energy conservation is also taken into consideration because the sensors will be placed for a long period without changing the battery [50], [51]. In general, it is difficult to develop universal protocol for WSNs.

Technical challenges that always exist from the nature is noise. Noisy sensor arises from imperfect sensing and electronic limitation such as saturated gain. That factor effects the accuracy of sensor reading. Noise also exist in wireless channel between sensor and FC. Several problems deal with the channel, such as interference, link contention, unidirectional link and path loss. Sensors also suffer from some environmental factors, such as rain, snow, wide variance of temperature, corrosive substance, and high wind speed. These factors can endanger typical electronic device. These factors are necessary considering in WSNs.

WSNs can be classified into several types based on the location of deployment and their characteristics into five types [50]. Each type has different challenges. There are terrestrial WSN, underground WSN, underwater WSN, multi-media WSN, and mobile WSN. Terrestrial WSNs usually consist of hundreds of sensor nodes deployed in a given area, either structured or unstructured deployment [1]. Structured deployment means that sensors are deployed in a certain area with pre-planning. While, unstructured deployment means that sensors are deployed randomly. Sensor can be spread into areas of interest using plane or

balloon. Because the number of sensors is large, WSNs have a large data redundancy. As a consequence, we need an optimal data collection algorithm. If sensors are located far from FC, we need a routing protocol which provides the shortest path to FC. Other consequence is interference. When sensors are dense, interference can limit the performance of WSNs.

Second type is underground WSNs. The underground WSNs consist of sensors deployed inside of cliff, cave or mining sites. Sensors are deployed with structured method. However, sensors may not be managed after deployment due to harsh terrain. Sensors are like put and left. In this case, sensors must be energy efficient in terms of sensing, processing, and communication. Selective radio frequency for this condition is necessary. In [52], an experimental result for underground WSNs is presented using frequency of 300-900 MHz. Characterization includes effect of underground channel due to volumetric water content, effect of attenuation due to frequency, and high temporal stability of underground channel. Instead of using electromagnetic waves, there is an alternative signal propagation techniques, such as magnetic induction. Using electromagnetic waves often suffers some technical problems such as large attenuation and large antenna size due to low frequency usage. To address these challenges, magnetic induction is used because channel conditions of that remain constant in a soil medium [53]. Transmission and reception can be conducted using small coil of wire.

Third type is underwater WSNs. Underwater WSNs is a promising technology for seismic imaging and oil offshore related. Sensor can be deployed for monitoring and data acquisition below the sea. However, sensor device may be more expensive than terrestrial WSNs. The sensor should be chosen appropriately following the characteristic of the environment. There are three major challenges regarding underwater WSNs [54]. First is communication medium. Conventional electromagnetic waves are suffered by a large attenuation. They can transmit signal up to 100 cm for same transmit power used in terrestrial WSN. To address this

problem, an acoustic waves are proposed in [54]. The second challenge is limited bandwidth and large propagation delay. This is a consequence of using acoustic waves which only have speed of sound. The last is energy conservation. Sensors are equipped with only limited battery. Communication and processing mechanism should comply with energy efficient manner. Moreover, another way to prolong the lifetime of sensor is using energy harvesting mechanism from vibration generated by, such as ocean waves [55].

The fourth type is multi-media WSNs. This category consists of sensors equipped with cameras and microphones. As a consequence, WSNs produce large amount of data. To send this data, the system requires a high data rate. A high data rate needs a large bandwidth and increases energy for communication as well. To overcome these challenges, WSNs require intermediate processing such as compressing, filtering, and merging the data.

The last category is mobile WSNs. This WSNs characterize that sensors can move and interact with the physical phenomena. Typical sensors for this WSNs may have more advanced technology that allows the sensor to move and adapt the environmental condition. The promising advantage of this WSNs is reducing of uncovered area. However, there are still some challenges, such energy efficiency, localization, self-organization, and maintenance.

In the next section, we present two fundamental theories which are used in this dissertation. First is the estimation theory. The estimation theory explains how to estimate the information after it has been transmitted through channel(s). The received signals, then, are used to reconstruct the information. Then, the different value between the estimated information and the original information is used to measure the performance of system. Second is the mutual information. The mutual information is used to measure the dependence of two random variables. In communications, the mutual information is translated to channel capacity, which measures how much information can be transmitted over noisy channels.

1.3 Fundamental of Estimation Theory

Estimation theory comes from the problem where we want to extract information/parameters from given data/observation. An estimator can be a function that assigns a value θ for each realization x . Because the data is random, we use a statistical method to get the estimated data.

In general, estimation theory can be classified into parametric estimator and non-parametric estimator. Parameteric estimator is the estimator with the knowledge of probability density function (pdf) of the data and of the parameter being estimated. The result of this estimator depends on the goodness of these pdfs. Non-parametric estimator is the estimator without dependence on pdf of the data.

In addition, estimator, based on the prior knowledge of pdf of the parameter being estimated, can be distinguished into classical estimator and Bayesian estimator. In the classical estimator, the parameter is modeled by deterministic vector with unknown value. Given observation $\mathbf{x}$, we can write pdf parameterized by unknown vector value θ as $f(\mathbf{x}; \theta)$. In Bayesian estimator, the estimator has a prior knowledge of pdf of the parameter being estimated, $f(\theta)$. Hence, pdf of observation is a density function conditioned on the value of the parameter, $f(\mathbf{x}|\theta)$.

In particular, this thesis focuses on Bayesian estimator. The interested reader can read the excellent book [56] about classical estimator. As mentioned before, Bayesian estimator requires the prior knowledge of pdf of the parameter being estimated, $f(\theta)$. The pdf of observation conditioned by parameter θ can be written as $f(\mathbf{x}|\theta)$. Following Bayes' theorem, we can write a posterior density function as

$$f(\theta|\mathbf{x}) = \frac{f(\mathbf{x}|\theta)f(\theta)}{f(\mathbf{x})}. \quad (1.1)$$

To measure the quality of estimation, Bayesian mean squared error (MSE) is used as

$$B_{mse}(\hat{\theta}) = \int \left[\int (\theta - \hat{\theta})^2 f(\theta|\mathbf{x}) d\theta \right] f(\mathbf{x}) d\mathbf{x} \quad (1.2)$$

where $p(x) \geq 0$ for all $\mathbf{x}$. The integral in brackets can be minimized for each x , then the Bayesian MSE is minimized. We want to obtain the optimal estimator given the Bayesian MSE as

$$\begin{aligned} \frac{\delta}{\delta \hat{\theta}} \int (\theta - \hat{\theta})^2 f(\theta|\mathbf{x}) d\theta &= \int \frac{\delta}{\delta \hat{\theta}} (\theta - \hat{\theta})^2 f(\theta|\mathbf{x}) d\theta \\ &= \int -2(\theta - \hat{\theta}) f(\theta|\mathbf{x}) d\theta \\ &= -2 \int \theta f(\theta|\mathbf{x}) d\theta + 2 \int \hat{\theta} f(\theta|\mathbf{x}) d\theta \end{aligned} \quad (1.3)$$

By setting the left hand side equal to zero, we have

$$\begin{aligned} \hat{\theta} &= \int \theta f(\theta|\mathbf{x}) d\theta \\ &= E(\theta|\mathbf{x}) \end{aligned} \quad (1.4)$$

We can see that the optimal estimator is mean of posterior pdf, $f(\theta|\mathbf{x})$. Thus, the result of estimator can be substituted into eq. (1.2) as

$$\begin{aligned} B_{mse}(\hat{\theta}) &= \int \left[\int (\theta - E(\theta|\mathbf{x}))^2 f(\theta|\mathbf{x}) d\theta \right] f(\mathbf{x}) d\mathbf{x} \\ &= \int \text{var}(\theta|\mathbf{x}) f(\mathbf{x}) d\mathbf{x}. \end{aligned} \quad (1.5)$$

Again, we can see that the Bayesian MSE is variance of the posterior pdf.

1.4 Fundamental of Mutual Information

Mutual information is an amount of dependence between two random variables. This is introduced by Shannon in terms of digital communication [57]. In other words, mutual information measures the amount of uncertainty about one random variable given by knowledge of the other random variable. In addition, mutual information has a close relation with entropy, which measures the uncertainty of the random variable. Following the definition of Kullback Leibler, the distance between two random variables is given by

$$\begin{aligned} D(p||q) &= \sum_{x \in X} p(x) \log \frac{p(x)}{q(x)} \\ &= E_p \log \frac{p(X)}{q(X)}. \end{aligned} \tag{1.6}$$

We can see that the result is always non-negative. It equals zero if and only if the two random variables are independent ($p = q$). Let us consider the relationship between entropy and mutual information. Since the mutual information measures dependence between two random variables, we can write that as the relative entropy between the joint distribution and the product of each distribution.

$$\begin{aligned} I(X; Y) &= \sum_{x \in X} \sum_{y \in Y} p(x, y) \log \frac{p(x, y)}{p(x)p(y)} \\ &= D(p(x, y) || p(x)p(y)) \\ &= E_{p(x, y)} \log \frac{p(X, Y)}{p(X)p(Y)}. \end{aligned} \tag{1.7}$$

By the definition $p(x|y) = \frac{p(x,y)}{p(y)}$, we can write

$$\begin{aligned}
I(X;Y) &= \sum_{x,y} p(x,y) \log \frac{p(x|y)}{p(x)} \\
&= - \sum_{x,y} p(x,y) \log p(x) + \sum_{x,y} p(x,y) \log p(x|y) \\
&= - \sum_{x,y} p(x,y) \log p(x) - \left(- \sum_{x,y} p(x,y) \log p(x|y) \right) \\
&= H(X) - H(X|Y)
\end{aligned} \tag{1.8}$$

Mutual Information over Noisy Channel

When information is conveyed over noisy channel, mutual information can be determined as channel capacity. We can model the output channel as, $Y = X + N$, where X is input data and N is zero mean i.i.d. Gaussian with variance σ_n^2 . Channel capacity for band-limited, power-limited over Gaussian channel can be formulated,

$$C = \max_{P_x: E[X^2] \leq P} I(X;Y) \tag{1.9}$$

Because of symmetric properties of mutual information, we can write as

$$I(X;Y) = H(Y) - H(Y|X) \tag{1.10}$$

where $H(Y) = \frac{1}{2} \log(2\pi \text{var}(Y))$ and the base of the algorithm is 2, so that Entropy has a unity in bits. $H(Y|X)$ can be derived as

$$\begin{aligned}
H(Y|X) &= \int p_X(x) H(Y|X=x) dx \\
&= \int p_X(x) \left(\int p_{Y|X}(y|x) \log \frac{1}{p_{Y|X}(y|x)} dy \right) dx \\
&= \int p_X(x) \left(\int p_N(y-x) \log \frac{1}{p_N(y-x)} dy \right) dx \\
&= \int p_X(x) \left(\int p_N(n) \log \frac{1}{p_N(n)} dy \right) dx \\
&= H(N) = \frac{1}{2} \log(2\pi \sigma_n^2).
\end{aligned} \tag{1.11}$$

Thus mutual information can be written as

$$\begin{aligned}
I(X, Y) &= H(Y) - H(N) \\
&= \frac{1}{2} \log(2\pi \text{var}(Y)) - \frac{1}{2} \log(2\pi \sigma_n^2).
\end{aligned} \tag{1.12}$$

The upper bound of mutual information can be achieved by making Y being Gaussian, i.e., by taking X being Gaussian with variance P

$$\begin{aligned}
\max_{X: E[X^2] \leq P} I(X; Y) &= \frac{1}{2} \log(2\pi[P + \sigma_n^2]) - \frac{1}{2} \log(2\pi \sigma_n^2) \\
&= \frac{1}{2} \log \left(1 + \frac{P}{\sigma_n^2} \right).
\end{aligned} \tag{1.13}$$

1.5 Short Review on Distributed Estimation and Capacity Analysis in WSNs

As mentioned earlier, WSNs have many potentials of technologies across application areas including environmental monitoring, health, security and surveillance, and robotic exploration [1]. The problem is how to infer information in the area of interest being useful for the user. One of the inference problem in WSNs is the distributed estimation. In the distributed estimation, the distributed sensors observe phenomena from the target(s) and transmit to FC. Received signal at FC is estimated using an estimation technique.

Distributed estimation by considering power consumption, has attracted much attention in [14], [58], [60]–[63]. Because the sensors deployed in a certain region are difficult to change the batteries, the low power consumption is important to guarantee a lifetime of the sensors. The distributed estimation is also applied on the orthogonal MAC model [58], [62], [64] and the coherent MAC model [14], [61], that considers Single-Input Single-Output (SISO). To save energy, Multiple-Input Multiple-Output (MIMO) system has been analyzed by involving signaling overhead [65] and can offer substantial energy savings in WSN. Cooperative MIMO with data collection also has been investigated in [66], [67].

Moreover, the goal of WSNs is often to deliver the sensing data from all sensors to a FC and then conduct further analysis at the FC. Thus, data collection is important in sensor network applications [72]. Theoretical measure that captures the limits of collection processing in sensor network is the capacity of data collection. Capacity of data collection reflects how fast FC can collect sensing data from all sensors [4]. Understanding the capacity of the network is important for network designers in a feasibility of a large scale network deployment [5], particularly, to improve the performance of WSNs [72]. Furthermore, such understanding is essential in the development of efficient protocols [74].

Capacity limits of data collection in wireless sensor networks have been studied in the literature [72]–[75]. In [73] and [74], they introduced the transport capacity of many-to-one in dense sensor networks. The authors in [6], [7], investigated the capacity of data collection with complex physical layer techniques. The capacity that involves with multiple selected sources and destination has been studied in [5]. The capacity of data collection of single and multi sinks (FC) are investigated [9]. In [4], the authors derive capacity of data collection in arbitrary WSNs. A data collection capacity that considers delay and compressive sensing has been, recently, investigated in [75]. Most of the literatures calculate based on either the physical model or the protocol model. Physical model also known as the SINR model, is based on practical transceiver designs of communication system that treats interference as noise. Further, capacity calculation is based on Shannon’s formula. The other model is the protocol model. The model defines that a successful transmission occurs when a sensor falls inside the transmission range of its intended transmitter and falls outside the interference ranges of other non-intended transmitters. However, the protocol model is relatively inaccurate, when simultaneous transmissions are allowed in the network [5], [59].

The extensive survey on distributed estimation and capacity analysis in WSNs are investigated in Chapter 2.

1.6 Outline and Contributions of the Dissertation

This dissertation has novel contributions on distributed estimation and capacity of data collection. Optimizations of both distributed estimation and capacity of data collection are derived analytically. Correspondingly, the proposed method and performance evaluation are conducted by computer simulation. A general overview of each chapter is presented here as well as the main contributions. Moreover, the position of each chapter in covering several

research fields is shown in Fig. 1.2.

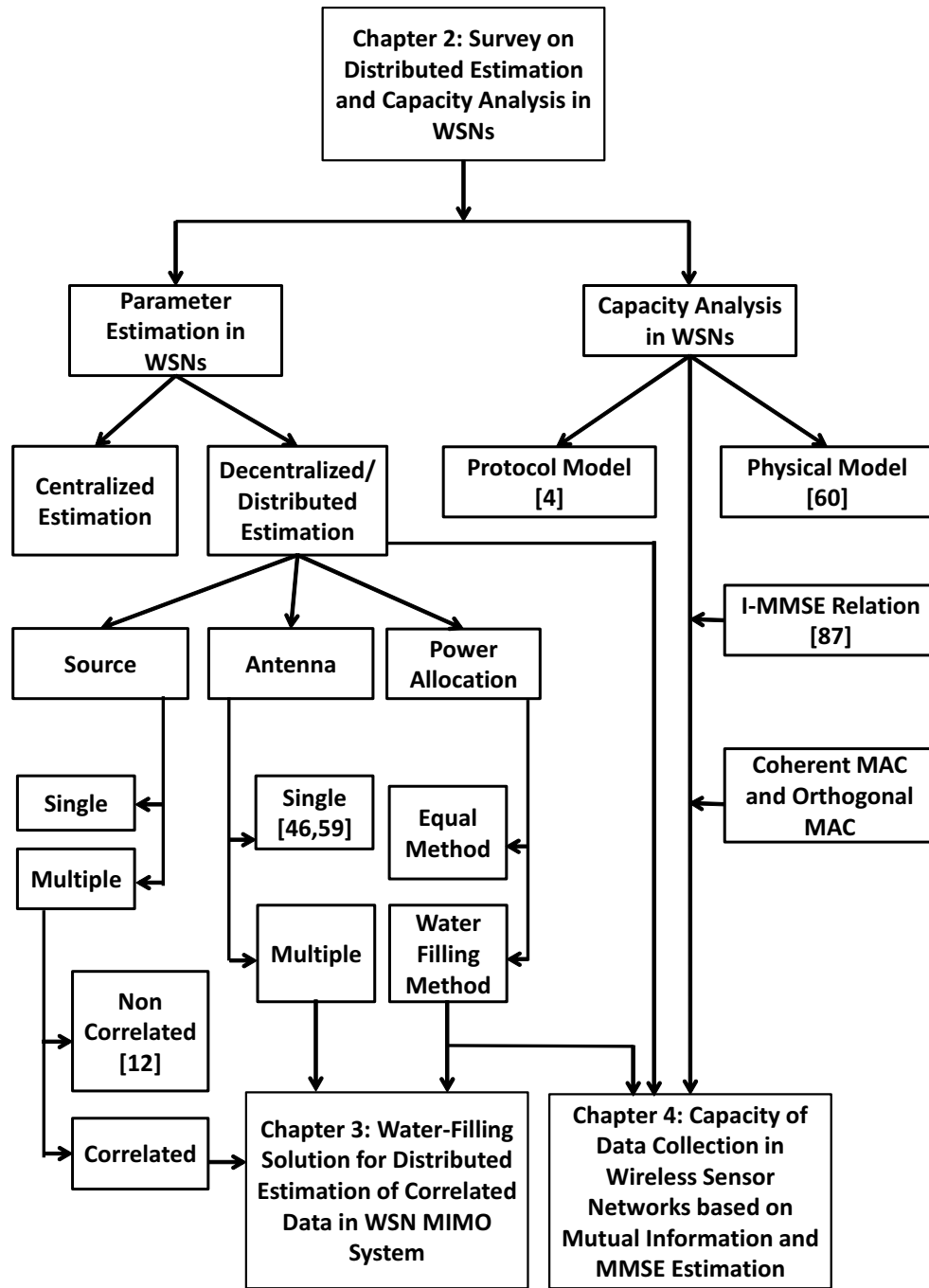


Figure 1.2: Position of each chapter in the existing research fields.

Chapter 3 introduces multiple targets estimation that are spatially distributed. If the targets are close to each other, the data observation will be potentially correlated. In addition, if the sensors require to deliver a large data observation such as motion sensor which is equipped by camera, multiple antenna transmission can be used. Previous work [14] does not consider correlation among data observation. In [58] they consider only single target transmitted using single antenna configuration. Our proposal is minimizing the distortion of estimation in terms of MSE by considering the correlated data observation and multiple antenna configuration. Given power constraint, we obtain closed-form solution that follows water-filling algorithm. Notations used in this chapter are summarized in Table 1.2.

Contributions: The proposed methods show that the distortion increases as the data becomes correlated. Simulation results show that the noisy channel is more harmful than the observation noise. By taking into account the effects of correlation, observation noise, channel noise, the proposed methods has a better performance than the equal power method.

Chapter 4 introduces channel capacity maximization in WSN called capacity of data collection. Maximization is conducted over two MAC models, i.e., coherent MAC and orthogonal MAC. Previous works [5], [59], capacity of data collection is calculated based on protocol model. Protocol model states that a successful transmission occurs when a sensor falls inside the transmission range of its intended transmitter and falls outside the interference ranges of other non-intended transmitters. However, the protocol model is relatively inaccurate, when simultaneous transmissions are allowed in the network. Our proposal is deriving capacity of data collection from the error estimation. We maximize the channel capacity by considering individual power constraint and sum power constraint. Optimization algorithms under both coherent MAC and orthogonal MAC are presented. Notations used in this chapter are summarized in Table 1.3.

