SECRECY CONSTRAINED DISTRIBUTED INFERENCE IN WIRELESS SENSOR NETWORKS

by

Jun Guo

A dissertation
submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy in Electrical and Computer Engineering
Boise State University

May 2017

BOISE STATE UNIVERSITY GRADUATE COLLEGE

DEFENSE COMMITTEE AND FINAL READING APPROVALS

of the dissertation submitted by

Jun Guo

Dissertation Title: Secrecy Constrained Distributed Inference In Wireless Sensor Networks

Date of Final Oral Examination: 7th April 2017

The following individuals read and discussed the dissertation submitted by student Jun Guo, and they evaluated his presentation and response to questions during the final oral examination. They found that the student passed the final oral examination.

Hao Chen, Ph.D.	Chair of the Supervisory Committee
John Chiasson, Ph.D.	Member, Supervisory Committee
Leming Qu, Ph.D.	Member, Supervisory Committee
Qi Cheng, Ph.D.	External Examiner

The final reading approval of the dissertation was granted by Hao Chen, Ph.D., Chair of the Supervisory Committee. The dissertation was approved by the Graduate College.

dedicated to my family

ACKNOWLEDGMENTS

In his book, *The Ph.D. Grind*, Philip Guo wrote, “there would be no Ph.D. without ten thousand hours of unglamorous, hard-nosed grinding”. I have the same feeling about my pursuit of Ph.D., which is closing to an end. However, I do not think hard work alone is enough for me to come so far without the guidance, encouragement and support from my advisor, professors, family and friends. So, I would like to express my gratitude to all of them on this page.

I am deeply indebted to my advisor, Dr. Hao Chen, for his belief in me from the very beginning and patient guidance on my research during the course of my Ph.D. study. I have learned a great deal on statistical inference from Dr. Chen, but one thing I am most impressed is his mathematical intuition. “Our intuition is based on accumulated and compiled experiences, not on magic”, mentioned in the book, *The power of intuition*, which means it is not something that one can pick up in a short amount of time; however it motivates me to keep accumulating, learning and thinking for the rest of my life. Besides, his sharpness and humor make me feel that he is more like a mentor and a caring friend, who also gave me freedom and support to collaborate with other professors and researchers on signal processing and machine learning projects.

I am grateful to Dr. John Chiasson, Dr. Leming Qu, and Dr. Qi Cheng for serving in my committee. Dr. Chiasson’s machine learning course motivated me to dive into the field of deep learning. We often discuss things beyond engineering topics and his humor definitely made my Ph.D. life more pleasant.

I would like to thank my close collaborator, Uri Rogers. I had the distinct pleasure of working with and learning from Uri on statistical inference and his way of writing.

I am really fortunate to have friends, Kehan Zhu and Mucun Tian during the course of my doctoral study. Throughout these years, we encourage and help out each other on research and daily lives.

This thesis is dedicated to my mother for the unconditional love, support and belief in me. To my sister, Jin, who always cares me deeply in her heart and taught me to be a resilient person. To my wife, Yuwen, who helps me find my inner peace and gives strength to keep exploring the road less traveled. To my child, who always reminds me to have a beginner's mind.

Last but not the least, the writing of this thesis heavily relies on open-source operating system and tools including GNU/Linux, LaTeX, LaTeX draw, Vim, Emacs, Python, git, Gummi and so on. I revere the visionary leaders in free software revolution, Richard Stallman and Linus Torvalds who made free software possible at the very beginning. I am also thankful for all the programmers and donors who contribute to open-source software. Your great work indeed makes the world a better place.

ABSTRACT

Comprised of a large number of low-cost, low-power, mobile and miniature sensors, wireless sensor networks are widely employed in many applications, such as environmental monitoring, health-care, and diagnostics of complex systems. In wireless sensor networks, the sensor outputs are transmitted across a wireless communication network to legitimate users such as fusion centers for final decision-making.

Because of the wireless links across the network, the data are vulnerable to security breaches. For many applications, the data collected by local sensors are extremely sensitive, and care must be taken to prevent that information from being leaked to any malicious third parties, e.g., eavesdroppers. Eavesdropping is one of the most significant threats to wireless sensor networks, where local sensors are tapped by an eavesdropper in order to intercept information.