Contributions: The proposed methods show that the capacity of data collection scale

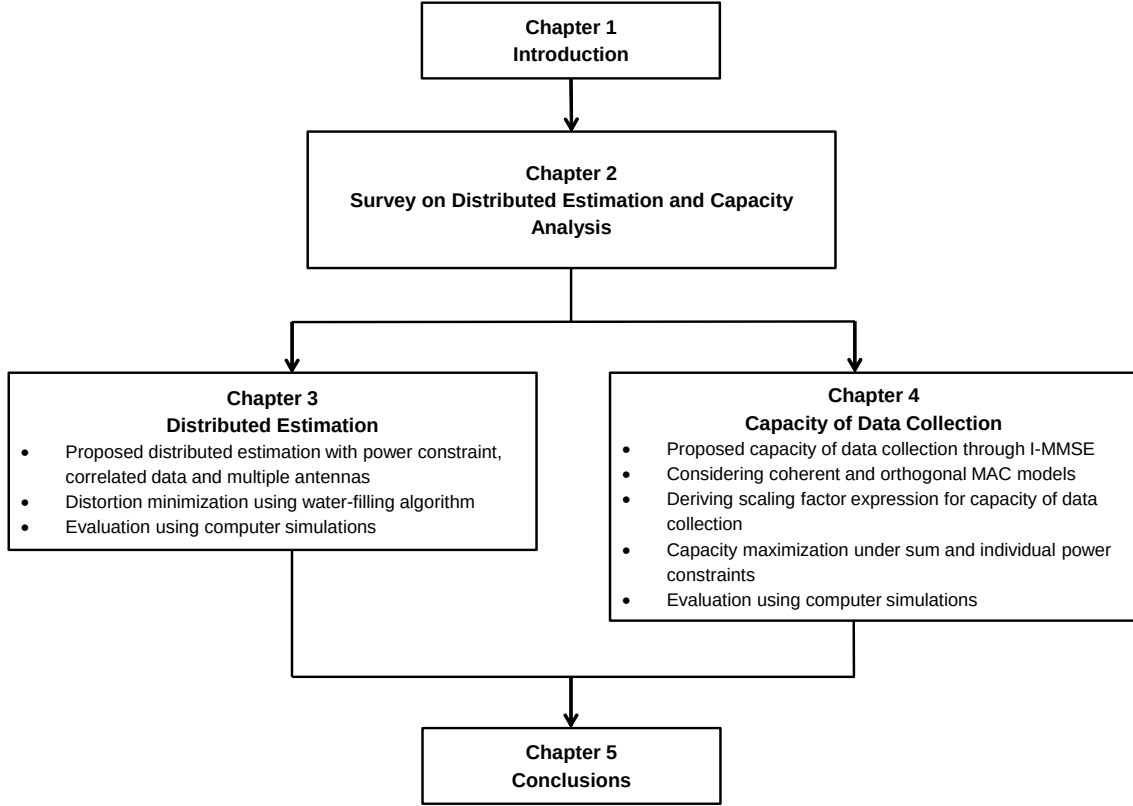


Figure 1.3: Configuration of this dissertation

as $\Theta((1/2)\log(1 + L))$ when the number of sensors L grows to infinity. The capacity of the optimal power allocation for both MAC models are larger than that of the equal power allocation. Simulation results show that the capacity of coherent MAC is larger than that of orthogonal MAC, particularly when the number of sensors L is large and the total power is finite. This is a consequence of using orthogonal link from the sensors to FC where the corrupted channel cannot be eliminated even when the number of sensors grows to infinity.

Chapter 5 summarizes results and contribution of this dissertation. Fig. 1.3 describes the configuration of this dissertation.

Table 1.1: Notation used in Chapter 2

Notation	Discription
N	Number of sensors
$f_i(.)$	Probability density function of sensor i
θ	Original parameter
$\hat{\theta}$	Estimated parameter
x_i	Observation of sensor i
P_e	Detection error probability
$\mathbf{m}$	Signal messages
$\mathbf{A}$	Precoding matrix
$Q(.)$	Quantization function
B	Total bit rate
b_i	Quantization bit rate of sensor i
V	Dynamic range of sensor
$\mathbf{y}$	Received signal
h	Channel gain
u	Channel noise
P	Total power
K	Ricean factor
σ^2	Noise variance
d_i	Distance of node i to the origin
W	Total channel capacity
λ	Throughput each sensor
r	Radius of node
A_r	Size of node area
μ	Minimum uncovered area
H	Number of cluster heads
Δ	Incremental distance of each node
ϵ	Arbitrary number close to zero
T	BFS Tree
R	Radius of interference area
α	Constant parameter
δ	Path loss exponent
F_i	Path i
G_i	i th hop
c	Number of leafs in BFS Tree
C	Capacity
n	Number of sensor interferers
D	Bandwidth for all channels
ρ	Distance of sensor to the FC

Table 1.2: Notation used in Chapter 3

Notation	Discription
L	Number of sensors
p	Number of source signals
$\mathbf{s}$	Vector of source signals
$\hat{\mathbf{s}}$	Vector of estimated source signals
$\bar{\mathbf{s}}$	Vector of estimated source signals due to observation noise
$\bar{\bar{\mathbf{s}}}$	Vector of estimated source signals due to channel noise
$\mathcal{C}$	Complex number
$\mathbf{F}_l$	Sensing matrix of l th sensor
k	Number of measurements
$\mathbf{x}_l$	Sensor observations of l th sensor
$\mathbf{n}_l$	Observation noise of l th sensor with zero mean and variance σ^2
$\mathbf{r}$	Correlation matrix
N	Number of Encoded messages
$\mathbf{A}_l$	Coding matrix of l th sensor
$\mathbf{H}_l$	Channel matrix l th sensor
$\mathbf{y}$	Received signal at FC
$\mathbf{v}$	Additive white Gaussian noise (AWGN) with zero mean and variance σ^2
$\mathbf{B}_l$	Intermediate variable of $\mathbf{H}_l \mathbf{A}_l$
D_t	Total distortion in terms of mean square error (MSE)
$\tilde{D}_t$	Approximate of total distortion
D_o	Distortion due to observation noise
D_c	Distortion due to channel noise
D_{low}	Lower bound of total distortion
$\mathbf{R}_{xo}$	Covariance matrix of sensor observations
$\mathbf{R}_{nl}$	Covariance matrix of observation noise
$\mathbf{R}_s$	Covariance matrix of source signals
$\mathbf{R}_v$	Covariance matrix of channel noise
$\mathbf{R}_{sx}$	Covariance matrix of between sensors and sources
$\mathbf{R}_{xc}$	Covariance matrix of channel
$\mathbf{R}_{xl}$	Covariance matrix of l th sensor
P	Total power
$\text{tr}(\cdot)$	Trace of matrix $(\cdot)$
d	Distance between sources
ρ	Spatial correlation coefficient
α_i, β_i	i th eigenvalues of positive semidefinite matrices $\mathbf{X}$ and $\mathbf{Y}$, respectively
$\mathbf{U}$	Unitari matrix
$\hat{\mathbf{U}}$	The first p th columns of matrix $\mathbf{U}$
$\mathbf{V}$	Eigenvector
$\Sigma(\cdot)$	Eigenvalue matrix of $(\cdot)$
f_i, g_i, w_i, t_i, n_i	i th eigenvalue of f, g, w, t, n , respectively
$\mathbf{T}$	Intermediate variable of $\mathbf{B} \mathbf{R}_{xo}^{1/2}$
$\mathbf{W}$	Intermediate variable of $\mathbf{B} \mathbf{R}_{xc}^{1/2}$
μ	Water level of water-filling algorithm

Table 1.3: Notation used in Chapter 4

Notation	Discription
X, Y, N	Input , output, and noise variables, respectively
$E[.]$	Expected value of $[.]$
C	Channel capacity
C_{upp1}	Upper bound of coherent MAC capacity
C_{upp2}	Upper bound of orthogonal MAC capacity
C_{eq1}	Coherent capacity of equal power method
C_{eq2}	Orthogonal capacity of equal power method
C_{opt1}	Coherent capacity of optimal power allocation
σ_X^2	Variance of X
L	Number of sensors
x_i	Signal observation of i th sensor
s	Source signal with zero mean and variance σ_s^2
n_i	Observation noise of i th sensor with zero mean and variance $\sigma_{n_i}^2$
P_i	Average transmit power of i th sensor
P_i^{opt}	Optimal average transmit power of i th sensor
P	Total transmit power of system
a_i	Amplification factor of i th sensor
α_i	signal to noise ratio (SNR) of i th sensor
g_i	Channel gain of i th sensor
v	Channel noise with zero mean and variance σ_v^2
y	Received signal
J	Mean Square Error (MSE)
J_0	Lower bound of MSE
J_{eq1}	Coherent MSE of equal power method
γ	Total SNR observation
β	Total SNR
β_{opt}	Optimal total SNR
c	Broadcast constant from fusion center (FC)
c_i	Local constant of i th sensor
c_0	Local threshold of i th sensor
c_g	Normalized constant
y_i	Received signal of i th sensor in orthogonal multiple access channel (MAC)
v_i	Channel noise of i th sensor in orthogonal MAC
L_1	Number of active sensors
d	Uniform random variable
ϵ	Path loss parameter
$\mathcal{L}(.)$	Lagrangian function
t	Slack variable
μ	The first water level
η	The second water level

Chapter 2

Survey on Distributed Estimation and Capacity Analysis in WSNs

In this Chapter, we review several works related to distributed estimation and capacity analysis in WSNs. Because the distributed estimation is a part of distributed inference, we start the discussion from the distributed inference. Furthermore, we review the capacity analysis in the next section.

2.1 Distributed Inference

In WSNs, one of the purposes is how to infer information from the area of interest or observation to be known and useful for user applications. Information is gathered by a distributed sensor in the field. Thus, a new paradigm, a distributed inference is necessary to efficiently coordinate the large number of sensor networks to achieve high level information processing [90]–[94].

Distributed inference in WSNs can be classified into four categories, i.e. estimation, de-

tection, localization, and tracking. In estimation, sensors observe a source(s) and send to FC. FC receives signals from them. Using some estimation techniques such as Minimum Mean Square Error (MMSE), Maximum Likelihood Estimator (MLE), and Best Linear Unbiased Estimator (BLUE), the FC estimates the sources. Moreover, in case of detection, FC reconstructs the state of the sources. The state can be defined as the presence of the sources before the sensors estimate the attributes such as position and velocity [95]. Distributed localization has a purpose to know where the sources are. Distributed localization is different from the conventional location services and should be operated in energy efficient, self-organizing, and robust [96]. In terms of large sensors, distributed localization is also expected to be deployed in a scalable sensor networks. Furthermore, in tracking application, given target locations by sensors at time instants in the past, it is possible to fit the data samples into a dynamic model to predict future target locations [97]–[98]. Such kind of scenarios can be applied for single or multiple targets. For a single moving target, sufficiently accurate tracking may be accomplished by fitting the data into a linear or polynomial model using a least square fit. Tracking is a complicated problem when multiple targets are present. Target tracks can cross paths resulting in the data association problem [99].

Information sharing and collaborative processing among sensors are also essential to achieve enough precision, because the data from a single low cost sensor is too coarse to derive a reliable inference. Collaborative processing is a multi-sensor data fusion layer, which is to fuse the data received from multiple sensors and extract the useful information. Thus, in the context of information sharing and collaborative processing, there are two types of processing, i.e. centralized and decentralized inferences.

Centralized and Decentralized Inference

A network can be classified as centralized inference if FC receives all raw data from sensor observations. The sensors are responsible for transmitting their observation to FC without any channel distortion [10]. For example, FC has the knowledge of the sensor noises such as pdf and/or variance of noises. Under these conditions, sensor can perfectly send their observations to FC. Benefit of centralized estimation, we can serve as a performance benchmark of the system. In case of parameter estimation, FC can use MLE and BLUE [56]. If FC knows pdf of sensor noises $f_i(\cdot)$, ($i = 1, \dots, N$) as prior information, the centralized MLE to estimate parameter θ on observation $\mathbf{x}$ is

$$\begin{aligned}\bar{\theta}_{ML} &= \arg \max_{\theta} f(\mathbf{x}|\theta) \\ &= \arg \max_{\theta} \sum_{i=1}^N \log f_i(x_i - \theta).\end{aligned}\tag{2.1}$$

BLUE can be applied if FC knows noise variance of sensors. BLUE is a combination of observation $\mathbf{x}$ with weights inversely proportional to noise variance of i th sensor, σ_i^2 as

$$\bar{\theta}_{BLUE} = \left(\sum_{i=1}^N \frac{1}{\sigma_i^2} \right)^{-1} \sum_{i=1}^N \frac{x_i}{\sigma_i^2}\tag{2.2}$$

where N is the number of sensors.

However, characterizing the exact noise pdf of each sensor is difficult in WSNs. Moreover, sending all raw data requires sufficient bandwidth and high transmission energy. These drawbacks make the centralized inference impractical, because WSNs should operate in energy and bandwidth efficient.

Instead of sending all raw data to FC, local processing at sensors can perform such as,

compression [45], quantization [100], or encoding [101] to reduce the dimension of the data. Reducing the dimension of the data can reduce communication cost including bandwidth and energy. This paradigm is referred to decentralized inference. It can be accomplished as follows. First, each sensor performs a local encoding into spatio-temporal sensor data in an independent and identically distributed (i.i.d.) fashion. This encoding compresses i th sensor observation, $\mathbf{x}_i$ to $\mathbf{m}_i$ by mapping $\mathbf{m}_i = f_i(\mathbf{x}_i)$, where $f_i(\cdot)$ is a linear mapping function. The encoded message $\mathbf{m}_i$ is then transmitted to FC. Second, all the encoded messages are combined at FC to produce a global inference. In decentralized estimation, the quality is measured by the MSE criterion:

$$\text{MSE} = E(|\bar{\theta} - \theta|^2) \quad (2.3)$$

where $\bar{\theta}$ and θ are estimated parameter and original parameter, respectively. In decentralized detection, the quality is measured by the Detection Error Probability (DEP):

$$P_{e,\theta} = P_{\theta}(\bar{\theta} \neq \theta). \quad (2.4)$$

2.2 Distributed Estimation

As a part of distributed inferences, we focus only on distributed estimation. Moreover, we should not confuse of using decentralized and distributed terminology, because these have the same meaning.

2.2.1 Distributed Estimation in Bandwidth Constraint

Capacity of WSN is limited, because wireless channel is shared across the networks. Hence, considering the bandwidth constraint is important. There are two models: the first is limiting the binary bits being sent to FC and the second is limiting the number of real-value messages being sent by sensors to FC. The first model is similar with digital transmission scenario [76]–[82], while the second model is similar with analog transmission scenario [83]–[86], [88].

In case of the first model, sensors compress the dimension of sensor observations. Sensors send the compressed data to FC in a fixed number of real-messages. Compressing the observation data has a meaning that the vector dimension of transmitted data is less than the vector dimension of sensor observation. The simple application of this technique is a linear mapping [9], [101]. The number of real-messages can be determined by the degrees of freedom of the channel between a sensor to FC. For example, the observations $\mathbf{x}_i$ from sensor i will be mapped into real messages $\mathbf{m}_i$, we have

$$\mathbf{m}_i(\mathbf{x}_i) = \mathbf{A}_i \mathbf{x}_i \quad (2.5)$$

where $\mathbf{A}_i$ is a precoding matrix of sensor i .

In case of the second model, sensor observations are digitized and quantized into one or several number of bits. Every sensor performs a local quantization,

$$\tilde{x}_i = Q_i(x_i) \quad (2.6)$$

where $Q_i(\cdot)$ is a quantization function. Given total bit rate B , the objective function is

derived as minimization of MSE with respect to total bit rate

$$\begin{aligned} \min \quad & \text{MSE} \\ \text{s.t.} \quad & \sum_{i=1}^N b_i \leq B, \quad k = 1, \dots, N \end{aligned} \quad (2.7)$$

where b_i is bit rate of i th sensor quantization. Furthermore, the sensor quantization describes as follows :

- Suppose sensor observation is bounded as $x \in [-V, V]$, where V is decided by sensors dynamic range.
- Variance of observation noise is assumed as σ^2 .
- The range $[-V, V]$ is divided uniformly using intervals of length $\delta = (2V)/(2^{b_i} - 1)$.
- As shown in [102][103], the variance of unbiased estimator can be expressed

$$\text{MSE} \leq \sigma^2 + \frac{V^2}{(2^b - 1)2} \quad (2.8)$$

where $\frac{V^2}{(2^b - 1)2}$ is the upper bound of the quantization noise variance.

An equivalent of 1-bit MSE function is also introduced in [103] to determine the optimal quantization bit rate of each sensor. The 1-bit MSE function is derived as

$$g(\sigma^2, b) = b \cdot \left(\sigma^2 + \frac{V^2}{(2^b - 1)2} \right). \quad (2.9)$$

Using 1-bit MSE function, they propose quantization for homogeneous and heterogeneous sensor networks. Moreover, they investigate a theoretical upper bound of estimation MSE

with factor of 2.2872 as

$$\text{MSE} \leq 2.2872 \left(\sum_{i=1}^{N^{opt}} \frac{1}{\sigma_i^2} \right)^{-1} \quad (2.10)$$

where N^{opt} is the number of optimal sensors with small observation noise variances.

2.2.2 Distributed Estimation in Coherent and Orthogonal MAC

Transmission between sensors and FC can be conducted in two fashions, coherent MAC and Orthogonal MAC [9]. Coherent also means that the signals from sensors add up coherently at FC. Transmissions are simultaneous and in the same frequency band, keeping the utilized bandwidth independent of the number of sensors. As a consequence, FC cannot select individual transmission. Sensors are synchronized before they transmit the signals. For example, the received signal at FC can be expressed as

$$\mathbf{y} = \sum_{i=1}^N h_i m_i(x_i) + u \quad (2.11)$$

where N and h_i are the number of sensors and channel gain of sensor i , respectively. $m_i(\cdot)$ and x_i are transmitted messages and sensor observation, respectively. u is noise at FC. The noise may represent as channel noise or temperature noise. We can see that the noise incurs only once, while orthogonal MAC means that the transmissions do not interfere with each other. Signals arrive at FC individually. FC can choose some signals to improve the quality of estimation. Because the signals reach FC individually, we can write signal reception from sensor i as

$$\mathbf{y}_i = h_i m_i(x_i) + u_i. \quad (2.12)$$

We can see that the signals are received individually at FC. As consequence, the signals also experience an individual noise, which has disadvantages for low power signals. The signals degrades easily as the noise power increases.

2.2.3 Distributed Estimation with and without Channel State Information (CSI)

When the sensors experience fading, sensors need a channel knowledge to improve performance of estimation. In [83], the performance of estimation under availability of CSI is investigated. Considering that sensors communicate with FC using orthogonal MAC, several channel models such as Rayleigh and Ricean fadings are discussed. Performance analysis is conducted using an asymptotic MSE. The asymptotic MSE is computed when the number of sensors N goes to infinity.

If there is no CSI, the channel between sensors and FC simply adopts i.i.d. channel statistics. Given channel gain of i th sensor h_i , the asymptotic MSE under the channel statistics can be expressed as

$$\text{MSE}(\mathbf{h}) = \frac{\sigma^2 P + 1}{P|E[h_i]|^2} \quad (2.13)$$

where $\mathbf{h} = [h_i, \dots, h_N]$. P is total power. For Rayleigh fading channel, the value of asymptotic MSE does not converge to zero resulting poor performance of MSE. This is because the value of the inverse of an exponential random variable does not exist. In case of channel which has nonzero mean such as Ricean fading channel, the asymptotic MSE can be expressed as

$$\text{MSE}_{\text{NoCSI}} = \frac{\sigma^2 P + 1}{P} \frac{K + 1}{K} \quad (2.14)$$

where K is Rician factor. As K increases, the channel has less fading. In case $K \rightarrow 0$, the MSE grows without bound.

If there is perfect CSI, amplifier used by each sensor may depend on the channel response. The optimal MSE can be obtained by convex optimization by minimizing MSE with subject to total power. The optimum conditional MSE can be expressed as

$$\text{MSE}_{\text{CSI}}(\mathbf{h}) = \left(\sum_i^N \frac{1}{\sigma^2 + \frac{1}{P|h_i|^2}} \right)^{-1} \quad (2.15)$$

If the sensor amplifier does not support a large dynamic range in response of channel gain, the sensor can be considered for having only a constant gain. Then, sensor is only able to adjust for the phase of its channel. This condition can be called as phase only CSI. Considering Rayleigh fading, the asymptotic MSE under phase only CSI can be written as

$$\text{MSE}_{\text{POCSI}} = \frac{\sigma^2 P + 1}{P} \frac{4}{\pi}. \quad (2.16)$$

Moreover, under Ricean fading channels can be expressed as

$$\text{MSE}_{\text{POCSI}} = \frac{\sigma^2 P + 1}{P} \frac{1}{(K + 1)\gamma^2(3/2)e_1^{-2K}F_1^2(3/2; 1; K)} \quad (2.17)$$

where γ and $F_1^2(.;.;.)$ are Gamma function and confluent hypergeometric function [104], respectively. Interestingly, the different value of asymptotic MSE between Rayleigh fading channel and Ricean fading channel lies between $\frac{4}{\pi}$ and 1 when K varies between 0 and ∞ .

2.2.4 Distributed Estimation in Energy Constraint

Distributed estimation in WSNs has to consider energy constraint, because sensors are equipped with small batteries. Recharging or replacing the batteries are costly. One of the major objectives of WSNs is to design energy efficient devices, protocol, and algorithm [2]. Given total power P and the number of sensors N , minimization of MSE is expressed as,

$$\begin{aligned} \min \quad & \text{MSE} \\ \text{s.t.} \quad & \sum_{i=1}^N P_i \leq P, \quad P_i > 0 \end{aligned} \quad (2.18)$$

where P_i is the allocated power per sensor.

Solving the objective function under total power constraint, we arrive at the tradeoff between a subset of active sensors and the energy used by each active sensor to minimize the MSE. The solution of the problem is a sensor scheduling. In [9], sensors are scheduled using parameters such as quality of sensor observation and channel quality between sensor and FC. Power is allocated to sensors which have good these parameters, otherwise sensors are considered being idle or deactivated.

2.3 Capacity Analysis

Each sensor measures source periodically and sends to FC at a certain time. This requires a channel to deliver these data to FC, and computing capacity becomes important. However, sensors have some challenges such as channel noise, limited power, and limited bandwidth. Moreover, the performance of channel capacity can be characterized by the rate at which sensing data can be collected and transmitted to FC. The theoretical measure that captures

the limitations of collection processing in WSNs is capacity for data collection in WSNs.

The study of capacity analysis in sensor networks is initiated by Gupta and Kumar [59]. They formulate the capacity expression based on ad hoc network, which sensors communicate to each other using peer to peer network. The data is routed to the destination node randomly. In [105], they use a mobile destination node to reduce the number of hops. Hence, the capacity can be improved. The advantages of using relay node to improve the capacity are investigated in [106].

However, the network topology of ad hoc network is different with WSNs. The network topology in WSNs exists the FC. Hence, the characteristic network changes being many-to-one scenario. The existing sensors between source and FC serve as channel encoder. In conventional approach, there are two models for investigating capacity analysis in WSNs, i.e., protocol model [73]–[74] and physical model [6], [7].

2.3.1 Capacity Analysis based Protocol Model

The protocol model is also known as a disk graph model. This model describes that successful transmission occurs when a sensor falls inside the transmission range of its intended transmitter and falls outside the interference ranges of other non-intended transmitters. Because of spatial separation, some sensors can successfully transmit their data and some others cannot. Thus, a fixed transmission range r can be defined such that the FC can successfully receive the signal sent by sensor i only if $\|d_{FC} - d_i\| \leq r$, where, $\|d_{FC} - d_i\|$ is the Euclidean distance between FC and sensor i . Capacity analysis based on protocol model is investigated under different network models including flat network, hierarchical network, and arbitrary network.

Capacity in Flat Network

The flat network architecture where sensors communicate with the FC via possibly one hop or multi-hop routes by using peer sensors as relays [73]. Given the total transmission capacity of the channel W , we can derive a trivial upper bound of throughput of each sensor. The maximum throughput can be achieved when FC is 100% busy. If N sensors communicate to FC simultaneously, the achieved throughput by any single sensor λ , must satisfy

$$\begin{aligned} N\lambda &\leq W \\ \lambda &\leq \frac{W}{N} \end{aligned} \tag{2.19}$$

We can see that W/N bits/s per sensor is the upper bound for throughput that can be achieved on average by each sensor. Moreover, this formulation is valid if only sensor can directly reach FC. FC can use Time Division Multiple Access (TDMA) scheme to schedule sensor transmission. Thus, each sensor gets an equal share of the channel as $\lambda = W/N$.

The successful transmission between sources and FC can be limited by using transmission range r . Sensors which have a greater transmission range than r , have a high probability that the throughput is below W/n . This is due to the maximum number of simultaneous transmissions. For every pair of source-FC communication, there is an interference limiting this transmission. We can easily derive an area for ideal transmission as

$$A_r = \pi r^2. \tag{2.20}$$

If there are N transmissions, there needs equal space among N sensors. The number of

simultaneous transmissions which can be accommodated is

$$N \leq \frac{1}{\pi r^2}. \quad (2.21)$$

Therefore, regardless of how the circles of sensors are arranged, there will be always some uncovered areas. Given $\mu = r^2(2\sqrt{3} - \pi)$ as the minimum uncovered area per source-FC transmission pair, we have

$$N \leq \frac{1}{\pi r^2 + \mu}. \quad (2.22)$$

After an elegant derivation [73], we get

$$\lambda \leq \frac{W(\pi r^2 + \sqrt{\epsilon})}{N(\pi r^2 + \mu)} < \frac{W}{N} \quad (2.23)$$

where ϵ can be made arbitrarily close to zero, as $N \rightarrow \infty$. Thus, we can see that $\lambda = \frac{W}{N}$ is not achievable.

Capacity in Hierarchical Network

In this network, sensors closely located to each other send the data through a cluster head. The cluster head communicates with FC using single hop transmission. Cluster head cannot transmit and receive simultaneously. The main purpose of using cluster head is an expectation that the sensor can achieve throughput as $\lambda = W/N$, particularly when cluster heads have the same capacity W . Hence, to achieve the maximum capacity, the FC should be busy all the time. The cluster head transmits the same amount of data, because the cluster head is assumed to have the same number of sensors. Given the number of cluster heads H , it follows that each cluster head sends $1/H$ fraction of the time or leaves $1 - 1/H$ for receiving

from sensor members.

It is also interesting to know the number of clusters required to achieve $\lambda = W/N$. It can be formulated as

$$H \geq \frac{5\pi r^2 + 4\pi r\Delta + \pi\Delta^2}{\pi r^2 - \sqrt{\epsilon} + (4\pi r^2 + 4\pi r\Delta + \pi\Delta^2)\sqrt{\epsilon}} \quad (2.24)$$

where Δ is an incremental distance between two sensors to avoid interference. Furthermore, if we use the same analogy with the flat network, the cluster head is similar with the cluster head against the FC. Hence, we can write

$$H \leq \frac{1}{\pi(2r + \Delta)^2}. \quad (2.25)$$

Using the two previous equations, we arrive at

$$\frac{20r^4 + 36\Delta r^3 + 25\Delta^2 r^2 + 8\Delta^3 r + \Delta^4}{r^2 - \sqrt{\epsilon}(4r^2 + 4r\Delta + \Delta^2 - 1/\pi)} \leq \frac{1}{\pi}. \quad (2.26)$$

In case $\Delta = 0$, where the transmission range of each cluster head to FC is equal, we can compute the transmission range as $r < \sqrt{\frac{1}{20\pi}}$. In case $\Delta = r$, we can express the transmission range being $r < \sqrt{190\pi}$.

Finally, given the transmission range r , we can formulate the number of cluster heads as

$$H \geq 3 - \pi r^2. \quad (2.27)$$

Capacity in Arbitrary Network

Sensor deployment in WSNs may not be large as in theory [4]. Therefore, it is necessary to study the capacity analysis in an arbitrary network. Starting from the disk graph model,

the authors propose an algorithm called Breadth First Search (BFS) tree. The purpose of using the algorithm is to achieve the theoretical upper bound as W/N , where W and N are total capacity and the number of sensors, respectively.

The algorithm of BFS consists of two steps, i.e. data collection formation and data collection scheduling.

Data collection formation defines that given an BFS tree T under protocol interference model, the maximum interference number δ_i on every path F_i can be bounded by a constant $8\alpha^2$, where $\delta_i \leq 8\alpha^2$. $n^{F_i}(n_j)$ is the number of sensors on path F_i which are inside in the interference range of interferer n_j . The area of the interference region of n_j can be denoted as πR^2 and the radius of sensor inside in the interference region is $\frac{r}{2}$. Hence, we can express at most $\frac{\pi R^2}{\pi(\frac{r}{2})^2} = 4\alpha^2$. Using Pigeonhole principle, we can state that there are more than $\frac{8\alpha^2}{4\alpha^2} = 2$ sensors inside the single small disk with radius of $\frac{r}{2}$.

The next step is branch scheduling. Branch scheduling is an algorithm to count how many time slots that the FC needs to receive all data from one snapshot transmission. The algorithm can be written as

Algorithm 1. Branch Scheduling on BFS Tree.

Input: T : BFS tree, c : the number of leaves on T ,

1. **for** each snapshot **do**
2. **for** $t = 1$ to c **do**
3. Collect data on path F_i . All sensors on F_i transmit data to FC using time scheduler.
4. The collection is terminated when sensors on branch G_i do not have data for this snapshot. The total slots used are at most $\delta_i \cdot |G_i|$, where $|G_i|$ is the hop length of G_i .
5. **end for**
6. **end for**

Using this algorithm, capacity of W/N can be achieved.

2.3.2 Capacity Analysis based Physical Model

The physical model is also known as a transceiver model. The capacity expression is modeled as function of SINR. Capacity analysis in WSNs is conducted using many-to-one channel wireless networks. The pioneer work is presented by Gupta and Kumar in a large scale ad hoc wireless networks [59]. The result suggests that the capacity of ad hoc network scales as $C(\sqrt{N})$, where N is the number of sensors. However, the ad hoc network is different with WSNs. The ad hoc network is peer to peer communication, which information flows from an arbitrary sensor to another arbitrary sensor, while WSNs have a difference topology, which information is collected a single node, i.e. FC.

In this subsection, we present several capacity calculations based on different scenarios. In [7], they introduce cooperative Time Reversal Communication (TRC) with help of cluster heads. In this scenario, the multiple sensors send information to the same destination. A group of sensors cooperates to transmit a common data stream to FC. During a training phase, FC transmits a short sequence of training pulses received by all sensors. Thus, sensors estimate its receive pulses in terms of arrival time, duration, and shape of the received pulses. After the training phase, the sensing data are transmitted from sensors to FC. Before transmission, sensors disseminate the information to all sensors then sensors separately transmit the information to the sink using synchronous transmission. The capacity expression of TRC can be written as

$$C(N_t) = D \log \left[1 + \frac{\left(\sum_{i \in N_t} A m_i \rho_i^{-\delta/2} \right)^2}{D\sigma^2 + \sum_{k \in I} P_k \rho_k^{-\delta}} \right] \quad (2.28)$$

where D represents common bandwidth for all links on the network. Am_i is the average amplitude of transmitted signal from sensor i . N_t and ρ_i are the set of cooperative sensors and the distance of sensor i to FC, respectively.

In the absence of cluster head, the capacity computation using one hop scenario is presented in [6]. The independent data streams are transmitted from different sensors to FC. Given the number of sensors N , C_N is defined as the maximum number of bits that can be transmitted from sensors N to FC. If there are two positive constant c_1, c_2 , capacity expression can be written

$$c_1 \log(N) \leq C_N \leq c_2 \log(N) \quad (2.29)$$

as $N \rightarrow \infty$. $\log(.)$ is assumed to have a base 2. The eq. (2.28) presents upper and lower bounds. The upper bound can be obtained by assuming a genie informs every source to all sensors. Hence, this network is similar with one to one channel with N transmit antennas and one receive antenna. The capacity of this scenario can be expressed as

$$c_2 \log(N) = \log(1 + NP_{\text{total}}) \quad (2.30)$$

where P_{total} is total power.

To provide lower bound, every sensor is allowed to distribute its information to closely located sensor. Assuming there is a dense sensor networks, the cost of the cooperative communication is relatively small. Using beamforming technique, the received power is considered to increase as logarithmic. If there are N sensors which transmit simultaneously,

we have a lower bound as

$$C_N \geq c_1 \log(N) \quad (2.31)$$

where

$$c_1 = \frac{\delta}{2\delta + 1} - \epsilon/2 > 0$$

where δ and ϵ are path loss exponent and an arbitrary number > 0 , respectively.

To this end, we present the capacity analysis based on protocol model and physical model. In Chapter 4, we propose a new derivation of capacity expression in WSNs using the relationship between Mutual Information and Minimum Mean Square Error (I-MMSE). The analysis provides the capacity expression under coherent and orthogonal MAC models.

Chapter 3

Water-Filling Solution for Distributed Estimation of Correlated Data in WSN MIMO System

3.1 Introduction

In Chapter 2, much work on the distributed estimation have been done including bandwidth constraint, power constraint, MAC model, and the presence/absence of CSI. In this Chapter, we focus on the distributed estimation considering source correlation and multiple antenna scenario. We also takes into account those factors, which are discussed in Chapter 2. The correlated data has been considered in [68], [69]. In [68], the correlated data is assumed due to a set of spatially distributed sensor nodes. In [69], the correlated data is assumed due to a result of colored observation noise. Here, we consider a practical scenario, if the sources are close to each other, the data will be potentially correlated. Considering multiple sources, the receiver can be ordered to estimate such as the number of sources and position

of each source [108]. When the sources are being correlated, the sources become difficult to be identified, counted, and known of their positions. In this chapter, we consider multiple targets that are spatially distributed. The targets are observed by multiple sensors that apply an analog forwarding scheme. This scheme will multiply the observed data with a designed coding matrix in each sensor, which results in encoded messages. The encoded messages, which are signals from the sensors, are transmitted to the FC over a coherent MAC. We consider only one hop communication, because we focus on the effect of source correlation, noisy observation, channel noise, and power allocation method. The channel follows MIMO model that has multiple antennas at transmitter and at receiver. At the FC, the received signals are estimated using Linear Minimum Mean Square Error (LMMSE) rule. To perform the estimation, we calculate a distortion in terms of MSE. We also consider the separation of the estimation under two conditions, i.e., distortion due to noisy observation and distortion due to channel noise. Both of the distortions should be minimized by designing coding matrices under total power constraint. To derive the coding matrix, we use Singular Value Decomposition (SVD) technique. We show that the equations can be formulated as a convex optimization problem. We derive a closed-form solution that can be solved using water-filling algorithm, which is one of the optimal way in terms of power allocation to minimize the distortion. The designed coding matrices will be sent back to the sensors and will optimize transmit power of them. Under the above scenario, the proposed method will be compared to the equal power method. The equal power method allocates equal transmit power to each sensor. Compared to the existing literature, the contribution of this work lies in the following aspects: we consider MIMO model based on the water-filling algorithm and a spatial correlated data as an extension of [14], [58]. We obtain a closed-form solution for the optimization problem and give the coding matrices for each sensor. We also derive a lower bound of the MSE.

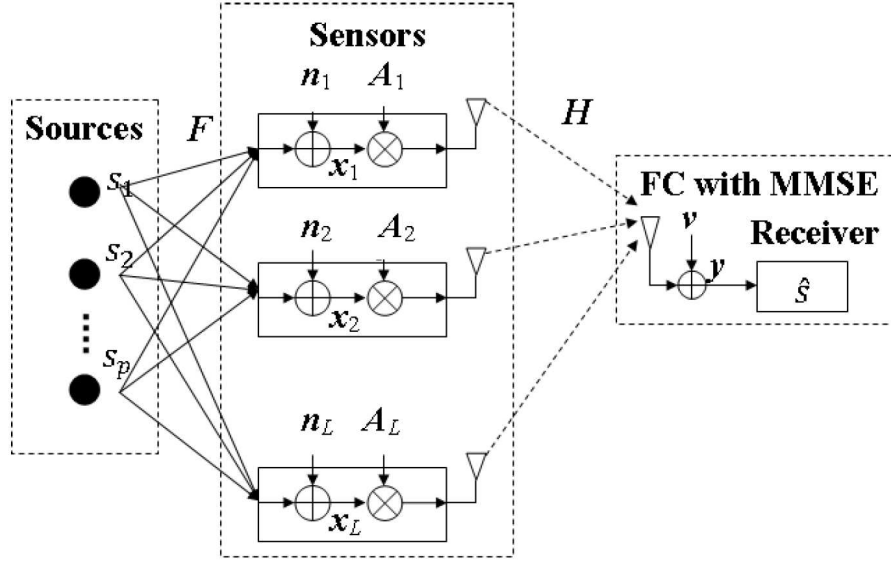


Figure 3.1: Distributed estimation over a wireless sensor network.

The rest of this chapter is organized as follows. Section 3.2 describes the network setup under consideration. In Section 3.3, we formulate the linear coding matrix based on total power constraint and correlated sources. Section 3.4 presents some numerical simulation examples and the conclusion is drawn in Section 3.5.

3.2 Problem Formulation

Assume that there are L sensors for estimating p random source signals, written in a vector form $\mathbf{s} = [s_1, \dots, s_p]^T \in \mathbb{C}^p$, as shown in Fig. 3.1. The sensors observe the sources through a sensing matrix $\mathbf{F} \in \mathbb{C}^{k \times p}$ and each sensor has k measurements given by

$$\mathbf{x}_l = \mathbf{F}_l \mathbf{s} + \mathbf{n}_l, \quad 1 \leq l \leq L \quad (3.1)$$

where $\mathbf{n}_l \in \mathbb{C}^k$ is the additive noise. We assume that the targets are close to each other. They become potentially correlated. Therefore, the sensing matrix, $\mathbf{F}$, can be modeled as $\mathbf{F} = \mathbf{F}_o \mathbf{r}^{1/2}$ where $\mathbf{F}_o$ is a matrix with size $k \times p$ that has i.i.d. zero mean circularly symmetric complex Gaussian (ZMCSCG) entries with unit variance, and $\mathbf{r}$ is the $p \times p$ spatial correlation matrix. If we have L sensors, eq. (3.1) can be written as

$$\mathbf{x} = \mathbf{F}\mathbf{s} + \mathbf{n} \quad (3.2)$$

where $\mathbf{x} = [\mathbf{x}_1^T, \dots, \mathbf{x}_L^T]^T \in \mathbb{C}^{Lk}$, $\mathbf{F} = [\mathbf{F}_1^T, \dots, \mathbf{F}_L^T]^T \in \mathbb{C}^{Lk \times p}$, and $\mathbf{n} = [\mathbf{n}_1^T, \dots, \mathbf{n}_L^T]^T \in \mathbb{C}^{Lk}$ with $Lk = \sum_{l=1}^L k_l$. The sensor observations are encoded using a linear coding matrix $\mathbf{A}_l \in \mathbb{C}^{N \times k}$, where N is the number of encoded messages transmitted from the l th sensor. The message vector is transmitted through channel matrix $\mathbf{H}_l \in \mathbb{C}^{N \times N}$ using N different frequencies. The received signal at FC can be written as

$$\begin{aligned} \mathbf{y} &= \sum_{l=1}^L \mathbf{H}_l \mathbf{A}_l (\mathbf{F}_l \mathbf{s} + \mathbf{n}_l) + \mathbf{v} \\ &= \mathbf{B}(\mathbf{F}\mathbf{s} + \mathbf{n}) + \mathbf{v} \end{aligned} \quad (3.3)$$

where $\mathbf{B} = [\mathbf{B}_1, \dots, \mathbf{B}_L] \in \mathbb{C}^{N \times Lk}$ with $\mathbf{B}_l = \mathbf{H}_l \mathbf{A}_l \in \mathbb{C}^{N \times k}$ and $\mathbf{v} \in \mathbb{C}^N$ is additive Gaussian noise.

FC employs LMMSE estimator to estimate parameter $\mathbf{s}$ based on the received signal $\mathbf{y}$ in eq. (3.3) [70]

$$\hat{\mathbf{s}} = E[\mathbf{s}\mathbf{y}^H](E[\mathbf{y}\mathbf{y}^H])^{-1}\mathbf{y} \quad (3.4)$$

and the total distortion in terms of MSE can be expressed as

$$D_t = \text{tr}(E[(\mathbf{s} - \hat{\mathbf{s}})(\mathbf{s} - \hat{\mathbf{s}})^H]) \quad (3.5)$$

We can view the total distortion as a summation of noise observation distortion, D_o and channel noise distortion, D_c as follows:

$$\begin{aligned}
D_t &= E[(\bar{\bar{\mathbf{s}}} - \mathbf{s})^2] \\
&= E[(\bar{\bar{\mathbf{s}}} - \bar{\mathbf{s}} + \bar{\mathbf{s}} - \mathbf{s})^2] \\
&= \underbrace{E[(\bar{\bar{\mathbf{s}}} - \bar{\mathbf{s}})^2]}_{:=D_c} + \underbrace{E[(\bar{\mathbf{s}} - \mathbf{s})^2]}_{:=D_o} + \underbrace{2E[(\bar{\bar{\mathbf{s}}} - \bar{\mathbf{s}})(\bar{\mathbf{s}} - \mathbf{s})]}_{:=D_m}
\end{aligned} \tag{3.6}$$

where D_m characterizes the mutual term. The $\bar{\bar{\mathbf{s}}}$ and $\bar{\mathbf{s}}$ are estimated parameters due to channel noise and due to observation noise, respectively. Based on *Cauchy-Schwarz* inequality, we can have the upper bound distortion D_t as follows:

$$\begin{aligned}
D_t &\leq E[(\bar{\bar{\mathbf{s}}} - \bar{\mathbf{s}})^2] + E[(\bar{\mathbf{s}} - \mathbf{s})^2] + 2\sqrt{E[(\bar{\bar{\mathbf{s}}} - \bar{\mathbf{s}})^2(\bar{\mathbf{s}} - \mathbf{s})^2]} \\
&= (\sqrt{D_c} + \sqrt{D_o})^2 \leq 2(D_c + D_o) = 2\tilde{D}_t.
\end{aligned} \tag{3.7}$$

To obtain eq. (3.7), we have used $(E[xy])^2 \leq E[x^2]E[y^2]$ in the first inequality and $(\sum_{k=1}^K G_k)^2 \leq K \sum_{k=1}^K G_k^2$ in the second inequality. Then, we have

$$\tilde{D}_t = D_c + D_o. \tag{3.8}$$

We can calculate distortion due to noisy observation, D_o , by assuming channel noise is free, $\mathbf{v} = 0$. Then, the received signal can be written as

$$\mathbf{y} = \sum_{l=1}^L \mathbf{H}_l \mathbf{A}_l \mathbf{x}_l = \mathbf{B}(\mathbf{F}\mathbf{s} + \mathbf{n}). \tag{3.9}$$

Based on eq. (3.9), we can derive D_o as

$$D_o = \text{tr}(\mathbf{R}_s - \mathbf{R}_{sx} \mathbf{B}^H [\mathbf{B} \mathbf{R}_{xo} \mathbf{B}^H]^{-1} \mathbf{B} \mathbf{R}_{sx}^H) \quad (3.10)$$

where $\mathbf{R}_{xo} = \sigma_s^2 \mathbf{F} \mathbf{r} \mathbf{F}^H + \mathbf{R}_{nl}$ with $\mathbf{r}$ being spatial correlation matrix. We assume covariance matrices of the targets, $\mathbf{R}_s = \sigma_s^2 \mathbf{I}_p$, noise observation, $\mathbf{R}_{nl} = \sigma_n^2 \mathbf{I}_k$ and channel noise, $\mathbf{R}_v = \sigma_v^2 \mathbf{I}_N$. Introducing $\mathbf{R}_{sx} = \mathbf{r}^{1/2} \mathbf{R}_s \mathbf{F}^H$ is covariance between $\mathbf{s}$ and $\mathbf{x}$. Moreover, we can calculate distortion due to channel noise, D_c , by assuming observation noise is free, $\mathbf{n}_l = 0$. Then, the received signal can be written as

$$\mathbf{y} = \sum_{l=1}^L \mathbf{H}_l \mathbf{A}_l \mathbf{x}_l + \mathbf{v} = \mathbf{B}(\mathbf{F} \mathbf{s}) + \mathbf{v}. \quad (3.11)$$

Based on eq. (3.11), we can derive D_c as

$$D_c = \text{tr}(\mathbf{R}_s - \mathbf{R}_{sx} \mathbf{B}^H [\mathbf{B} \mathbf{R}_{xc} \mathbf{B}^H + \mathbf{R}_v]^{-1} \mathbf{B} \mathbf{R}_{sx}^H) \quad (3.12)$$

where $\mathbf{R}_{xc} = \sigma_s^2 \mathbf{F} \mathbf{r} \mathbf{F}^H$.