I considered distributed inference in the presence of a global, greedy and informed eavesdropper who has access to all local node outputs rather than access. My goal is to develop secured distributed systems against eavesdropping attacks using a physical-layer security approach instead of cryptography techniques because of the stringent constraints on sensor networks energy and computational capability. The physical-layer security approach utilizes the characteristics of the physical layer, including transmission channels noises, and the information of the source. Additionally, physical-layer security for distributed inference is scalable due to the low computational complexity.

I first investigate secrecy constrained distributed detection under both Neyman-

Pearson and Bayesian frameworks. I analyze the asymptotic detection performance and proposed a novel way of analyzing the maximum performance trade-off using Kullback-Leibler divergence ratio between the fusion center and eavesdropper. Under the Neyman-Pearson framework, I show that the eavesdropper's detection performance can be limited such that her decision-making is no better than random guessing; meanwhile, the detection performance at the fusion center is guaranteed at the prespecified level. Similar analyses and proofs are provided under the Bayesian framework, where it was shown that an eavesdropper can be constrained to an error probability level equal to her prior information. Additionally, I derive the asymptotic error exponent and show that asymptotic perfect secrecy and asymptotic perfect detection are possible by increasing the number of sensors under both frameworks if the fusion center has noiseless channels to the sensors.

For secrecy constrained distributed estimation, I conducted similar analysis under both a classical setting and Bayesian setting. I derived the maximum achievable secrecy performance and show that under the condition that the eavesdropper has noisy channels and the fusion center has noiseless channels, both asymptotic perfect secrecy and asymptotic perfect estimation can be achieved under a classical setting. Similarly, under a Bayesian setting, I derived the performance trade-off using Fisher information ratio and show that the fusion center outperforms the eavesdropper significantly in the simulation section.

Secrecy constrained in distributed inference with Rayleigh fading binary symmetric channel is considered as well. Similarly, I derive the maximum achievable secrecy performance ratio for both detection and estimation.

The maximum achievable trade-off turns out to be almost the same in distributed estimation as in distributed detection. This suggests that a universal framework for generally structured inference problems are feasible. Further investigations are needed to justify this conjecture for more general applications.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

KLD – **K**ullback **L**eibler **D**ivergence

WSN – **W**ireless **S**ensor **N**etwork

FC – **F**usion **C**enter

IoT – **I**nternet of **T**hings

LRT – **L**ikelihood **R**atio **T**est

ROC – **R**eceiver **O**perating **C**haracteristic

MSE – **M**ean **S**quared **E**rror

CRLB – **C**ramér **R**ao **L**ower **B**ound

LRQ – **L**ikelihood **R**atio **Q**uantizer

BER – **B**it **E**rror **R**ate

BSC – **B**inary **S**ymmetric **C**hannel

SNR – **S**ignal **N**oise **R**atio

FI – **F**isher **I**nformation

BPSK – **B**inary **P**hase **S**hifting **K**eying

LIST OF SYMBOLS

α	sensor false alarm probability
β	sensor detection probability
η	decision threshold
P_f	false alarm probability at the decision center
P_d	detection probability at the decision center
P_m	missed detection probability at the decision center
T_E	prespecified performance threshold for Eve
T_F	prespecified performance threshold for the fusion center
P_e	probability of error at the decision center
ρ	bit error rate
γ	decision rule
ϵ	mean squared error
I	Fisher information for one sensor
$\mathbf{I}$	Fisher information for all sensors
$D(P_0(\cdot)\|\|P_1(\cdot))$	KLD for one sensor
$\mathbb{D}(P_0(\cdot)\|\|P_1(\cdot))$	total KLD for all sensors

CHAPTER 1

INTRODUCTION

1.1 Wireless Sensor Networks

Comprised of a large number of low-cost, low-power, mobile and miniature sensors, wireless sensor networks (WSNs) are systems of detecting phenomena, estimating parameters or measuring some physical properties of the environment, where sensors are densely deployed to the region of interest [1, 5, 83, 88]. Many WSNs have a dedicated node called sink node or fusion center (FC), of which the computational capability is more powerful than other sensing nodes because of data fusion requirements. Due to energy constraints, time-delay, bandwidth and memory limitations, the local nodes cannot send all the observed information directly to the FC where the final decision is made. The data observed by local sensors must be quantized or compressed before transmission over wireless channels to the FC. Therefore, one of the essential problems in WSNs is to design and optimize the local quantization rule for local nodes and fusion rule at the decision center in order to make the optimal inference at the FC based on the transmitted data from senors [5, 83, 88, 100].