3.2.1 Correlated Sources

We assume that data observed by sensors are spatially correlated. The targets are assumed being in line with spacing distance, d , as shown in Fig. 3.2 [69]. The model of correlation is

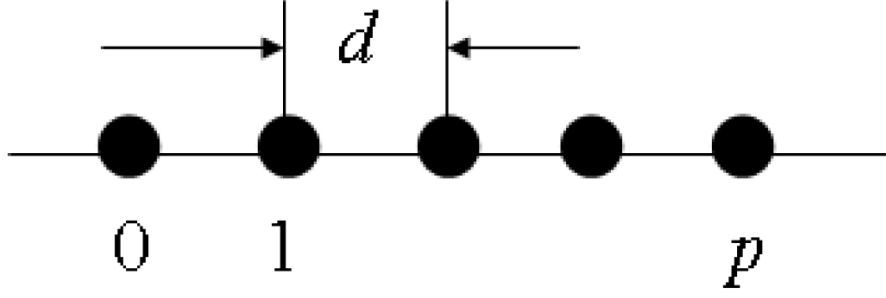


Figure 3.2: The targets are assumed being in line with spacing distance, d .

as follows:

$$\begin{aligned}
 [\mathbf{r}]_{i,j} &= \rho^{|d_{ij}|}; \quad i, j \in 1, \dots, p \\
 &= \begin{bmatrix} \rho^{|d_{11}|} & \dots & \rho^{|d_{1j}|} \\ \vdots & \ddots & \vdots \\ \rho^{|d_{i1}|} & \dots & \rho^{|d_{ij}|} \end{bmatrix}
 \end{aligned} \tag{3.13}$$

where d_{ij} is a distance between the i th and the j th targets, and ρ being spatial correlation coefficient. In general, $\mathbf{r}$ is Hermitian and positive definite, so we can write $\mathbf{r} = \mathbf{r}^{1/2} \mathbf{r}^{1/2}$.

3.2.2 Objective Function

According to the total distortion we need to minimize $\tilde{D}_t$ under a total power constraint, P .

The total transmit power for the L sensors is defined as

$$\sum_{l=1}^L \text{tr}(\mathbf{A}_l \mathbf{R}_{xl} \mathbf{A}_l^H) \leq P \tag{3.14}$$

where $\mathbf{R}_{xl} = \sigma_s^2 \mathbf{F}_l \mathbf{r} \mathbf{F}_l^H + \sigma_n^2 \mathbf{I}_k$.

Furthermore, based on eqs. (3.8) and (3.14), we have an objective function as follows:

$$\begin{aligned} & \min \quad \tilde{D}_t \\ & \text{subject to} \quad \sum_{l=1}^L \text{tr}(\mathbf{A}_l \mathbf{R}_{xl} \mathbf{A}_l^H) \leq P. \end{aligned} \quad (3.15)$$

3.3 Proposed Approach

3.3.1 Proposed Method

In this section, we aim to solve the objective function in eq. (3.15). First, we need to minimize it through SVD technique. Second, we formulate a convex optimization function by considering total power constraint. Afterwards, we further show that optimal solution can be obtained by a water filling algorithm.

Let us introduce a lemma (cf.[70]).

Lemma 3.1: For any two positive semidefinite matrices $\mathbf{X}$ and $\mathbf{Y}$ with size n , it holds that

$$\text{tr}(\mathbf{X}\mathbf{Y}) \leq \sum_{i=1}^n \alpha_i \beta_i \quad (3.16)$$

where α_i and β_i are the i th eigenvalues of $\mathbf{X}$ and $\mathbf{Y}$, respectively, in increasing order.

Since $\mathbf{R}_{xo}$ and $\mathbf{R}_{xc}$ are positive definite, they can be written as $\mathbf{R}_{xo} = \mathbf{R}_{xo}^{1/2} \mathbf{R}_{xo}^{1/2}$ and $\mathbf{R}_{xc} = \mathbf{R}_{xc}^{1/2} \mathbf{R}_{xc}^{1/2}$, respectively. By performing SVD, we can write them as

$$\mathbf{R}_{xo}^{-1/2} \mathbf{F} \mathbf{F}^H \mathbf{R}_{xo}^{-1/2} = \mathbf{U}_f \Sigma_f \mathbf{U}_f^H, \quad (3.17)$$

$$\mathbf{R}_{xc}^{-1/2} \mathbf{F} \mathbf{F}^H \mathbf{R}_{xc}^{-1/2} = \mathbf{U}_g \Sigma_g \mathbf{U}_g^H, \quad (3.18)$$

where $\Sigma \mathbf{f} = \text{diag}(f_1, \dots, f_p)$ with $f_1 \geq f_2 \geq \dots \geq f_p > 0$ and $\Sigma \mathbf{g} = \text{diag}(g_1, \dots, g_p)$ with $g_1 \geq g_2 \geq \dots \geq g_p > 0$. $\mathbf{U}_f \in \mathbb{C}^{Lk \times Lk}$ and $\mathbf{U}_g \in \mathbb{C}^{Lk \times Lk}$ are unitary, respectively.

Let $\mathbf{B}\mathbf{R}_{xo}^{1/2}$ and $\mathbf{B}\mathbf{R}_{xc}^{1/2}$ can be performed as SVD as follows:

$$\mathbf{T} = \mathbf{B}\mathbf{R}_{xo}^{1/2} = \mathbf{U}_t \Sigma \sqrt{\mathbf{t}} \mathbf{V}_t^H \quad (3.19)$$

$$\mathbf{W} = \mathbf{B}\mathbf{R}_{xc}^{1/2} = \mathbf{U}_w \Sigma \sqrt{\mathbf{w}} \mathbf{V}_w^H \quad (3.20)$$

where $\mathbf{U}_t \in \mathbb{C}^{N \times N}$ and $\mathbf{U}_w \in \mathbb{C}^{N \times N}$ are unitary, respectively. Matrix $\Sigma \sqrt{\mathbf{t}} = \text{diag}(\sqrt{t_1}, \dots, \sqrt{t_N})$ with $t_1 \geq \dots \geq t_N > 0$, $\Sigma \sqrt{\mathbf{w}} = \text{diag}(\sqrt{w_1}, \dots, \sqrt{w_N})$ with $w_1 \geq \dots \geq w_N > 0$, $\mathbf{V}_t \in \mathbb{C}^{Lk \times N}$ and $\mathbf{V}_w \in \mathbb{C}^{Lk \times N}$ have orthonormal columns.

Based on eqs. (3.17) and (3.19), the distortion due to observation noise, D_o , can be reformed as

$$\begin{aligned} D_o &= \text{tr}(\sigma_s^2 \mathbf{I}_p - \sigma_s^4 \mathbf{R}_{xo}^{-1/2} \mathbf{F} \mathbf{F}^H \mathbf{R}_{xo}^{-1/2} \mathbf{B}^H \mathbf{R}_{xo}^{1/2} \\ &\quad [\mathbf{B}\mathbf{R}_{xo}^{1/2} \mathbf{B}^H \mathbf{R}_{xo}^{1/2}]^{-1} \mathbf{B}\mathbf{R}_{xo}^{1/2}) \\ &= \text{tr}(\sigma_s^2 \mathbf{I}_p - \sigma_s^4 \mathbf{U}_f \sum \mathbf{f} \mathbf{U}_f^H \mathbf{T}^H [\mathbf{T} \mathbf{T}^H]^{-1} \mathbf{T}) \end{aligned} \quad (3.21)$$

By using *Lemma 3.1*, we have

$$\begin{aligned} D_o &\leq p\sigma_s^2 - \sigma_s^4 \sum_{i=1}^p \frac{f_i b_i}{b_i} \\ &\leq p\sigma_s^2 - \sigma_s^4 \sum_{i=1}^p f_i. \end{aligned} \quad (3.22)$$

According to eqs. (3.18) and (3.20), the distortion due to channel noise, D_c , can be reformed as

$$\begin{aligned}
D_c &= \text{tr}(\sigma_s^2 \mathbf{I}_p - \sigma_s^4 \mathbf{R}_{xc}^{-1/2} \mathbf{F} \mathbf{F}^H \mathbf{R}_{xc}^{-1/2} \mathbf{B}^H \mathbf{R}_{xc}^{1/2} \\
&\quad [\mathbf{B} \mathbf{R}_{xc}^{1/2} \mathbf{B}^H \mathbf{R}_{xc}^{1/2} + \mathbf{R}_v]^{-1} \mathbf{B} \mathbf{R}_{xc}^{1/2}) \\
&= \text{tr}(\sigma_s^2 \mathbf{I}_p - \sigma_s^4 \mathbf{U}_g \sum \mathbf{g} \mathbf{U}_g^H \mathbf{W}^H \\
&\quad [\mathbf{W} \mathbf{W}^H + \sigma_v^2 \mathbf{I}_N]^{-1} \mathbf{W}).
\end{aligned} \tag{3.23}$$

By using *Lemma 3.1*, we have

$$D_c \leq p\sigma_s^2 - \sigma_s^4 \sum_{i=1}^p \frac{g_i w_i}{w_i + \sigma_v^2}. \tag{3.24}$$

We have $\tilde{D}_t$ in terms of SVD as follows:

$$\begin{aligned}
\tilde{D}_t &\leq D_o + D_c \\
&\leq 2p\sigma_s^2 - \sigma_s^4 \left(\sum_{i=1}^p f_i + \frac{g_i w_i}{w_i + \sigma_v^2} \right).
\end{aligned} \tag{3.25}$$

For simplifying analysis, we set $\mathbf{V}_w^H \hat{\mathbf{U}}_g = [\mathbf{I}_N \quad 0]$ where $\hat{\mathbf{U}}_g = \mathbf{U}_g(:, 1 : N)$. This is intended to keep D_c diagonal. When the number of transmitters are set minimum, $N = p$, we have

$$\mathbf{V}_w = \hat{\mathbf{U}}_g; \quad \text{where} \quad \hat{\mathbf{U}}_g = \mathbf{U}_g(:, 1 : p) \tag{3.26}$$

where $\hat{\mathbf{U}}_g$ is taken from the first p columns of matrix $\mathbf{U}_g$. Then, from eq. (3.26), we can express eq. (3.20) as

$$\begin{aligned}\mathbf{B}\mathbf{R}_{xc}^{1/2} &= \mathbf{U}_w \sum \sqrt{\mathbf{w}} \hat{\mathbf{U}}_g^H \\ \mathbf{B} &= \mathbf{U}_w \sum \sqrt{\mathbf{w}} \underbrace{\hat{\mathbf{U}}_g^H \mathbf{R}_{xc}^{-1/2}}_{\mathbf{U}_R}\end{aligned}\quad (3.27)$$

where $\mathbf{U}_w \in \mathbb{C}^{N \times N}$ is unitary in eq. (3.27), thus we have $\mathbf{B} = \Sigma \sqrt{\mathbf{w}} \mathbf{U}_R$. Because of $\mathbf{B} = [\mathbf{B}_1, \dots, \mathbf{B}_l]$, we can express $\mathbf{B}_l$ as

$$\mathbf{B}_l = \Sigma \sqrt{\mathbf{w}} \mathbf{U}_{Rl} \quad (3.28)$$

where $\mathbf{U}_R = [\mathbf{U}_{R1}, \dots, \mathbf{U}_{RL}]$ with $\mathbf{U}_{Rl} \in \mathbb{C}^{p \times k}$.

Let us reform the total power constraint in terms of SVD. We have $\mathbf{A}_l = \mathbf{H}_l^{-1} \mathbf{B}_l$. We can rewrite the eq. (3.14) by taking into account the eq. (3.28) as follows:

$$\begin{aligned}\sum_{l=1}^L \text{tr}(\mathbf{H}_l^{-1} \mathbf{B}_l \mathbf{R}_{xl} \mathbf{B}_l^H \mathbf{H}_l^{-H}) &\leq P \\ \sum_{l=1}^L \text{tr}(\mathbf{H}_l^{-1} \Sigma \sqrt{\mathbf{w}} \mathbf{U}_{Rl} \mathbf{R}_{xl} \Sigma \sqrt{\mathbf{w}} \mathbf{U}_{Rl}^H \mathbf{H}_l^{-H}) &\leq P \\ \sum_{l=1}^L \text{tr}(\mathbf{H}_l^{-1} \mathbf{U}_{Rl} \mathbf{R}_{xl} \mathbf{U}_{Rl}^H \mathbf{H}_l^{-H} \Sigma \mathbf{w}) &\leq P \\ \sum_{i=1}^p n_i w_i &\leq P\end{aligned}\quad (3.29)$$

where $\sum_{l=1}^L (\mathbf{H}_l^{-1} \mathbf{U}_{Rl} \mathbf{R}_{xl} \mathbf{U}_{Rl}^H \mathbf{H}_l^{-H})$ have diagonal entries $n_i, 1 \leq i \leq p$. We also have $\mathbf{R}_{xl} = \sigma_s^2 \mathbf{F}_l \mathbf{r} \mathbf{F}_l^H + \sigma_n^2 \mathbf{I}_k$.

Form eqs. (3.25) and (3.29), we can express the objective function in terms of SVD as

follows:

$$\begin{aligned}
& \max \quad - \left(\sum_{i=1}^p f_i + \frac{g_i w_i}{w_i + \sigma_v^2} \right) \\
& \text{subject to} \quad \sum_{i=1}^p n_i w_i \leq P, \\
& w_i \geq 0, i = 1, \dots, p
\end{aligned} \tag{3.30}$$

It is a convex problem since the objective function is a linear combination of convex functions and the constraints are linear. To solve eq. (3.30), it can be written as Lagrangian equation as

$$\begin{aligned}
L(\mu, w_i) = & - \left(\sum_{i=1}^p f_i + \frac{g_i w_i}{w_i + \sigma_v^2} \right) \\
& + \mu \left(\sum_{i=1}^p n_i w_i - P \right) - \sum_{i=1}^p u_i w_i.
\end{aligned} \tag{3.31}$$

We can obtain the global optimum by solving the KKT conditions [71]

$$\begin{aligned}
\frac{g_i \sigma_v^2}{(w_i + \sigma_v^2)^2} &= \mu n_i - u_i, \quad i = 1, \dots, p \\
\mu \left(\sum_{i=1}^p n_i w_i - P \right) &= 0 \\
u_i w_i &= 0, \quad i = 1, \dots, p \\
\mu &\geq 0.
\end{aligned} \tag{3.32}$$

The solution of the above can be expressed as follows. For $i = 1, \dots, L$

$$w_i = \begin{cases} \left(\sqrt{\frac{g_i}{\mu n_i \sigma_v^2}} - 1 \right) \sigma_v^2, & \frac{g_i}{n_i \sigma_v^2} \geq \mu \\ 0, & \frac{g_i}{n_i \sigma_v^2} < \mu \end{cases}. \tag{3.33}$$

The parameter μ satisfies

$$\sum_{i=1}^L \left(\sqrt{\frac{g_i}{\mu n_i \sigma_v^2}} - 1 \right)^+ \sigma_v^2 n_i = P \quad (3.34)$$

where $(x)^+ = \max(0, x)$. We solve the parameters w_i and μ using water-filling algorithm [61]:

input: $\mathbf{g}_i = (g_1, \dots, g_p); \mathbf{n}_i = (n_1, \dots, n_p); \sigma_v^2; P$

output: $\mu; m_i, i = 1, \dots, p$

1. Reorder the sequence $\mathbf{t}_i = \frac{\mathbf{g}_i}{\mathbf{n}_i \sigma_v^2}$, in the increasing order, and set $m = 1$

2. **do**

($\mu \leftarrow t_m$

$\tau \leftarrow \sum_{i=m}^p \sqrt{t_i} \sigma_v^2 n_i / P + \sum_{i=m}^p \sigma_v^2 n_i$

$m = m + 1$

) **while**($\mu < \tau$ and $m \leq p$)

3. $\mu \leftarrow \tau$

$\omega_i \leftarrow \sigma_v^2 \left(\sqrt{\frac{g_i}{\mu n_i \sigma_v^2}} - 1 \right)^+$

After we obtain $\Sigma \sqrt{\mathbf{w}} = \text{diag}(\sqrt{w_1}, \dots, \sqrt{w_p})$, we can write coding matrix $\mathbf{A}_l$ as

$$\mathbf{A}_l = \mathbf{H}_l^{-1} \Sigma \sqrt{\mathbf{w}} \mathbf{U}_{Rl} \quad (3.35)$$

We also define a lower bound from eq. (3.25), as $P \rightarrow \infty$, we have $w_i \rightarrow \infty$ as

$$D_{low} = 2p\sigma_s^2 - \sigma_s^4 \left(\sum_{i=1}^p f_i + g_i \right) \quad (3.36)$$

3.3.2 Equal Power Method

Under equal power strategy, the transmit power for all sensors are set to be equal as follows:

$$P_l = \frac{P}{L}, \quad 1 \leq l \leq L \quad (3.37)$$

where P_l is the transmit power for the l th sensor. Then, we have the l th coding matrix, $\mathbf{A}_l = \lambda_l [\mathbf{I}_N \ 0]$ where $\lambda_l = \sqrt{P/[L\text{tr}(\mathbf{R}_{xl}(1:N, 1:N))]}$, so that $\text{tr}(\mathbf{A}_l \mathbf{R}_{xl} \mathbf{A}_l^H) = P_l$, $1 \leq l \leq L$. The $\mathbf{R}_{xl}(1:N, 1:N)$ denotes the first N rows and columns of $\mathbf{R}_{xl}$ and $[\mathbf{I}_N \ 0]$ is a $N \times k$ matrix with its diagonal entries equal to 1 and other entries equal to 0.

3.4 Simulation Results

We present simulation results to illustrate the estimation performance of the previous section. We use P to denote the total transmit power constraint across the network. In all simulations, the random vectors $\mathbf{s}$, $\mathbf{n}_l$ and $\mathbf{v}$ are complex Gaussian with zero mean and unit variance. Note that power P is taken relative to the channel noise power. Since the channel noise has unitary variance, thus we label the total transmit power in unit of dB. The channel matrix $\mathbf{H}_l$ is also Gaussian random variable with zero mean and unit variance. We set the number of encoding matrix, $N = 5$, equal to that of sources, $p = 5$. This is because distortion performance does not degrade when $N > p$ [10]. Assuming that targets are close to each other so that we set distance among them, $d = 1$. The channels between sensors and the fusion center are chosen to be independent and the average LMMSE is calculated over 500 times. Therefore, we calculate the total distortion, $\tilde{D}_t$, instead of D_t due to mathematical tractability.

Fig. 3.3 plots distortion in terms of average MSE performance comparison between the proposed method, the lower bound, and the equal power method, in which we take

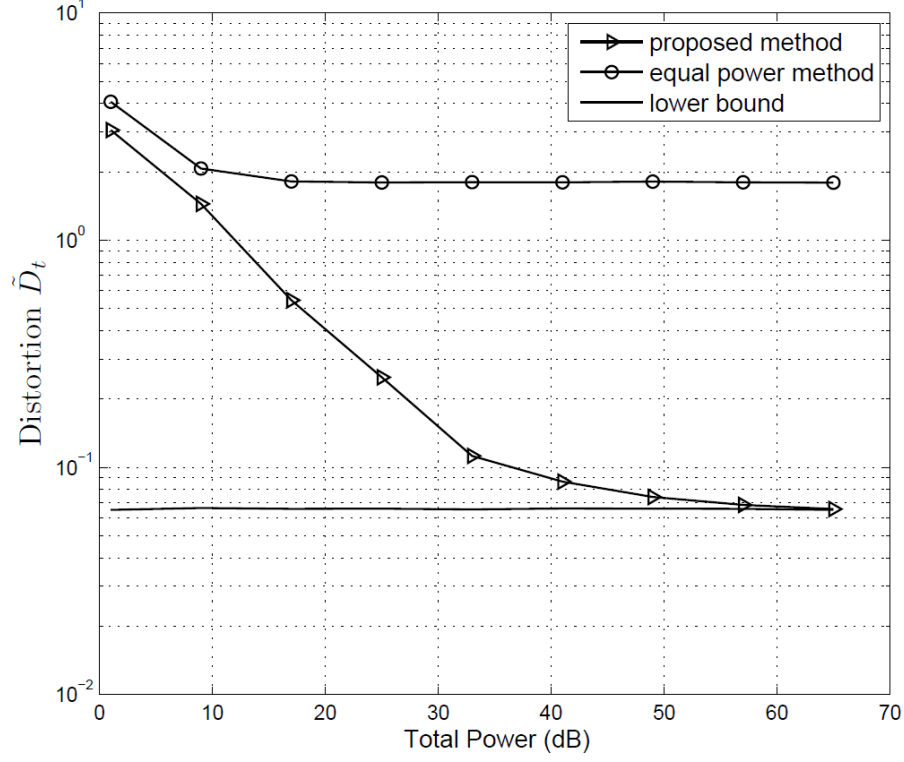


Figure 3.3: Distortion comparison between the proposed method, equal power method and the lower bound with parameters: $\rho = 0$, $L = 10$, $k = 8$, $N = 5$, $p = 5$.