1.1.1 Topologies in WSNs

Different WSNs have different network topologies, which determines different ways of communications within sensors and how the sensors send their data to the FC. The common topologies in WSNs include peer to peer networks, parallel (star) networks, tree networks, and mesh networks [72].

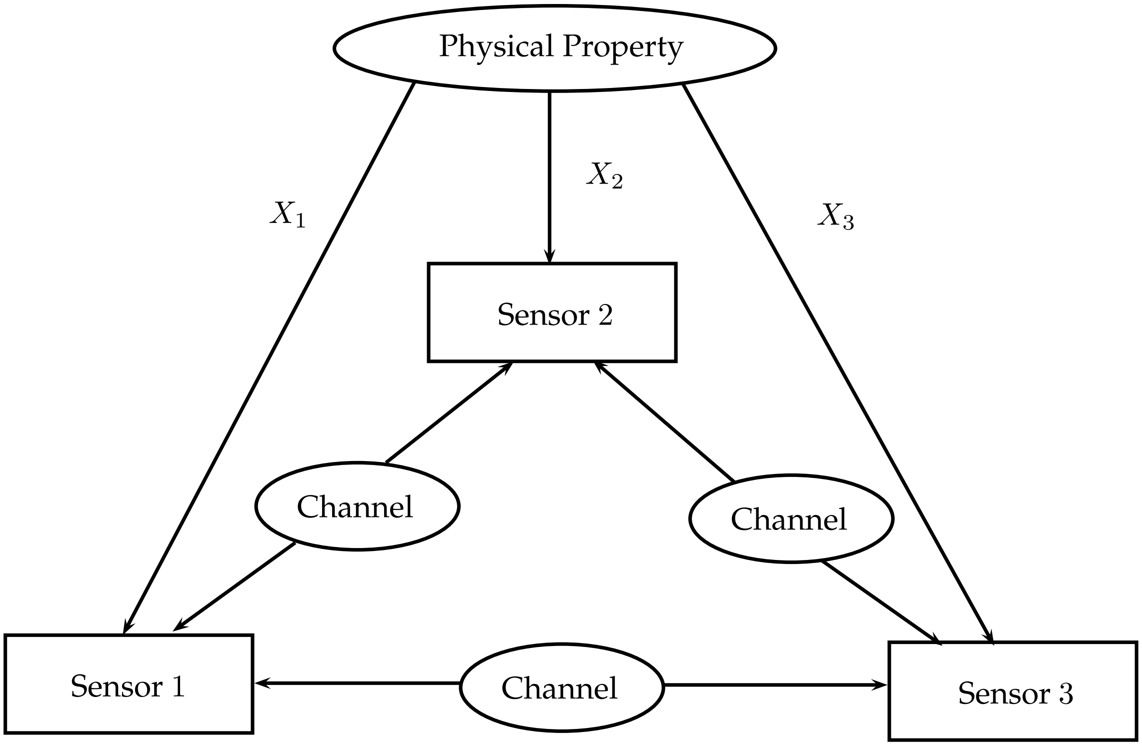

Figure 1.1: Peer-to-Peer Topology

In peer-to-peer network for three sensors shown in Figure 1.1, local sensors observe the physical property and they are able to send the outputs to each other across their respective channels. In this way, each sensor can be considered as the FC. Therefore, this network topology is flexible in a sense that when one sensor fails, another sensor could take over the job for decision making.