$L = 10$, $k = 8$ and the correlation coefficient, $\rho = 0$. From Fig. 3.3, we can see that the proposed method performs better than the equal power method. The distortion of the proposed method becomes close to the lower bound. Moreover, the gap of distortion between the proposed method and the equal power method becomes larger as P increases. This is because the proposed method allocates power by taking into account the effects of observation and channel noise, while the equal power method does not.

Fig. 3.4 shows distortion for uncorrelated, $\rho = 0$ and correlated, $\rho = 0.9$ data. We can see that the distortion of both methods becomes worse as the data being correlated. If the data is more correlated, sensors provide redundant information about the targets. It becomes difficult to estimate the targets. Once $\rho = 0.9$, the distortion of the proposed method

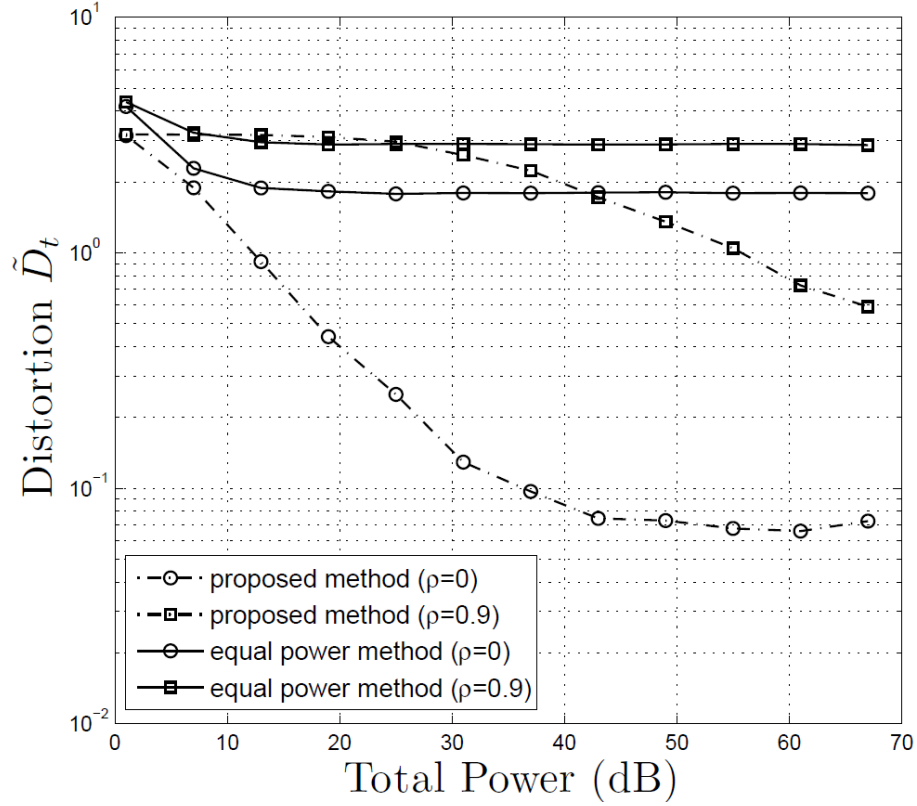


Figure 3.4: Distortion comparison between difference correlation coefficients with parameters: $\rho = 0, 0.9, L = 10, k = 8, N = 5, p = 5$.

remains constant for $P < 25$ dB. Moreover, the equal power method performs better than the proposed method particularly for P between 10–25 dB, because the proposed method does not have enough power to accomplish the threshold of water-filling. However, the distortion of the proposed method becomes smaller than that of the equal power method for $P > 25$ dB, because it allocates power by taking into account the effects of the correlation, observation, and channel noise.

The distortion of the proposed method and the lower bound with different power levels ($P = 5, 10$, and 20 dB) and spatial correlation coefficient, $\rho = 0.7$ are shown in Fig. 3.5. Both distortions become smaller as the number of measurements, k increases. This is because

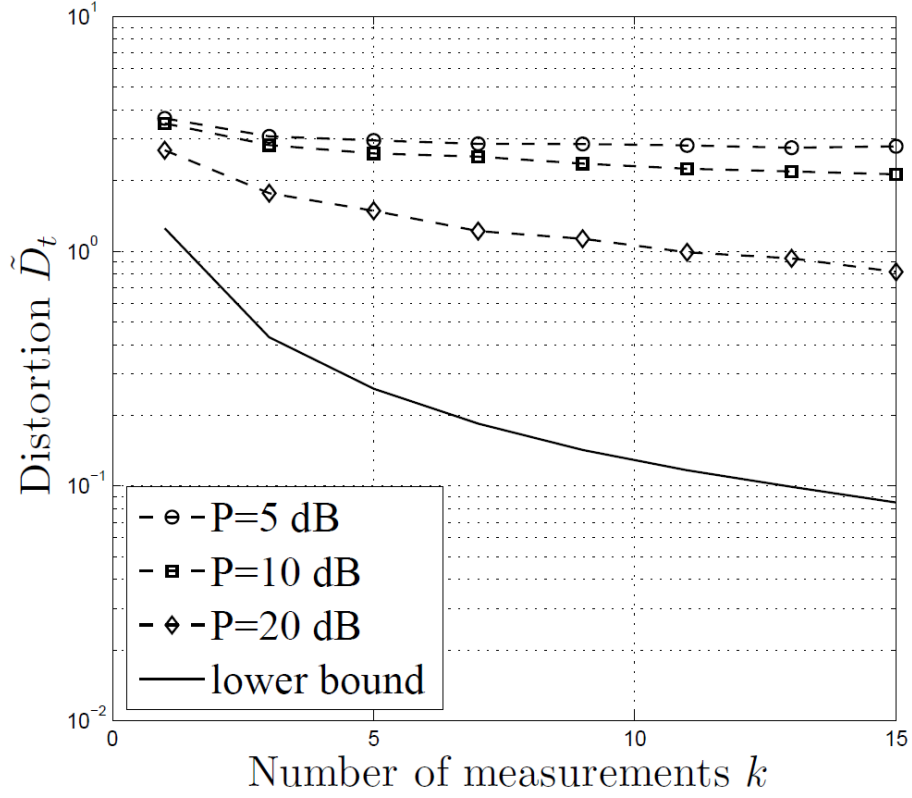


Figure 3.5: Distortion versus k with parameters: $\rho = 0.7$, $L = 10$, $N = 5$, $p = 5$, $P = 5, 10, 20$ dB.

increasing the number of k leads to increase of measurement power. In Fig. 3.5, we can also see that the gap of the distortion between the proposed method and the lower bound becomes larger. This is because the proposed method has a power constraint compared with the lower bound that has no power constraint.

From Fig. 3.6, we simulate the proposed method with different power levels ($P = 1, 2, 5, 8, 15, 20$ dB) and correlation coefficient, $\rho = 0.7$. The distortions of the proposed method become smaller as the number of sensors, L increases. However, they rise slightly after reached the optimum number of sensors. For example, the total power constraint is $P = 1$ dB, the distortion rises when the number of sensors is $L > 3$. This is owing to the fact

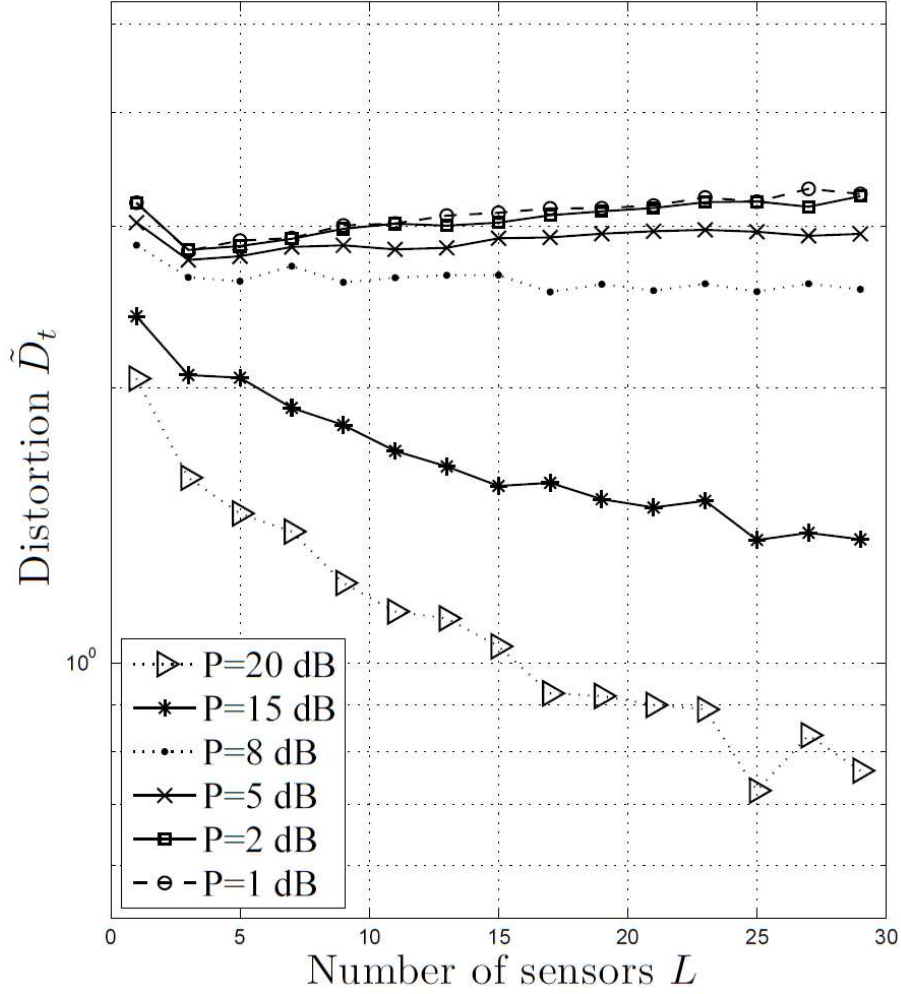


Figure 3.6: Distortion versus L with parameters: $\rho = 0.7, k = 8, N = 5, p = 5, P = 1, 2, 5, 8, 15, 20$ dB.

that the power allocated to each sensor becomes smaller as the increase of the number of sensors.

In Fig. 3.7, we show the simulation results of the proposed method with an ideal sensor condition and an ideal channel condition. The simulation with ideal sensor condition is taken by assuming that the sensor observations have no distortion. To take into account that assumption, we set the distortion due to observation noise, $D_o = 0$, in eq. (3.8). Then,

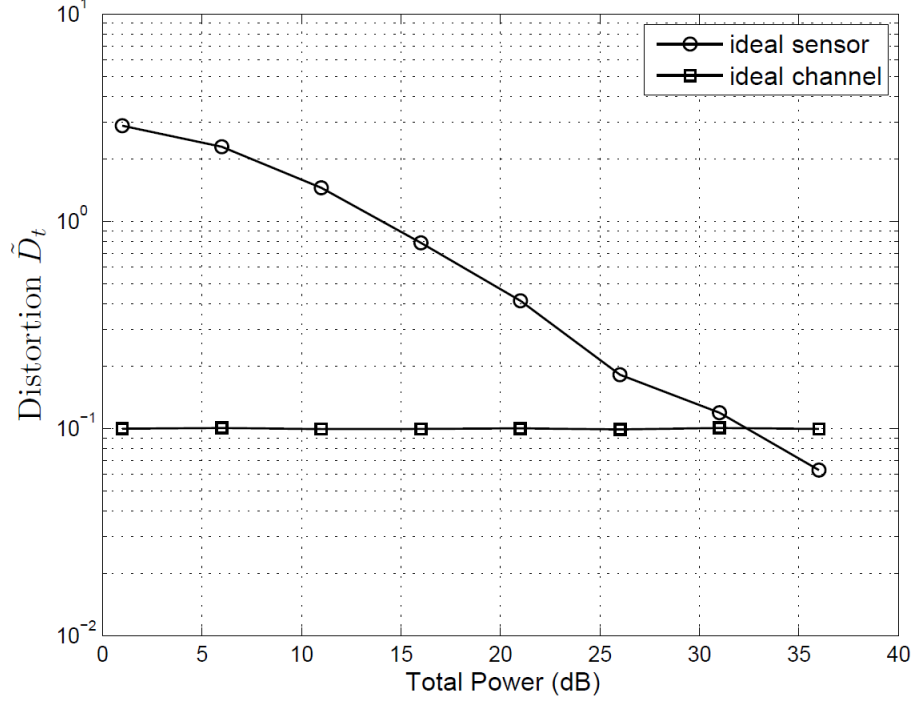


Figure 3.7: Distortion comparison of the proposed method with ideal sensor and ideal channel condition with parameters: $\rho = 0.5, k = 8, N = 5, p = 5$.

the simulation with ideal channel condition is taken by assuming that the fusion center has a perfect channel knowledge between the sensors and the fusion center. In a similar way to the ideal sensor condition, we set the distortion due to channel noise, $D_c = 0$, in eq. (3.8). We use a complex Gaussian random variable with zero mean and unit variance for observation noise, $\mathbf{n}$, and noisy channel, $\mathbf{v}$, respectively. The distortion with ideal sensor becomes smaller as P increases, while the distortion with ideal channel still remains constant. It means that the water-filling method is more effective to combat the noisy channel rather than the observation noise.

3.5 Conclusion

We studied distributed estimation of a random vector in MIMO sensor network by considering total power constraint and spatial correlated data. For the spatial correlated data, we obtained a water-filling based closed-form solution that follows water-filling strategy. We also derived lower bound of distortion to this system. From the simulation results, we showed that the distortion increases as data becomes more correlated, because the sensors become difficult to estimate the targets. Moreover, we showed that the noisy channel was more harmful than the observation noise. By taking into account the effects of correlation, observation, and channel noise, the proposed method has a better distortion performance than the equal power method.

Chapter 4

Capacity of Data Collection in Wireless Sensor Networks based on Mutual Information and MMSE Estimation

4.1 Introduction

In Chapter 3, we discuss the distributed estimation which can be classified as the inference problem. The inference problem points out how to measure the quality of data observed by sensors and arrived at the destination. We also can compute how much the data are distorted in terms of MSE. In this Chapter, we investigate the communication problem in terms of capacity analysis. It is necessary, because the capacity measures the achievable data rate over noisy channel. Capacity also reflects how fast the data are collected by FC.

The literatures In Chapter 2, the capacity analysis has been investigated on either the

physical model [6], [7] or the protocol model [73], [74]. The physical model takes into account such as time invariance and non-deterministic channels. The capacity derivation is very complex including antenna sharing, channel coding, and beamforming. The second model is the protocol model, which makes simplified assumptions to the link capacity. However, the protocol model is relatively inaccurate, when simultaneous transmissions are allowed in the network [5], [59]. Moreover, the protocol model does not capture physical characteristics such as time invariant and non-deterministic channels.

In this Chapter, we propose an alternative derivation regarding the capacity. The derivation is based on the relationship between mutual information and MMSE. We follow [4] for naming of the capacity as the capacity of data collection because, in fact, FC collects the data from the sensors. We focus on deriving capacity of data collection for random networks under coherent and orthogonal MAC scenario. We provide a new perspective of capacity calculation of data collection in WSNs that can be derived from error estimation of the target at the FC. The relationship between mutual information and MMSE has been revealed by Guo in [54], [55], [87]–[93]. First, we derive a capacity formulation on coherent MAC model. In the coherent MAC model, we assume that there is perfect synchronization between sensors and the fusion center so that the transmitted messages from local sensors can be coherently combined at the fusion center. With such an assumption, one key design consideration at local sensors and the fusion center is how to jointly process the sensed and received information in terms of capacity. We write a problem formulation for maximizing the capacity and then solve it through convex optimization technique. We derive the optimal power allocation strategy to maximize the capacity. The upper bound on the capacity of data collection with coherent MAC model is also derived as a benchmark. Second, we derive a capacity formulation on orthogonal MAC scenario. The motivation for using orthogonal multiple access schemes such as Frequency Division Multiple Access (FDMA), Time Divi-

sion Multiple Access (TDMA), or Code Division Multiple Access (CDMA), is the removal of the requirement on the carrier level synchronization among sensors [40]–[43], [53]. As the coherent MAC model, we also derive optimal power allocation strategy for the case where the capacity is maximized under certain power constraints. In the orthogonal model, the optimal power allocation is achieved by turning off certain sensors with bad channels and bad observation quality. The upper bound on this model is also derived and interestingly equal to the upper bound on coherent one.

The rest of this chapter is organized as follows. Section 4.2 describes the preliminary theory and system model. In Section 4.3, we formulate the capacity of data collection for the upper-bound, equal power allocation and optimal power allocation in coherent MAC model. Section 4.4, we formulate the capacity of data collection for the equal power allocation, optimal power allocation and the upper bound on orthogonal MAC model. Section 4.5 presents some simulation results and conclusion is drawn in Section 4.6.

4.2 Problem Formulation

As a preliminary, we start by explaining the relationship between mutual information and MMSE [87].

4.2.1 Capacity of the Gaussian Channel based I-MMSE approach

An input-output model can be written as

$$Y = \sqrt{\text{snr}}X + N \tag{4.1}$$

where $N \sim \mathcal{N}(0, 1)$ is standard Gaussian. We note here that snr in (4.1) coincides with the usual notation of signal-to-noise power ratio (SNR) only if $E[X^2] = 1$. Then, we refer to snr as SNR regardless of the input power. The MMSE of estimating the input X of the model given the noisy output Y can be denoted by

$$\begin{aligned} \text{mmse}(X, \text{snr}) &= \text{mmse}(X | \sqrt{\text{snr}} + N) \\ &= E[(X - E[X | \sqrt{\text{snr}} + N])^2] \end{aligned} \quad (4.2)$$

where $E[.]$ is the expected value.

The MMSE can be regarded as a function of SNR for every given distribution P_X . In particular, if $X \sim \mathcal{N}(m, \sigma_X^2)$, the MMSE can be denoted as

$$\text{mmse}(X, \text{snr}) = \frac{\sigma_X^2}{1 + \sigma_X^2 \text{snr}}. \quad (4.3)$$

Moreover, simple quantitative connections between MMSE and information measures are revealed in [87]. One of the results is

$$2 \frac{d}{d\text{snr}} I(X; Y) = \text{mmse}(X, \text{snr}) \quad (4.4)$$

for every $\text{snr} \geq 0$. The corresponding capacity of the model is

$$\begin{aligned} C &= \max I(X; Y) \\ &= \frac{1}{2} \int_0^{\text{snr}} \text{mmse}(X, \text{snr}) d\text{snr} \\ &= \frac{1}{2} \log(1 + \sigma_X^2 \text{snr}) \end{aligned} \quad (4.5)$$

where we adopt natural logarithms and use nats as the unit of all capacity measures.

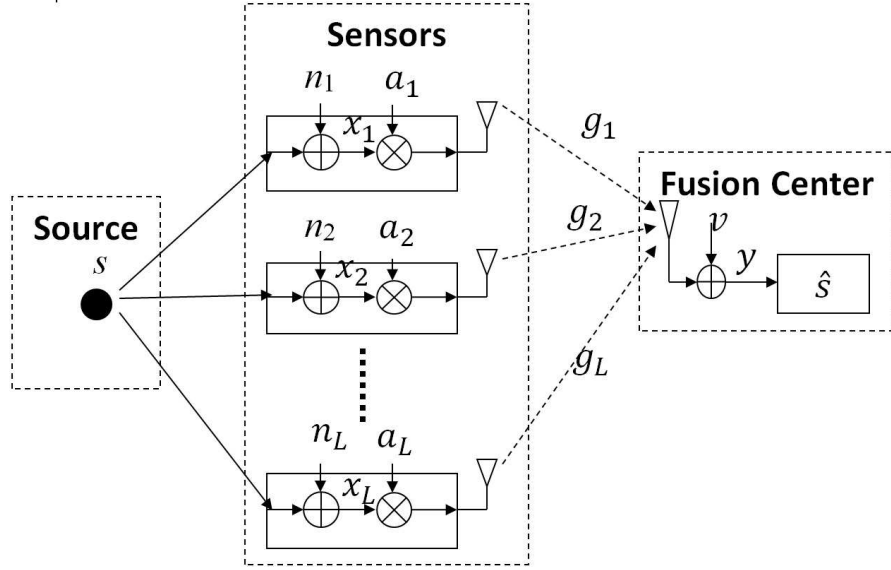


Figure 4.1: System model of the capacity of data collection in WSNs.

4.2.2 System Model

Suppose that there are L sensors, each making observation on a common unknown parameter s as Fig. 4.1. The sensors observe s with noisy observation n that has zero mean and variance, $\sigma_{n_i}^2$. We assume the sensor and FC communicate with coherent MAC. When source and observation are scalars, the observation model can be written as

$$x_i = s + n_i, \quad 1 \leq i \leq L. \quad (4.6)$$

Suppose that the corresponding analog amplifying and forwarding scheme is used, we have a power amplification factor a_i of i th sensor. The average transmit power of sensor i is

$$\begin{aligned} P_i &= a_i^2(\sigma_s^2 + \sigma_{n_i}^2) \\ &= a_i^2(1 + \alpha_i^{-1}) \end{aligned} \quad (4.7)$$

where we assume that $\sigma_s^2 = 1$. $\alpha_i = 1/\sigma_{n_i}^2$ is SNR observation of sensor i . After amplification, signals are transmitted to the FC. The received signal at FC is

$$\begin{aligned} y &= \sum_{i=1}^L g_i a_i x_i + v \\ &= \sum_{i=1}^L g_i a_i s + \sum_{i=1}^L g_i a_i n_i + v \end{aligned} \quad (4.8)$$

where g_i and v are channel gain and channel noise, respectively. Similarly, v is assumed to have zero mean and unit variance, σ_v^2 . The linear MMSE estimator of s from y is $\hat{s} = (E[sy]/E[y^2])y$, which has MSE, J , satisfying [70]

$$\frac{1}{J} = 1 + \left(1 + \sum_{i=1}^L g_i^2 a_i^2\right)^{-1} \left(\sum_{i=1}^L g_i a_i \sqrt{\alpha_i}\right)^2. \quad (4.9)$$

4.3 Capacity of Data Collection in Coherent MAC Model

4.3.1 Upper Bound on Capacity of Data Collection

Note that when all sensor observations $x_i = [x_1, \dots, x_L]^T$ are directly available to the FC, a centralized estimator $\hat{s}_0 = (E[sx]/E[x^2])x$, achieves an MSE, J_0 , as

$$\begin{aligned} \frac{1}{J_0} &= 1 + \sum_{i=1}^L \alpha_i \\ &= 1 + \gamma \end{aligned} \quad (4.10)$$

where α_i and γ are SNR observation of sensor i and the total SNR observation, respectively. Analytical proof is also available in Appendix A. Applying eq. (4.5), we can express the

capacity of data collection as follows:

$$\begin{aligned}
C_{\text{upp1}} &= \frac{1}{2} \int_0^{\text{SNR}_{\text{tot}}} J_0 d\gamma \\
&= \frac{1}{2} \int_0^{\text{SNR}_{\text{tot}}} \frac{1}{1+\gamma} d\gamma \\
&= \frac{1}{2} \log(1 + \text{SNR}_{\text{tot}}).
\end{aligned} \tag{4.11}$$

Because of the randomness of sensors deployment, we assume that noisy observation becomes i.i.d., with $\sigma_{n_1}^2, \dots, \sigma_{n_L}^2 = \sigma_n^2$. Then, the upper bound of the capacity of data collection can be expressed as

$$C_{\text{upp1}} = \frac{1}{2} \log \left(1 + \frac{L}{\sigma_n^2} \right). \tag{4.12}$$

Without loss of generality, we can express the capacity of the network scaled by $\Theta \left(\frac{1}{2} \log(1 + L) \right)$ as the number of sensors becomes infinity, $L \rightarrow \infty$.

4.3.2 Capacity of Data Collection for Equal Power Allocation

Suppose all sensors use the same transmit power, $P_i = P/L$ where P is the total transmit power. From (4.7), we get $a_i = \sqrt{P/L(\alpha_i^{-1} + 1)}$. Let $J_{\text{eq1}}(P)$ denote the achieved MSE with equal transmit power. From (4.9), $J_{\text{eq1}}(P)$ satisfies

$$\begin{aligned}
\frac{1}{J_{\text{eq1}}(P)} &= 1 + \left(\frac{L}{P} + \sum_{i=1}^L \frac{g_i^2}{\alpha_i^{-1} + 1} \right)^{-1} \\
&\quad \times \left(\sum_{i=1}^L \sqrt{\frac{\alpha_i}{\alpha_i^{-1} + 1}} g_i \right)^2.
\end{aligned} \tag{4.13}$$

With the same analogy in eq. (4.10), we define $\left(\frac{L}{P} + \sum_{i=1}^L \frac{g_i^2}{\alpha_i^{-1} + 1}\right)^{-1} \left(\sum_{i=1}^L \sqrt{\frac{\alpha_i}{\alpha_i^{-1} + 1}} g_i\right)^2$ as a total SNR of the system. Therefore, we can express the capacity of equal transmit power as follows:

$$\begin{aligned} C_{\text{eq1}}(P) &= \frac{1}{2} \log \left(1 + \left(\frac{L}{P} + \sum_{i=1}^L \frac{g_i^2}{\alpha_i^{-1} + 1} \right)^{-1} \right. \\ &\quad \left. \times \left(\sum_{i=1}^L \sqrt{\frac{\alpha_i}{\alpha_i^{-1} + 1}} g_i \right)^2 \right). \end{aligned} \quad (4.14)$$

For $P \rightarrow \infty$, we can write the capacity as

$$\begin{aligned} C_{\text{eq1}}(\infty) &= \frac{1}{2} \log \left(1 + \left(\sum_{i=1}^L \frac{g_i^2}{\alpha_i^{-1} + 1} \right)^{-1} \right. \\ &\quad \left. \times \left(\sum_{i=1}^L \sqrt{\frac{\alpha_i}{\alpha_i^{-1} + 1}} g_i \right)^2 \right) \\ &\leq C_{\text{upp1}}. \end{aligned} \quad (4.15)$$

We can summarize the above results that the capacity achieved by analog forwarding scheme where each sensor uses exactly the same transmit power of P/L . We can express for every finite P as

$$C_{\text{upp1}} \geq C_{\text{eq1}}(\infty) \geq C_{\text{eq1}}(P). \quad (4.16)$$

4.3.3 Capacity of Data Collection for Optimal Power Allocation

Here, we consider an optimal power allocation whereby the transmit power is optimally allocated among the sensors to achieve the maximum capacity. From the right hand side

(RHS) of eq. (4.9), we denote $\beta = \left(1 + \sum_{i=1}^L g_i^2 a_i^2\right)^{-1} \left(\sum_{i=1}^L g_i a_i \sqrt{\alpha_i}\right)^2$ as a total SNR. Therefore, we can easily express the capacity as follows:

$$\begin{aligned}
C &= \frac{1}{2} \log \left(1 + \left(1 + \sum_{i=1}^L g_i^2 a_i^2 \right)^{-1} \left(\sum_{i=1}^L g_i a_i \sqrt{\alpha_i} \right)^2 \right) \\
&= \frac{1}{2} \log (1 + \beta).
\end{aligned} \tag{4.17}$$

Let $C(P_1, \dots, P_L)$ denote the capacity achieved by optimally assigning P_i to sensor i . Maximizing the capacity under a sum power constraint can be written as

$$\begin{aligned}
&\max_{P_i; 1 \leq i \leq L} && C(P_1, \dots, P_L) \\
&\text{s.t.} && \sum_{i=1}^L a_i^2 (\alpha_i^{-1} + 1) \leq P, \\
&&& a_i (1 + \alpha_i^{-1}) \leq P_i^{\max} \quad 1 \leq i \leq L.
\end{aligned} \tag{4.18}$$

Maximizing capacity in eq. (4.18) is equivalent to maximize the total SNR as follows:

$$\begin{aligned}
&\max_{a_i^2; 1 \leq i \leq L} && \beta(a_1^2, \dots, a_L^2) \\
&\text{s.t.} && \sum_{i=1}^L a_i^2 (\alpha_i^{-1} + 1) \leq P, \\
&&& a_i (1 + \alpha_i^{-1}) \leq P_i^{\max} \quad 1 \leq i \leq L.
\end{aligned} \tag{4.19}$$

With the aid of Appendix B that follows the solution in [10], we get the best achievable

total SNR as

$$\beta_{\text{opt}} = \sum_{i=1}^L \frac{\alpha_i}{1 + (\alpha_i^{-1} + 1)/(g_i^2 P)}. \quad (4.20)$$

The optimal power allocation achieving the optimal total SNR is

$$P_i^{\text{opt}} = c_i P, \quad 1 \leq i \leq L \quad (4.21)$$

where

$$c_i = c \frac{g_i^2 \alpha_i^2 (\alpha_i^{-1} + 1)}{(\alpha_i^{-1} + 1 + g_i^2 P)^2}$$

and

$$c = \left(\sum_{i=1}^L \frac{g_i^2 \alpha_i^2 (\alpha_i^{-1} + 1)}{(\alpha_i^{-1} + 1 + g_i^2 P)^2} \right)^{-1}.$$

Implementing optimal power allocation, we need the FC to broadcast the constant c and P to the sensors. The sensors use c , P , and two local parameters g_i^2 , α_i to determine their individual transmit power.

Therefore, we can express the optimal capacity of data collection as

$$\begin{aligned} C_{\text{opt1}} &= \frac{1}{2} \log(1 + \beta_{\text{opt}}) \\ &= \frac{1}{2} \log \left(1 + \sum_{i=1}^L \frac{\alpha_i}{1 + (\alpha_i^{-1} + 1)/(g_i^2 P)} \right). \end{aligned} \quad (4.22)$$

4.4 Capacity of Data Collection in Orthogonal MAC Model

In this section, we adopt orthogonal channels between the sensors and the FC. We assume that the observed signal is analog and the observation noises are uncorrelated across sensors. In addition, we assume that the second moments of the signal and noise are known to the corresponding sensor and the FC. The FC deploys the MMSE estimator to generate estimates of the unknown signal. In this setting, we use an analog transmission system where observations are amplified and forwarded to the FC.

Suppose that the received signal of orthogonal MAC from sensor i to FC can be written as

$$y_i = g_i a_i s + g_i a_i n_i + v_i \quad (4.23)$$

where v_i and g_i are the channel noise with zero mean and unit variance of channel i and channel gain, respectively. For MMSE estimation, we can get an MSE, J [10] as

$$\frac{1}{J} = 1 + \sum_{i=1}^L \frac{g_i^2 a_i^2 \alpha_i}{1 + g_i^2 a_i^2}. \quad (4.24)$$

4.4.1 Capacity of Data Collection for Equal Power Allocation

For equal power method, $P_i = P/L$, thus we have $a_i^2 = P_i/(1 + \alpha_i^{-1})$. By changing the form of eq. (4.24), we get

$$\frac{1}{J} = 1 + \sum_{i=1}^L \frac{\alpha_i}{1 + (1 + \alpha_i^{-1})/(g_i^2 P_i)}. \quad (4.25)$$

Following the expression of eq. (4.5) and $\sum_{i=1}^L \frac{\alpha_i}{1+(1+\alpha_i^{-1})/(g_i^2 P_i)}$ as a total SNR of the system, we can write the capacity as

$$C_{\text{eq2}} = \frac{1}{2} \log \left(1 + \sum_{i=1}^L \frac{\alpha_i}{1 + (1 + \alpha_i^{-1})/(g_i^2 P_i)} \right). \quad (4.26)$$

For $P_i \rightarrow \infty$, we can write an upper bound on the capacity of data collection in orthogonal MAC as

$$\begin{aligned} C_{\text{upp2}} &= \frac{1}{2} \log \left(1 + \sum_{i=1}^L \alpha_i \right) \\ &= \frac{1}{2} \log \left(1 + \frac{L}{\sigma_n^2} \right) \\ &= C_{\text{upp1}}. \end{aligned} \quad (4.27)$$

Interestingly, we can see that the upper bound on capacity of data collection for orthogonal MAC and that for the coherent MAC are equal.

4.4.2 Capacity of Data Collection for Optimal Power Allocation

To maximize the capacity of data collection on orthogonal MAC under optimal power method, first, we need to minimize the MSE under total power constraint, P . The MSE for optimal power method of the orthogonal MAC for the case of scalar source and observations is given in [10]

$$\frac{1}{J} = 1 + \sum_{i=1}^{L_1} \alpha_i \left(1 - \frac{\sqrt{1 + \alpha_i^{-1}}}{c_0 g_i} \right). \quad (4.28)$$

where c_0 and L_1 are the threshold of $\frac{g_i^2}{1+\alpha_i^{-1}} \geq 1/c_0^2$ whether a sensor transmits or keeps silent and the number of active sensors, respectively. The threshold c_0 is defined by

$$c_0 = \frac{\left(\sum_{m=1}^{L_1} \frac{\alpha_m(\alpha_m^{-1}+1)}{g_m^2} + P \right)}{\left(\sum_{m=1}^{L_1} \frac{\alpha_m \sqrt{\alpha_m^{-1}+1}}{\sqrt{g_m^2}} \right)}$$

Following eq. (4.5), we can express the capacity as,

$$C_{\text{opt2}} = \frac{1}{2} \log \left(1 + \sum_{i=1}^{L_1} \alpha_i \left(1 - \frac{\sqrt{1 + \alpha_i^{-1}}}{c_0 g_i} \right) \right). \quad (4.29)$$

We note that the optimal power method for orthogonal MAC will allocate most of power to sensors that have good observation and channel qualities. Hence, the active sensors are sensors that have good observation and channel qualities.

4.5 Simulation Results

In Fig. 4.2, we plot the curves of capacity of data collection for coherent MAC model versus total transmit power P in dB (relative to the channel noise power) with the number of sensors $L = 10$. In the simulation, sensor observation noise variance is set $\sigma_n^2 = 0.5$. The channel gains, g_i , are taken as $c_g * d^{-\epsilon}$ where d is uniformly taken from real interval $[1, 10]$ and ϵ is a path loss parameter that we assume $\epsilon = 2$. Parameter c_g is a normalization constant to make $E(g_i) = 1$. Simulations are averaged over 5000 realizations. Those parameters are also used in all simulations. For coherent MAC model in Fig. 4.2, we can see that when P increases, equal power method and optimal power method converge to two different limits that are $C_{\text{eq1}}(\infty)$ and C_{upp1} , respectively. This is because the optimal power method allocates power by taking into account channel gain sensor observation while the equal power method does

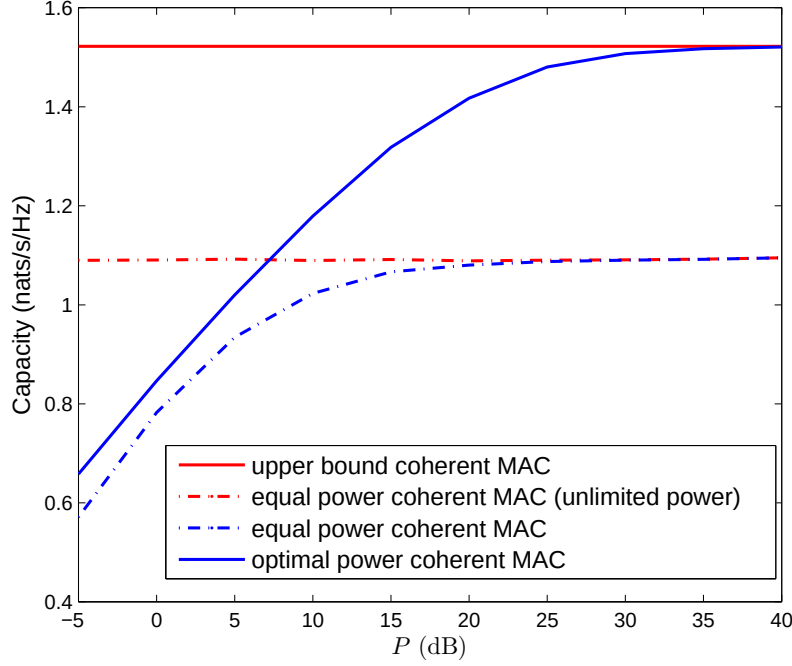


Figure 4.2: Capacity of data collection for equal power method versus optimal power method in coherent MAC as P increases. Note that power P is taken relative to the channel noise power. Since we assume that the channel noise has unitary variance, thus we label the total transmit power in unit of dB.

not. Moreover, the limit of the equal power method, $C_{\text{eq1}}(\infty)$, is due to inhomogeneous sensing environment.

We can see in Fig. 4.3 for orthogonal MAC model that the capacity of the optimal power method is larger than that of the equal power method. This is because the optimal power method allocates most of power to sensors that have good observation and channel qualities. Moreover, as P increases, both the optimal method and the equal one converge to the upper bound. In high power regime, each sensor has a redundant power to transmit the sensing data and can easily combat the channel noise.

In Fig. 4.4, we compare the optimal power method for both MAC models. We can see

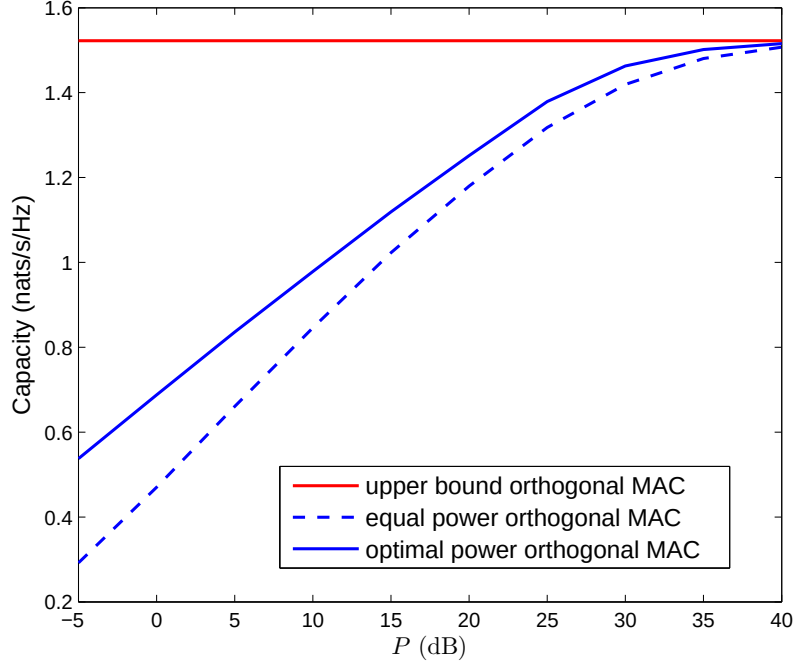


Figure 4.3: Capacity of data collection for equal power method versus optimal power method in orthogonal MAC as P increases. Note that power P is taken relative to the channel noise power. Since we assume that the channel noise has unitary variance, thus we label the total transmit power in unit of dB.

that the optimal power method for coherent MAC outperforms the orthogonal MAC. This is a consequence of using orthogonal links that have L different channel noises. We also compare the equal method of both MAC in Fig. 4.5. In high power regime ($P > 15$ dB), the equal method for orthogonal has larger capacity because the coherent MAC is limited by the finite number of sensor observations.

We calculate the capacity versus the number of sensors L with the total power being constant at $P = 20$ dB (relative to channel noise variance) for both models in Fig. 4.6. The capacity of both models increases as the total number of sensors increases. This is because as the number of sensors increases the total SNR also increases. However, we can see that

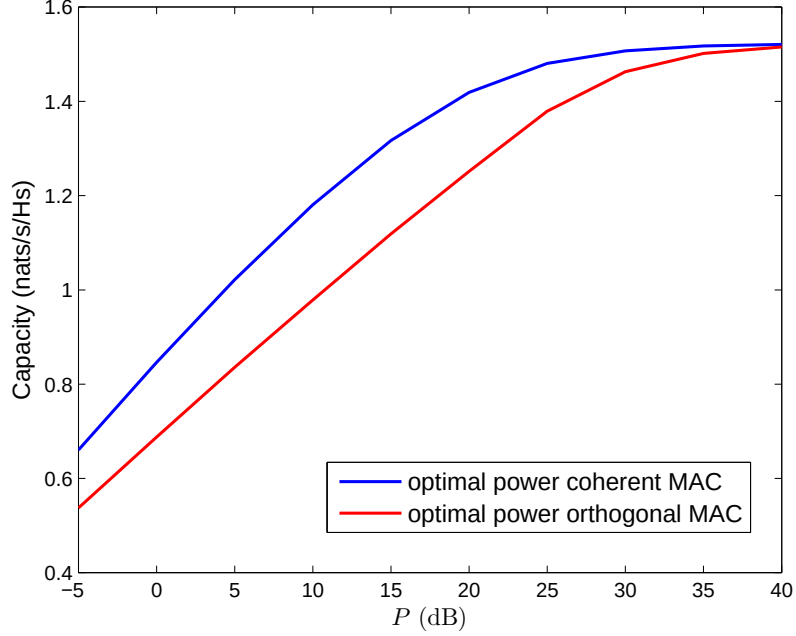


Figure 4.4: Comparison between the capacity of data collection in coherent MAC and orthogonal MAC for optimal power allocation method.

with this finite total power and a large number of sensors, the capacity of the coherent MAC is larger than that of the orthogonal MAC for both methods, equal and optimal power. This is because the corrupted-channels in orthogonal MAC cannot be eliminated even when L goes to infinity. However, in the coherent MAC model, channel noise incurs only once per reception at FC.