In Figure 1.2, we show the parallel (star) structure, where sensors observe phenomena in parallel and send their outputs to a FC through parallel channels. Unlike the peer-to-peer topology, each node cannot directly communicate with one another, and the FC is a fixed receiver. Parallel network is one of the most widely used structure in

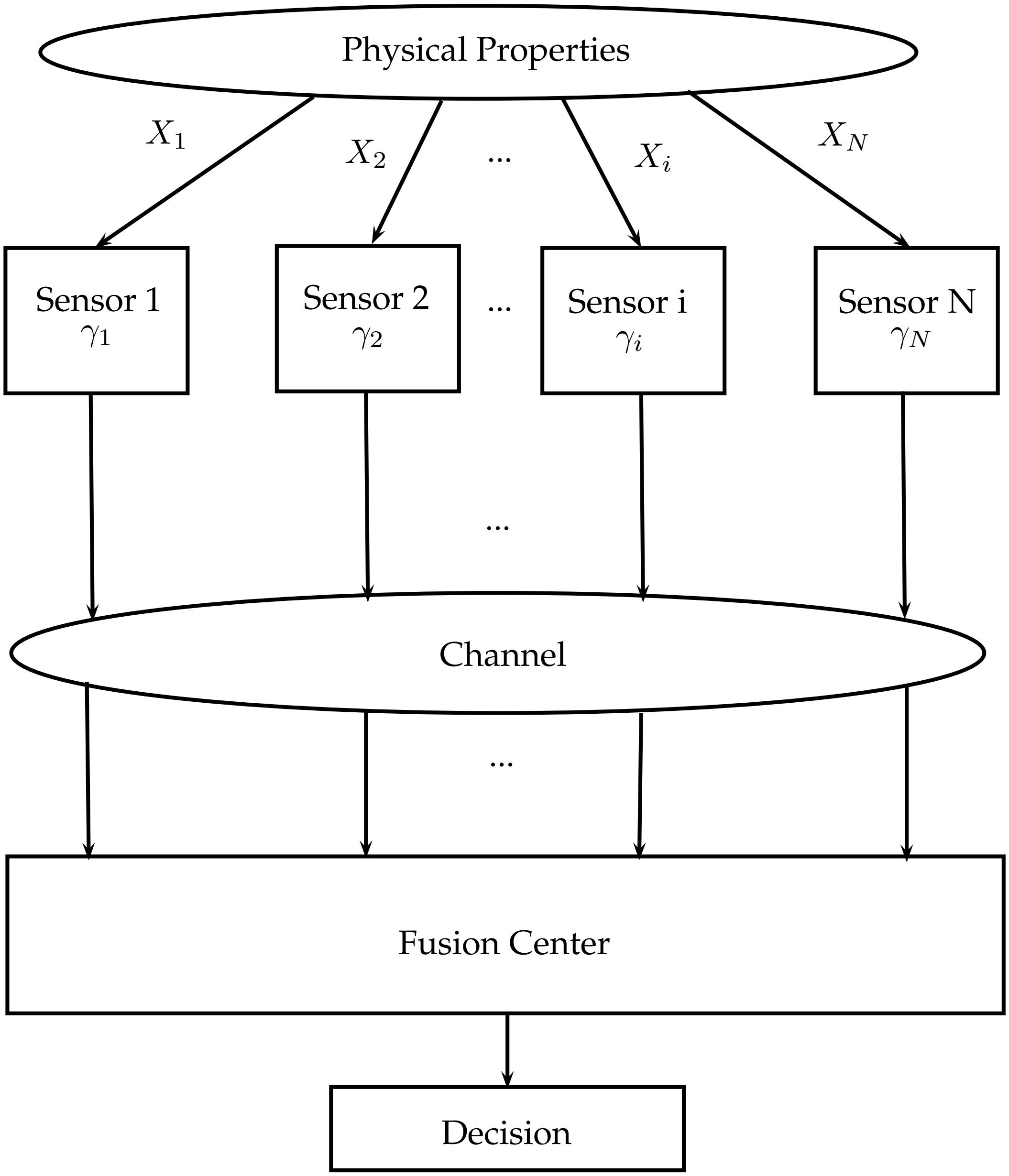

Figure 1.2: Parallel Topology

WSNs due to its simplicity and robustness. Under such a setting, the failure of a small portion of sensors will not deteriorate the performance of the network significantly.