In Fig. 4.7, we plot the percentage of active sensors versus the total transmission power, where we set $L = 100$ in the simulation for optimal method in orthogonal MAC. We note that the number of active sensors is less than L when the total power budget is small. This confirms that the optimal power allocation for orthogonal MAC allocates most of power to only the sensors that have good observation and channel qualities. Activating only the sensors that have good observation and channel qualities can be used to conserve energy of

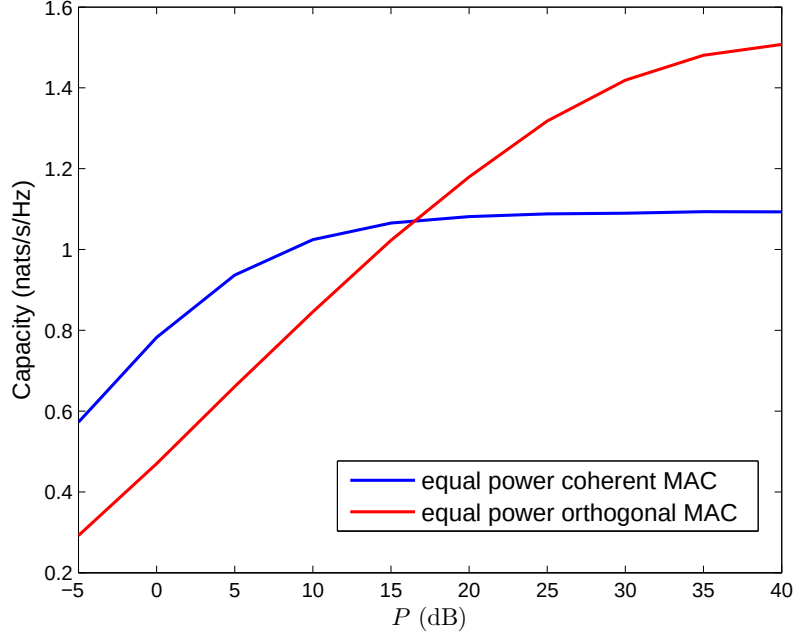


Figure 4.5: Comparison between the capacity of data collection in coherent MAC and orthogonal MAC for equal power allocation method.

the sensors and extend sensor's lifetime.

4.6 Conclusion

We studied the capacity of data collection in wireless sensor networks by considering power allocation strategy. We considered a scenario in which a number of sensors observe a target being estimated at FC using MMSE estimator. Based on the relationship between mutual information and MMSE, we derived the capacity of data collection in both coherent MAC model and orthogonal MAC model. Considering power constraint, we derived the capacity under two scenarios: equal power allocation and optimal power allocation of both models. We also provided the upper bound of capacity as a benchmark. In particular, we showed

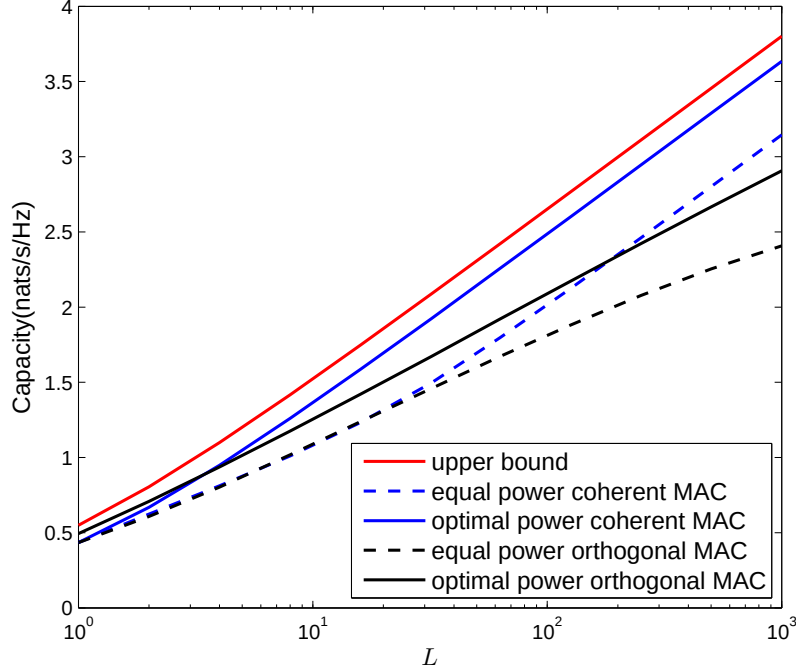


Figure 4.6: Capacity of data collection of equal power method versus optimal power method as L increases for either coherent and orthogonal MAC in a finite power budget, $P = 20$ dB.

that the capacity of data collection scaled as $\Theta\left(\frac{1}{2}\log(1+L)\right)$ when the number of sensors L grows to infinity. We verified by simulation results as follows: 1. For coherent MAC model, we derived the optimal power allocation strategy that maximizes the capacity. The capacity of the optimal power is larger than that of the equal power, because the optimal power method takes into account the SNR observation and channel gain to determine their individual transmit power. 2. For orthogonal MAC model, we derived the optimal power allocation strategy that maximizes the capacity. The capacity of the optimal power is larger than that of the equal power, because the optimal power method allocates most of power to only the sensors that have good SNR observation and channel qualities, while the sensors with bad observation and bad channel qualities will be turned off. Turning off the sensors

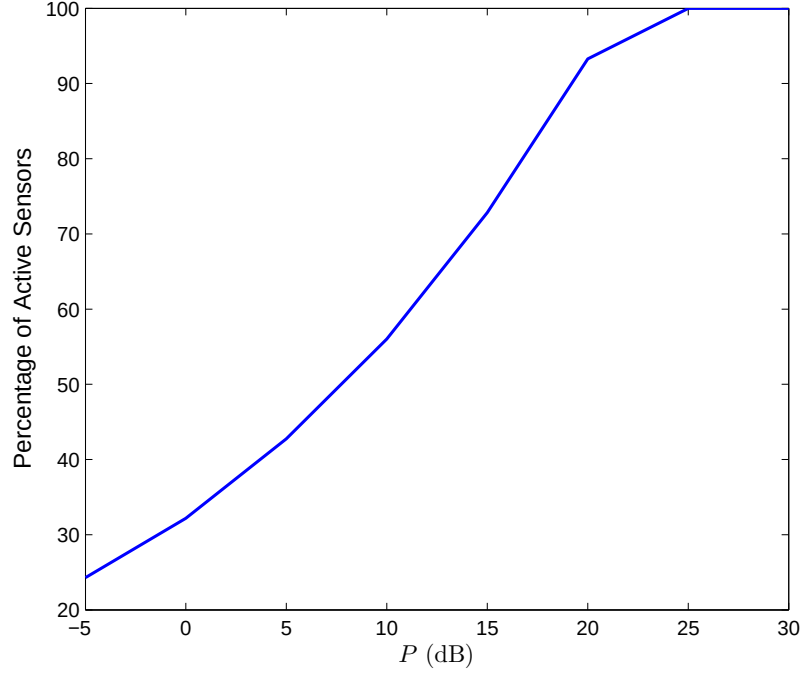


Figure 4.7: Percentage of active sensors as P increases for optimal power method in orthogonal MAC.

with bad observation and bad channel qualities can be used to conserve energy of the sensors and extend sensor's lifetime. Moreover, we showed that the capacity of coherent MAC is larger than that of orthogonal MAC, particularly when the number of sensors L is large and the total power P is fixed. This is a consequence of using orthogonal link from the sensors to FC where the corrupted-channel cannot be eliminated even when L goes to infinity.

Chapter 5

Conclusion

In this Chapter, we conclude this dissertation and discuss future work.

5.1 Contributions

Distortion minimization in distributed estimation mostly focused on single target and single antenna configuration. In this dissertation, we take the use of multiple antenna and multiple targets estimation. Hence, we consider the correlated data observation. To guarantee the quality of estimation, it is necessary to analyze the channel capacity between sensors and FC. Accordingly, study on channel capacity of data collection is essential.

In conclusion, we summarize the contributions of this dissertation as following:

- **Distributed Estimation over Correlated Data:** Our proposal is minimizing the distortion of estimation in terms of MSE by considering the correlated data observation and multiple antenna configuration. Given power constraint, we obtain closed-form solution that follows water-filling algorithm. The proposed methods show that the distortion increases as the data becomes correlated. Simulation results show that the noisy chan-

nel is more harmful than the observation noise. By taking into account the effects of correlation, observation noise, channel noise, the proposed methods have a better performance than the equal power method.

- **Capacity of Data Collection:** We analyze capacity of data collection using I-MMSE. We maximize the channel capacity by considering individual power constraint and sum power constraint. Optimization algorithms under both coherent MAC and orthogonal MAC are presented. We show that the capacity of data collection scale as $\Theta((1/2) \log(1+L))$ when the number of sensors L grows to infinity. The capacity of the optimal power allocation for both MAC models are larger than that of the equal power allocation. Simulation results show that the capacity of coherent MAC is larger than that of orthogonal MAC, particularly when the number of sensors L is large and the total power is fixed. This is a consequence of using orthogonal link from the sensors to FC where the corrupted channel cannot be eliminated even when the number of sensors goes to infinity.

5.2 Future Work

There are still many challenges in distributed estimation and capacity of data collection.

First, we focus on channel model between sensors and FC. Investigating the distributed estimation and capacity of data collection on various channel models will bring the analysis into more realistic scenario. In general, taking into account channel model is very challenging analysis, particularly for a large number of sensors.

Second, considering routing and data compression into both distributed estimation and capacity data collection is an interesting topic. Data compressing will reduce the amount of data being sent to the FC whether it will affect to the performance of estimation and the

capacity.

Third, the distributed estimation and capacity of data collection are enough for taking performance evaluation in WSN, the other evaluation methods such as outage distortion probability and outage capacity probability are still open problems.

Appendix A

Instead of intuitive assumption, we provide analytical derivation of the lower bound on MMSE estimation that can be achieved when the total power, $P \rightarrow \infty$ and the channel gain, $g_i = 1$. As $g_i = 1$, we can write eq. (4.9) as

$$\frac{1}{J} = 1 + \left(1 + \sum_{i=1}^L a_i^2\right)^{-1} \left(\sum_{i=1}^L a_i \sqrt{\alpha_i}\right)^2 \quad (\text{A.1})$$

Based on *Cauchy-Schwarz* inequality [89] that $(\sum_{i=1}^n k_i l_i)^2 \leq (\sum_{i=1}^n k_i^2) (\sum_{i=1}^n l_i^2)$, we can rewrite eq. (A.1) as

$$\frac{1}{J} \leq 1 + \left(1 + \sum_{i=1}^L a_i^2\right)^{-1} \left(\sum_{i=1}^L a_i^2\right) \left(\sum_{i=1}^L \alpha_i\right). \quad (\text{A.2})$$

We have power amplification factor, $a_i = \sqrt{P/L(\alpha_i^{-1} + 1)}$, then we have

$$\begin{aligned} \frac{1}{J} &\leq 1 + \left(1 + P \sum_{i=1}^L \frac{1}{L(\alpha_i^{-1} + 1)}\right)^{-1} \\ &\quad \times \left(P \sum_{i=1}^L \frac{1}{L(\alpha_i^{-1} + 1)}\right) \left(\sum_{i=1}^L \alpha_i\right). \end{aligned} \quad (\text{A.3})$$

Thus, as $P \rightarrow \infty$, $\frac{P \sum_{i=1}^L \frac{1}{L(\alpha_i^{-1}+1)}}{1+P \sum_{i=1}^L \frac{1}{L(\alpha_i^{-1}+1)}} \rightarrow 1$. Then, the lower bound of MMSE is

$$\frac{1}{J_0} = 1 + \sum_{i=1}^L \alpha_i. \quad (\text{A.4})$$

Appendix B

We know that eq. (4.19) is not a convex problem, but we can transform it into an equivalent convex form. Let suppose the optimal solution of eq. (4.19) is $\beta_{\text{opt}}(P)$, then $\frac{1}{\beta_{\text{opt}}(P)}$ should be monotonically decreasing as P increases. Then, we reform the objective function in eq. (4.19) as minimizing the total power consumption subject to a given inverse of total SNR constraint as follows:

$$\begin{aligned} \min_{P_i} \quad & P = \sum_{i=1}^L P_i \\ \text{s.t.} \quad & \frac{1}{\beta(P_1, \dots, P_L)} \leq \frac{1}{\beta}. \end{aligned} \tag{B.1}$$

In terms of power amplification factor a_i and straightforward modification of eq. (B.1), we get

$$\begin{aligned} \min_{a_i, \dots, a_L} \quad & \sum_{i=1}^L a_i^2 (1 + \alpha_i^{-1}) \\ \text{s.t.} \quad & \left(1 + \sum_{i=1}^L g_i^2 a_i^2 \right)^{-1} \left(\sum_{i=1}^L g_i a_i \sqrt{\alpha_i} \right)^2 \geq \beta^{-1}. \end{aligned} \tag{B.2}$$

We know that eq. (B.2) is still convex in terms of $a_{i=1}^L$. Therefore, we use a slack variable

$$t = \sum_{i=1}^L g_i a_i \sqrt{\alpha_i} \text{ as}$$

$$\begin{aligned} \min_{a_i, \dots, a_L} \quad & \sum_{i=1}^L a_i^2 (1 + \alpha_i^{-1}) \\ \text{s.t.} \quad & \left(1 + \sum_{i=1}^L g_i^2 a_i^2 \right) \geq \beta t^2 \\ & \sum_{i=1}^L g_i a_i \sqrt{\alpha_i} - t = 0. \end{aligned} \tag{B.3}$$

The Lagrangian function for eq. (B.3) is

$$\begin{aligned} \mathcal{L}(a_i, t, \mu, \eta) = & \sum_{i=1}^L a_i^2 (1 + \alpha_i^{-1}) \\ & + \mu \left(1 + \sum_{i=1}^L g_i^2 a_i^2 - \beta t^2 \right) \\ & + \eta \left(t - \sum_{i=1}^L g_i a_i \sqrt{\alpha_i} \right) \end{aligned} \tag{B.4}$$

where $\eta \in \mathbb{R}$ and $\mu \geq 0$. From the Lagrangian function, we can derive solution based on KTT condition [71]:

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial t} &= -2\mu\beta t + \eta = 0 \\ \frac{\partial \mathcal{L}}{\partial a_i} &= 2(1 + \alpha_i^{-1} + \mu g_i^2) a_i - g_i \sqrt{\alpha_i} \eta = 0, \quad 1 \leq i \leq L \\ \mu \left(1 + \sum_{i=1}^L g_i^2 a_i^2 - \beta t^2 \right) &= 0 \\ \sum_{i=1}^L g_i a_i \sqrt{\alpha_i} &= 0. \end{aligned} \tag{B.5}$$

From the second KKT conditions, we have

$$a_i = \frac{g_i \sqrt{\alpha_i} \eta}{2(1 + \alpha_i^{-1} + \mu g_i^2)}. \quad (\text{B.6})$$

We use the slack variable $t = \sum_{i=1}^L g_i a_i \sqrt{\alpha_i}$ and eq. (B.6) into the first KKT conditions, we get

$$\sum_{i=1}^L \frac{\mu g_i^2 \alpha_i}{1 + \alpha_i^{-1} + \mu g_i^2} = \beta^{-1}. \quad (\text{B.7})$$

Now, we plug in the first KKT conditions, $t = ((\beta^{-1})\eta)/2\mu$ and eq. (B.6) into the third KKT conditions as

$$\eta = 2 \left(\frac{\beta^{-1}}{\mu^2} - \sum_{i=1}^L \frac{g_i^4 \alpha_i}{(1 + \alpha_i^{-1} + \mu g_i^2)^2} \right)^{-1/2}. \quad (\text{B.8})$$

Thus, the optimal power allocation of sensor i in terms of μ and η is

$$P_i = a_i^2 (1 + \alpha_i^{-1}) = \frac{\eta^2}{4} \frac{g_i^2 \alpha_i (1 + \alpha_i^{-1})}{(1 + \alpha_i^{-1} + \mu g_i^2)^2}. \quad (\text{B.9})$$

From the fact that

$$\begin{aligned} P &= \sum_{i=1}^L P_i \\ &= \frac{\eta^2}{4} \sum_{i=1}^L \underbrace{\frac{g_i^2 \alpha_i (1 + \alpha_i^{-1})}{(1 + \alpha_i^{-1} + \mu g_i^2)^2}}_{=c^{-1}} \\ &\stackrel{(a)}{=} \mu. \end{aligned} \quad (\text{B.10})$$

By direct calculation that involves eqs. (B.7), (B.8), and (B.9), (a) is held. From eq. (B.6), we obtain that the optimal SNR, β_{opt} is a function of the total power, P as

$$\beta_{\text{opt}} = \sum_{i=1}^L \frac{\alpha_i}{1 + (\alpha_i^{-1} + 1)/(g_i^2 P)}. \quad (\text{B.11})$$

Bibliography

- [1] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, “A survey on sensor networks,” *IEEE Commun. Mag.*, vol. 40, no. 8, pp. 102–114, Aug. 2002.
- [2] J. Lin and G. AlRegib, “Distributed estimation in energy-constrained wireless sensor networks,” *IEEE Trans. Signal Process.*, vol. 57, no. 10, pp. 3746–3758, May 2009.
- [3] C. Wang, J. Jiang, Y. Liu, X. Y. Li, S. Tang, “Aggregation capacity of wireless sensor networks: Extended network case,” *IEEE Trans. Comput.*, vol. 63, no. 6, pp. 1351–1364, June 2014.
- [4] S. Chen, M. Huang, S. Tang, and Y. Wang, “Capacity of data collection in arbitrary wireless sensor networks,” *IEEE Trans. Parallel Distrib. Syst.*, vol. 23, no. 1, pp. 52–60, Jan. 2012.
- [5] P. Santi, “On the data gathering capacity and latency in wireless sensor networks,” *IEEE J. Sel. Areas Commun.*, vol. 28, no. 7, pp. 1211–1221, Sep. 2010.
- [6] H. E. Gamal, “On the scaling laws of dense wireless sensor networks: The data gathering channel,” *IEEE Trans. Inf. Theory*, vol. 51, no. 3, pp. 1229–1234, Mar. 2005.
- [7] R. Zheng and R. J. Barton, “Toward optimal data aggregation in random wireless sensor networks,” in *Proc. IEEE INFOCOM*, pp. 249–257, Houston, USA, May 2007.

- [8] B. Liu, D. Towsley, and A. Swami, "Data gathering capacity of large scale multihop wireless networks," in *Proc. IEEE MASS*, pp. 124–132, Atlanta, USA, Oct. 2008.
- [9] S. Chen, Y. Wang, X.-Y. Li, and X. Shi, "Capacity of data collection in randomly-deployed wireless sensor networks," *Wireless Netw.*, vol. 17, no. 2, pp. 305–318, Feb. 2011.
- [10] J. J. Xiao, S. Cui, Z. Q. Luo, and A. J. Goldsmith, "Linear Coherent Decentralized Estimation," *IEEE Trans. Signal Process.*, vol. 56, no. 2, pp. 757–770, Feb. 2008.
- [11] S. Cui, J. J. Xiao, A.J. Goldsmith, Z. Luo, and H. V. Poor, "Estimation diversity and energy efficiency in distributed sensing," *IEEE Trans. Signal Process.*, vol. 55, no. 9, pp. 4683–4695, Sep. 2007.
- [12] J. Kurose, V. Lesser, E. de Sousa e Silva, A. Jayasumana, and B. Liu, in *Sensor Networks Seminar*, CMPSCI 791L, University of Massachusetts, Amherst, MA, 2003.
- [13] K. Sohrawy, D. Minoli, and T. Znati, *Wireless Sensor Networks: Technology, Protocols, and Applications*. New Jersey, USA: John Wiley and Sons, 2007.
- [14] C.-H. Wu and C.-A. Lin, "Distributed estimation for vector signal in linear coherent sensor networks," *IEICE Trans. Commun.*, vol. E95-B, no. 2, pp. 460–465, Feb. 2012.
- [15] R. C. Luo and O. Chen, "Mobile sensor node deployment and asynchronous power management for wireless sensor networks," *IEEE Trans. Ind. Electron.*, vol. 59, no. 5, pp. 2377–2385, May 2012.
- [16] S. Mini, S. K. Udgate, and S. L. Sabat, "Sensor deployment and scheduling for target coverage problem in wireless sensor networks," *IEEE Sensors J.*, vol. 14, no. 3, pp. 636–644, Mar. 2014.

- [17] Y. Song, B. Wang, Z. Shi, K. R. Pattipati, and S. Gupta, "Distributed algorithms for energy-efficient even self-deployment in mobile sensor networks," *IEEE Trans. Mobile Comput.*, vol. 13, no. 5, pp. 1035–1047, May 2014.
- [18] W. Liang and Y. Liu, "Online data gathering for maximizing network lifetime in sensor networks," *IEEE Trans. Mobile Comput.*, vol. 6, no. 1, pp. 2–11, Jan. 2007.
- [19] M. Ma and Y. Yang, "SenCar: An energy-efficient data gathering mechanism for large-scale multihop sensor networks," *IEEE Trans. Parallel Distrib. Syst.*, vol. 18, no. 10, pp. 1476–1488, Oct. 2007.
- [20] M. Zhao, J. Li, and Y. Yang, "A framework of joint mobile energy replenishment and data gathering in wireless rechargeable sensor networks," *IEEE Trans. Mobile Comput.*, vol. 13, no. 12, pp. 2689–2705, Dec. 2014.
- [21] H. Zhang and H. Shen, "Balancing energy consumption to maximize network lifetime in data-gathering sensor networks," *IEEE Trans. Parallel Distrib. Syst.*, vol. 20, no. 10, pp. 1526–1539, Oct. 2009.
- [22] D. Takaishi, H. Nishiyama, N. Kato, and R. Miura, "Toward energy efficient big data gathering in densely distributed sensor networks," *IEEE Trans. Emerg. Topics Comput.*, vol. 2, no. 3, pp. 388–397, Sept. 2014.
- [23] J. Yuan and W. Yu, "Joint source coding, routing and power allocation in wireless sensor networks," *IEEE Trans. Commun.*, vol. 56, no. 6, pp. 886–896, June 2008.
- [24] H. Zhang and H. Shen, "Energy-efficient beaconless geographic routing in wireless sensor networks," *IEEE Trans. Parallel Distrib. Syst.*, vol. 21, no. 6, pp. 881–896, June 2010.

- [25] C. Pandana and K. J. R. Liu, "Robust connectivity-aware energy-efficient routing for wireless sensor networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 10, pp. 3904–3916, Oct. 2008.
- [26] S. J. Baek, G. de Veciana, and X. Su, "Minimizing energy consumption in large-scale sensor networks through distributed data compression and hierarchical aggregation," *IEEE J. Sel. Areas. Commun.*, vol. 22, no. 6, pp. 1130–1140, Aug. 2004.
- [27] C. Caione, D. Brunelli, and L. Benini, "Distributed compressive sampling for lifetime optimization in dense wireless sensor networks," *IEEE Trans. Ind. Informat.*, vol. 8, no. 1, pp. 30–40, Feb. 2012.
- [28] G. Mergen, Q. Zhao, and L. Tong, "Sensor networks with mobile access: energy and capacity considerations," *IEEE Trans. Commun.*, vol. 54, no. 11, pp. 2033–2044, Nov. 2006.
- [29] V. Rodoplu and T. H. Meng, "Bits-per-joule capacity of energy-limited wireless networks," *IEEE Trans. Commun.*, vol. 6, no. 3, pp. 857–865, Mar. 2007.
- [30] X. Zhang, H. V. Poor, and M. Chiang, "Optimal power allocation for distributed detection over MIMO channels in wireless sensor networks," *IEEE Trans. Signal Process.*, vol. 56, no. 9, pp. 4124–4140, Sept. 2008.
- [31] Y. Yilmaz and X. Wang, "Sequential distributed detection in energy-constrained wireless sensor networks," *IEEE Trans. Signal Process.*, vol. 62, no. 12, pp. 3180–3193, June 2014.
- [32] J. Park, G. Shevlyakov, and K. Kim, "Robust distributed detection with total power

- constraint in large wireless sensor networks,” *IEEE Trans. Wireless Commun.*, vol. 10, no. 7, pp. 2058–2062, July 2011.
- [33] S. Appadwedula, V. V. Veeravalli, and D. L. Jones, “Energy-efficient detection in sensor networks,” *IEEE J. Sel. Areas Commun.*, vol. 23, no. 4, pp. 693–702, Apr. 2005.
- [34] F. Jiang, J. Chen, and A. L. Swindlehurst, “Optimal power allocation for parameter tracking in a distributed amplify-and-forward sensor network,” *IEEE Trans. Signal Process.*, vol. 62, no. 9, pp. 2200–2211, May 2014.
- [35] S. Sengupta, M. Chatterjee, and K. A. Kwiat, “A game theoretic framework for power control in wireless sensor networks,” *IEEE Trans. Comput.*, vol. 59, no. 2, pp. 231–242, Feb. 2010.
- [36] X. Shi and G. Stromberg, “SyncWUF: An ultra low-power MAC protocol for wireless sensor networks,” *IEEE Trans. Mobile Comput.*, vol. 6, no. 1, pp. 115–125, Jan. 2007.
- [37] G. Vannucci, A. Bletsas, and D. Leigh, “A software-defined radio system for backscatter sensor networks,” *IEEE Trans. Wireless Commun.*, vol. 7, no. 6, pp. 2170–2179, June 2008.
- [38] A. Dunkels, B. Gronvall, and T. Voigt, “Contikia a lightweight and flexible operating system for tiny networked sensors,” in *Proc. EmNets*, pp.455–462, Sweden, Nov. 2004.
- [39] C.-C. Han, R. Kumar, R. Shea, E. Kohler, and M. Srivastava “A dynamic operating system for sensor nodes,” in *Proc. ACM MobiSys*, pp. 163–176, Washington, USA, Sep. 2005.
- [40] S. Bhatti, J. Carlson, H. Dai, J. Deng, J. Rose, A. Sheth, B. Shucker, C. Gruenwald, A. Torgerson, and R. Han, “MANTIS OS: An embedded multithreaded operating system

- for wireless micro sensor platforms,” *ACM J. Mobile Netw.*, vol. 10, pp. 563–579, Aug. 2005.
- [41] A. Eswaran, A. Rowe, and R. Rajkumar, “Nano-RK: An energy-aware resource-centric RTOS for sensor networks,” in *Proc. IEEE RTSS*, pp. 255–265, Miami, USA, Dec. 2005.
- [42] H. Cha , S. Choi , I. Jung, H. Kim and H. Shin, “RETOS: Resilient, expandable, and threaded operating system for wireless sensor networks,” in *Proc. ACM/IEEE IPSN*, pp. 148–157, Cambridge, USA, Oct. 2007.
- [43] Q. Cao, T. F. Abdelzaher, and J. A. Stankovic, “The LiteOS operating system: towards Unix-like abstractions for wireless sensor networks,” in *Proc. ACM/IEEE IPSN*, pp. 233–244, St. Louis, USA, Sep. 2008.
- [44] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, “Wireless sensor networks: a survey,” *Comput. Netw.*, vol. 38, no. 4, pp. 393–422, Mar. 2002.
- [45] J. J. Xiao, A. Ribeiro, Z. Q. Luo, and G. B. Giannakis, “Distributed compression-estimation using wireless sensor networks,” *IEEE Signal Process. Mag.*, vol. 23, no. 4, pp. 27–41, July 2006.
- [46] L. Cheng, C. D. Wu, and Y. Z. Zhang, “Indoor robot localization based on wireless sensor networks,” *IEEE Trans. Consum. Elect.*, vol. 57, no. 3, pp. 1099–1104, Aug. 2011.
- [47] V. Jelcic, M. Magno, D. Brunelli, G. Paci, and L. Benini, “Context-adaptive multimodal wireless sensor network for energy-efficient gas monitoring,” *IEEE Sensors J.*, vol. 13, no. 1, pp. 328–338, Jan. 2013.
- [48] G. Tolle, J. Polastre, R. Szewczyk, D. Culler, N. Turner, K. Tu, S. Burgess, T. Dawson,

- P. Buonadonna, D. Gay, and W. Hong, "A macroscope in the redwoods," in *Proc. 3rd SenSys*, pp. 51–63 San Diego, USA, Nov. 2005.
- [49] S. Savazzi, U. Spagnolini, L. Goratti, D. Molteni, M. Latva-aho, and M. Nicoli, "Ultra-wide band sensor networks in oil and gas explorations," , *IEEE Commun. Mag.*, vol. 51, no. 4, pp. 150–160, Apr. 2013.
- [50] J. Yick, B. Mukherjee, and D. Ghosal, "Wireless sensor network survey," *Computer Netw.*, vol. 52, no. 12, pp. 2292-2330, Aug. 2008
- [51] I. F. Akyildiz, D. Pompili, and T. Melodia, "Challenges for efficient communication in underwater acoustic sensor networks," *ACM SIGBED Rev.*, vol. 1, no. 2, pp. 3-8, Jul. 2004.
- [52] A. R. Silva and M. C. Vuran, "Empirical Evaluation of wireless underground-to-underground communication in wireless underground sensor networks," in *Proc. IEEE DCOSS*, pp. 231–244, Marina Del Rey, USA, June 2009.
- [53] I.F. Akyildiz, Z. Sun, and M.C. Vuran, "Signal propagation techniques for wireless underground communication networks." *Physical Commun.*, vol 2, no. 3, pp. 167-183, Sep. 2009.
- [54] J. Heidemann, Y. Wei, J. Wills, A. Syed, and L. Yuan, "Research challenges and applications for underwater sensor networking," in *Proc. IEEE WCNC*, vol. 1, pp. 228–235, Apr. 2006.
- [55] G. W. Taylor, J. R. Burns, S. A. Kammann, W. B. Powers, and T. R. Welsh, "The Energy Harvesting Eel: a small subsurface ocean/river power generator," *IEEE J. Ocean. Engineer.*, vol. 26, no. 4, pp. 539–547, Oct. 2001.

- [56] S. M. Kay, *Fundamental of statistical signal processing: estimation theory*, Upper Saddle River, NJ: Prentice-Hall, 1993.
- [57] C. E. Shannon, "A mathematical theory of communication," *Bell System Technical Journal*, vol. 27, pp. 379–423 and 623–656, July and Oct. 1948.
- [58] S. Talarico, N. A. Schmid, M. Alkhweldi, and M. C. Valentin, "Distributed estimation of a parametric field: Algorithms and performance analysis," *IEEE Trans. Signal Process.*, vol. 62, no. 5, pp. 1041–1053, Nov. 2013.
- [59] P. Gupta and P.R. Kumar, "The capacity of wireless networks," *IEEE Trans. Inf. Theory*, vol. 46, no. 2, pp. 388–404, Mar. 2000.
- [60] J. Fang and H. Li, "Power constrained distributed estimation with correlated sensor data," *IEEE Trans. Signal Process.*, vol. 57, no. 8, pp. 3292–3297, Aug. 2009.
- [61] W. Guo, J. J. Xiao, and S. Cui, "An efficient water-filling solution for linear coherent joint estimation," *IEEE Trans. Signal Process.*, vol. 56, no. 10, pp. 5301–5305, Oct. 2008.
- [62] I. Bahceci and A. Khandani, "Linear estimation of correlated data in wireless sensor networks with optimum power allocation and analog modulation," *IEEE Trans. Commun.*, vol. 56, no. 7, pp. 1146–1156, July 2008.
- [63] J. Fang, H. Li, Z. Chen, and S. Li, "Optimal precoding design and power allocation for decentralized detection of deterministic signals," *IEEE Trans. Signal Process.*, vol. 60, no. 6, pp. 3149–3163, June 2012.
- [64] H. Senol and C. Tepedelenlioglu, "Performance of distributed estimation over unknown parallel fading channels," *IEEE Trans. Signal Process.*, vol. 56, no. 12, pp. 6057–6068, Dec. 2008.

- [65] S. K. Jayaweera, "Virtual MIMO-based cooperative communication for energy constrained wireless sensor networks," *IEEE Trans. Wireless Comm.*, vol. 5, no. 5, pp. 984–989, May 2006.
- [66] Y. Gai, L. Zhang, and X. Shan, "Energy efficiency of cooperative MIMO with data aggregation in wireless sensor networks," in *Proc. IEEE WCNC*, pp. 791–796, Kowloon, Mar. 2007.
- [67] Q. Gao, Y. Zuo, J. Zhang, and X. H. Peng, "Improving energy efficiency in a wireless sensor network by combining cooperative MIMO with data aggregation," *IEEE Trans. Veh. Technol.*, vol. 59, no. 8, pp. 3956–3965, Oct. 2010.
- [68] M. H. Chaudhary and L. Vandendorpe, "Power constrained linear estimation in wireless sensor networks with correlated data and digital modulation," *IEEE Trans. Signal Process.*, vol. 60, no. 2, pp. 570–584, Feb. 2012.
- [69] J. Wu and N. Sun, "Optimum sensor density in distortion-tolerant wireless sensor networks," *IEEE Trans. Wireless Commun.*, vol. 11, no. 6, pp. 2056–2064, June 2012.
- [70] A. W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and Its Applications*. New York: Academic, 1979.
- [71] S. Boyd and L. Vanderberghe, *Convex Optimization*, Cambridge univ. Press, Cambridge, U.K., 2003.
- [72] C. Wang, C. Jiang, Y. Liu, X. Li, and S. Tang, "Aggregation capacity of wireless sensor networks: Extended network case," *IEEE Trans. Comp.*, vol. 63, no. 99, pp. 1–10, Sept. 2012.

- [73] E. J. Duarte-Melo and M. Liu, “Data-gathering wireless sensor networks: Organization and capacity,” *Computer Netw.*, vol. 43, pp. 519–537, Nov. 2003.
- [74] D. Marco, E. J. Duarte-Melo, M. Liu, and D.L. Neuhoff, “On the many-to-one transport capacity of a dense wireless sensor network and the compressibility of its data,” in *Proc. IPSN*, pp. 1–16, Palo Alto, USA, Apr. 2003.
- [75] H. Zheng, S. Xiao, X. Wang, X. Tian, and M. Guizani, “Capacity and delay analysis for data gathering with compressive sensing in wireless sensor networks,” *IEEE Trans. Wireless Commun.*, vol. 12, no. 2, pp. 917–927, Feb. 2013.
- [76] S. Marano, V. Matta, and P. Willett, “Distributed estimation in large wireless sensor network via a locally optimum approach,” *IEEE Trans. Signal Process.*, vol. 56, no. 2, pp. 748–756, Feb. 2008.
- [77] A. Ribeiro and G. B. Giannakis, “Bandwidth-constrained distributed estimation for wireless sensor networks-Part I: Gaussian case,” *IEEE Trans. Signal Process.*, vol. 54, no. 3, pp. 1131–1143, Mar. 2006.
- [78] A. Ribeiro and G. B. Giannakis, “Bandwidth-constrained distributed estimation for wireless sensor networks-Part II: Unknown probability density function,” *IEEE Trans. Signal Process.*, vol. 54, no. 7, pp. 2784–2796, July 2006.
- [79] Z.-Q. Luo “Universal decentralized estimation in a bandwidth constrained sensor network,” *IEEE Trans. Inf. Theory*, vol. 51, no. 6, pp. 2210–2219, June 2005.
- [80] J.-J. Xiao and Z.-Q. Luo, “Universal decentralized estimation in an inhomogeneous sensing environment,” *IEEE Trans. Inf. Theory*, vol. 51, no. 10, pp. 3564–3575, Oct. 2005.

- [81] K. Liu, H. El-Gamal, and A. Sayeed, "On optimal parametric field estimation in sensor networks," *Proc. IEEE WSSP*, pp. 1170–1175, Novosibirsk, July 2005.
- [82] G. Mergen and L. Tong, "Type-based estimation over multiaccess channels," *IEEE Trans. Signal Process.*, vol. 54, no. 2, pp. 613–626, Feb. 2006.
- [83] M. K. Banavar, C. Tepedelenlioglu, and A. Spanias, "Estimation over fading channels with limited feedback using distributed sensing," *IEEE Trans. Signal Process.*, vol. 58, no. 1, pp. 414–425, Jan. 2010.
- [84] H. Senol and C. Tepedelenlioglu, "Performance of distributed estimation over unknown parallel fading channels," *IEEE Trans. Signal Process.*, vol. 56, no. 12, pp. 6057–6068, Dec. 2008.
- [85] T. J. Goblick, "Theoretical limitations on the transmission of data from analog sources," *IEEE Trans. Inf. Theory*, vol. 11, no. 4, pp. 558–567, Oct. 1965.
- [86] M. Gastpar, "Uncoded transmission is exactly optimal for a simple Gaussian sensor network," *IEEE Trans. Inf. Theory*, vol. 54, no. 11, pp. 5247–5251, Nov. 2003.
- [87] D. Guo, S. Shamai, and S. Verdú, "Mutual information and minimum mean-square error in Gaussian channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 4, pp. 1261–1282, Apr. 2005.
- [88] F. Jiang, J. Chen, and A. L. Swindlehurst, "Optimal power allocation for parameter tracking in a distributed amplify-and-forward sensor network," *IEEE Trans. Signal Process.*, vol. 62, no. 9, pp. 2200–2211, Feb. 2014.
- [89] J. M. Steele, *The Cauchy-Schwarz Master Class*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

- [90] J. Agre and L. Clare, "An integrated architecture for cooperative sensing networks," *Computer*, vol. 33, no. 5, pp. 106–108, May 2000.
- [91] C. Y. Chong and S. Kumar, "Sensor networks: Evolution, opportunities and challenges," *Proceedings of IEEE*, vol. 91, no. 8, pp. 1247–1256, Aug. 2003.
- [92] S. Kumar, F. Zhao, and D. Shepherd, "Special issue on collaborative information processing in microsensor networks," *IEEE Signal Proces. Mag.*, vol. 19, pp. 13–14, Mar. 2002.
- [93] Z. Zhao, F., J. Liu, L. Guibas, and J. Reich, "Collaborative signal and information processing: An information-directed approach," *Proceedings of the IEEE*, vol. 91, no. 8, pp. 1199–1209, Aug. 2003.
- [94] F. Zhao, J. Shin, and J. Reich, "Information-driven dynamic sensor collaboration," *IEEE Signal Proces. Mag.*, vol. 19, no. 2 pp. 61–72, Mar. 2002.
- [95] V. V. Veeravalli and J. Chamberland, "Wireless sensors in distributed detection application," *IEEE Signal Process. Mag.*, vol. 24, no. 3, pp. 16–25, May 2007.
- [96] K. Langendoen and N. Reijers, "Distributed localization in wireless sensor networks: a quantitative comparison," *Comput. Netw.*, vol. 43, no. 3, pp. 499–518, Nov. 2003.
- [97] R. Brooks, C. Griffin, and D. S. Friedlander, "Self-Organized distributed sensor network entity tracking," *International Journal of High Performance Computing Applications*, vol. 16, no. 3, pp. 207–219, Aug. 2002.
- [98] R. Brooks and C. Griffin, "Traffic model evaluation of ad hoc target tracking algorithms," *International Journal of High Performance Computing Applications*, vol. 16, no. 3, pp. 221–234, Aug. 2002.

- [99] Dan Li, K. D. Wong, Y. H. Hu, and A. M. Sayeed, "Detection, classification, and tracking of targets in distributed sensor networks," *IEEE Signal Process. Mag.*, vol. 19, no. 2, pp. 17–29, Mar. 2002.
- [100] J. A. Gubner, "Distributed estimation and quantization," *IEEE Trans. Inf. Theory*, vol. 39, no. 4, pp. 1456–1459, Jul. 1993.
- [101] K. Liu, H. E. Gamal, and A. Sayeed, "Decentralized inference over multiple access channels," *IEEE Trans. Signal Process.*, vol. 55, no. 7, pp. 3445–3455, Jul. 2007.
- [102] J.-J. Xiao, S. Cui, Z.-Q. Luo, and A. J. Goldsmith, "Power scheduling of universal decentralized estimation in sensor networks," *IEEE Trans. Signal Process.*, vol. 54, no. 2, pp. 413–422, Feb. 2006.
- [103] J. Li and G. AlRegib, "Rate constraint distributed estimation in wireless sensor networks," *IEEE Trans. Signal Process.*, vol. 55, no. 5, pp. 1634–1643, May 2007.
- [104] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions*, New York: Courier Dover, 1965.
- [105] M. Grossglauser and D.N.C. Tse., "Mobility increases the capacity of ad hoc wireless networks," *IEEE/ACM Trans. Netw.*, vol. 10, no. 4, pp. 477–486, Aug. 2002.
- [106] M. Gastpar and M. Vetterli. "On the capacity of wireless networks: the relay case," in *Proc. IEEE INFOCOM*, vol. 3, pp. 1577–1586, Jun. 2002.

Appendix C

C.1 Journals

- [1] A. S. Arifin and T. Ohtsuki, “Water-Filling Solution for Distributed Estimation of Correlated Data in WSN MIMO System,” *ISRN Sensor Netw.*, vol. 2013, 7 pages, Aug. 2013.
- [2] A. S. Arifin and T. Ohtsuki, “Capacity of Data Collection in Wireless Sensor Networks based on Mutual Information and MMSE Estimation,” *ISRN Sensor Netw.*, vol. 2014, 9 pages, Feb. 2014.

C.2 International Conferences

- [1] A. S. Arifin and T. Ohtsuki, “Linear Precoding for Distributed Estimation of Correlated Sources in WSN MIMO System,” in *Proc. IEEE VTC*, pp. 15, Dresden, Germany, June 2013.
- [2] A. S. Arifin and T. Ohtsuki, “Analysis of Outage Probability for Full-Duplex SISO System with Self-Interference,” in *Proc. APCC*, pp., Pattaya, Thailand, Oct. 2014.

- [3] A. S. Arifin and T. Ohtsuki, “Global SNR Maximization for Distributed Estimation and Capacity of Data Collection in Wireless Sensor Networks Using Individual Power Constraint,” in *Proc. APSIPA ASC*, pp., Siem Reap , Cambodia, Dec. 2014.

C.3 Domestic Conferences

- [1] A. S. Arifin and T. Ohtsuki, “Optimal Precoding Design and Power Allocation for Distributed Estimation of Correlated Data in WSN MIMO System,” in *Proc. IEICE Tech. Rep.*, vol. 112, no. 242, pp. 155–160, Fukuoka, Japan Oct. 2013.
- [2] A. S. Arifin and T. Ohtsuki, “Many to One Transport Capacity of Wireless Sensor Networks based on I-MMSE,” in *Proc. IEICE Tech. Rep.*, vol. 113, no. 328, pp. 85–89, Tokyo, Japan, Nov. 2013.
- [3] A. S. Arifin and T. Ohtsuki, “Many to One Transport Capacity of Orthogonal Wireless Sensor Networks based on I-MMSE,” in *Proc. IEICE Tech. Rep.*, vol. 113, no. 361, pp. 125–129, Kanagawa, Japan, Des. 2013.
- [4] A. S. Arifin and T. Ohtsuki, “Asymptotic Ergodic Capacity for Full-Duplex MIMO Relay Networks,” in *Proc. IEICE Tech. Rep.*, vol. 113, no. 456, pp. 377–382, Tokyo, Japan, Mar. 2014.
- [5] A. S. Arifin and T. Ohtsuki, “Outage Probability Analysis in Bidirectional Full-Duplex SISO System with Self-Interference,” in *Proc. IEICE Tech. Rep.*, vol. 114, no. 164, pp. 257–260, Kyoto, Japan, Aug. 2014.