Tree network, shown in Figure 1.3, however, is a hierarchical structure, where low

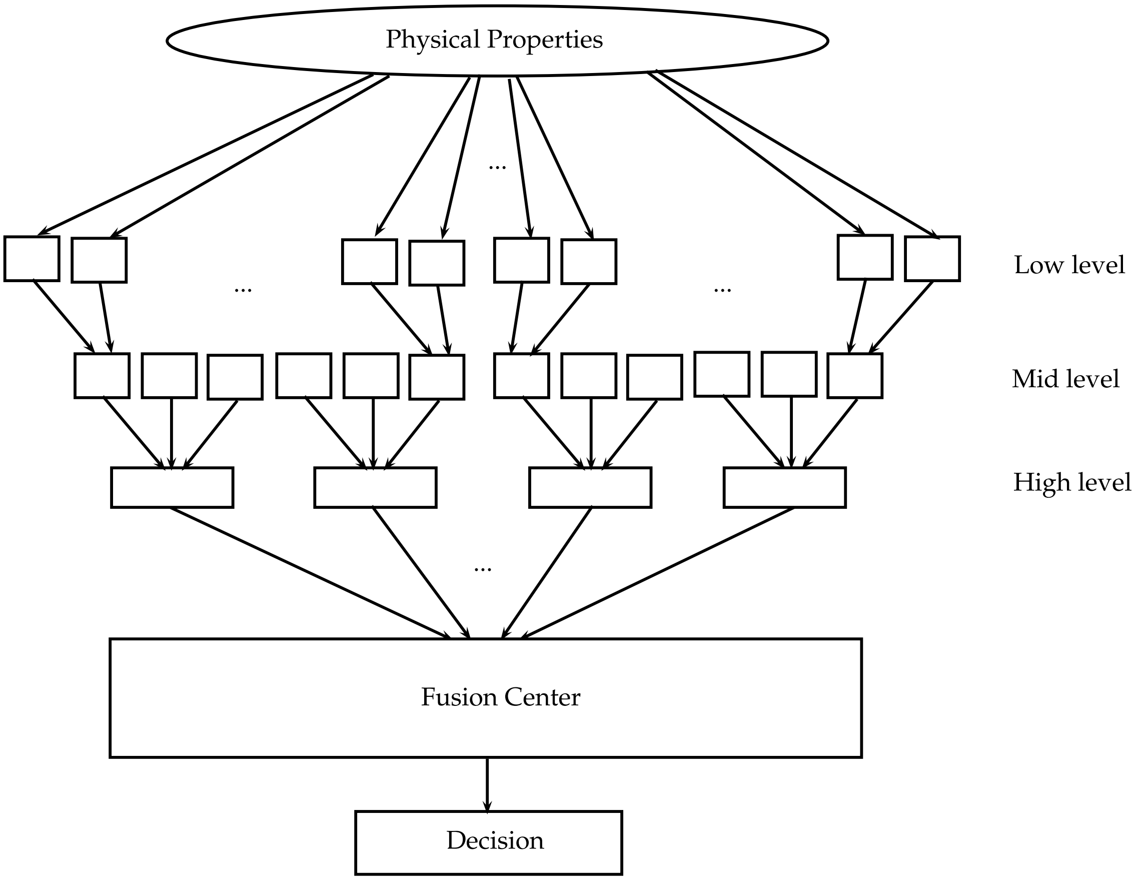

Figure 1.3: Tree Topology

level sensors observe the physical properties and then send their outputs to the next level and they keep doing so until the FC receives the output from high level nodes. This multi-hop communication is expected to consume less power than the single hop communication. Furthermore, the transmission power is low [97].

Mesh structured network with four sensors is shown in Figure 1.4. This setup allows data to hop from sensor to sensor. Similar to peer-to-peer topology, mesh network allows sensors to directly transmit the data to another sensor and the FC does not need to be fixed. Another advantage of this structure is that data can be transmitted from different routes to the desired location. The mesh network is the

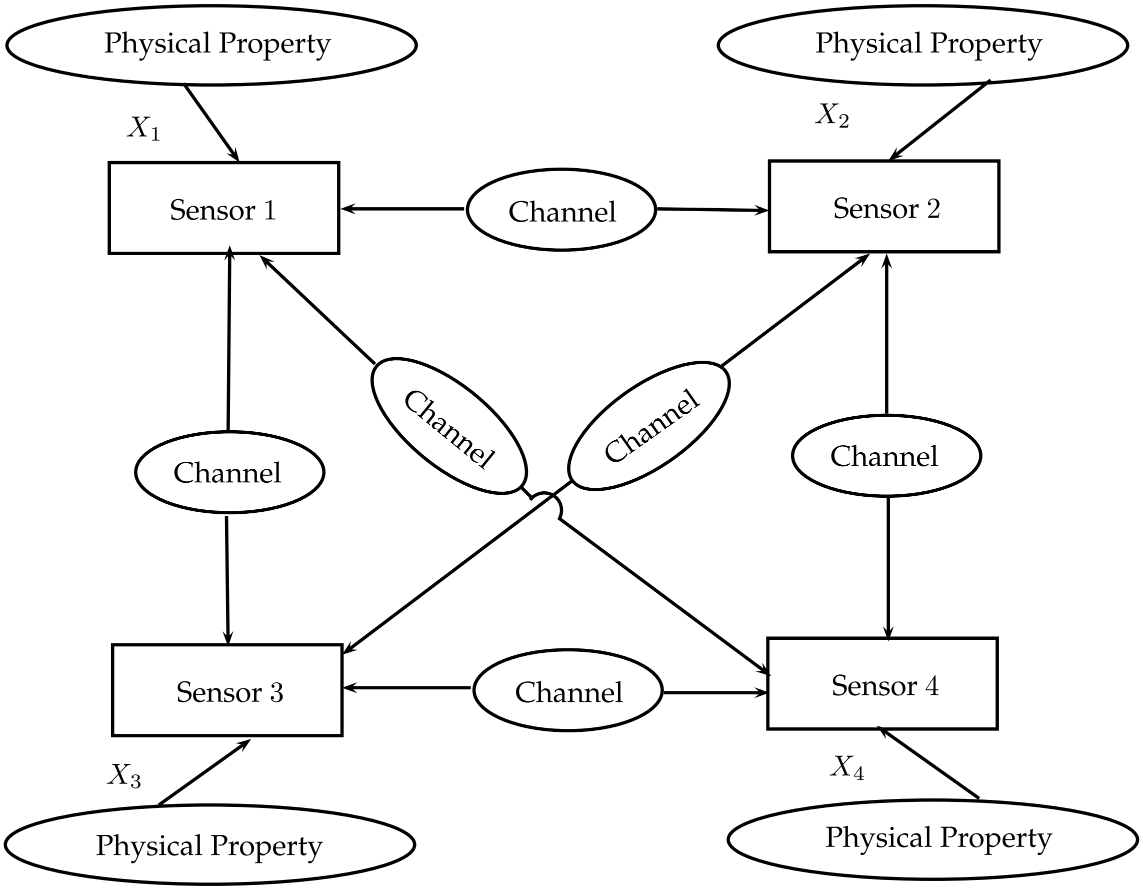

Figure 1.4: Mesh Topology

most complicated structure.

To sum up, the characteristics of common topologies for WSNs are summarized in Table 1.1. The topology of a WSN does not necessarily remain the same because the sensors could be mobile and their locations may change from one place to another.

1.1.2 Sensors Communication Protocol

With the topologies of WSNs defined, we know the structure of communications in WSNs. However, without the definition of communication protocol, sensors still could not communicate with each other. Such communication protocols are used

Table 1.1: **Characteristics of Common Topologies for WSNs**

Topology Name	Advantages	Disadvantages
Peer-to-Peer	flexible	not robust to sensor failures
Parallel	simple; robust in terms of sensor failures and network performance	not flexible
Tree	flexible; energy-efficient	not robust to backbone sensor failures
Mesh	flexible; robust	complicated

by sensor-to-sensor and sensor-to-FC. A protocol diagram is illustrated in Figure 1.5, which consists of the application layer, transport layer, network layer, data link layer, physical layer, power management plane, mobility management plane, and task management plane [1].

Each layer has their own functionalities and knows how to respond to the requests from the layer below or above. The main functions are summarized as follows:

- The application layer interacts with the end users and specifies how the data are requested.
- The transport layer sends and receives data upon request from the application layer. It is especially needed when the system needs to access through external networks.

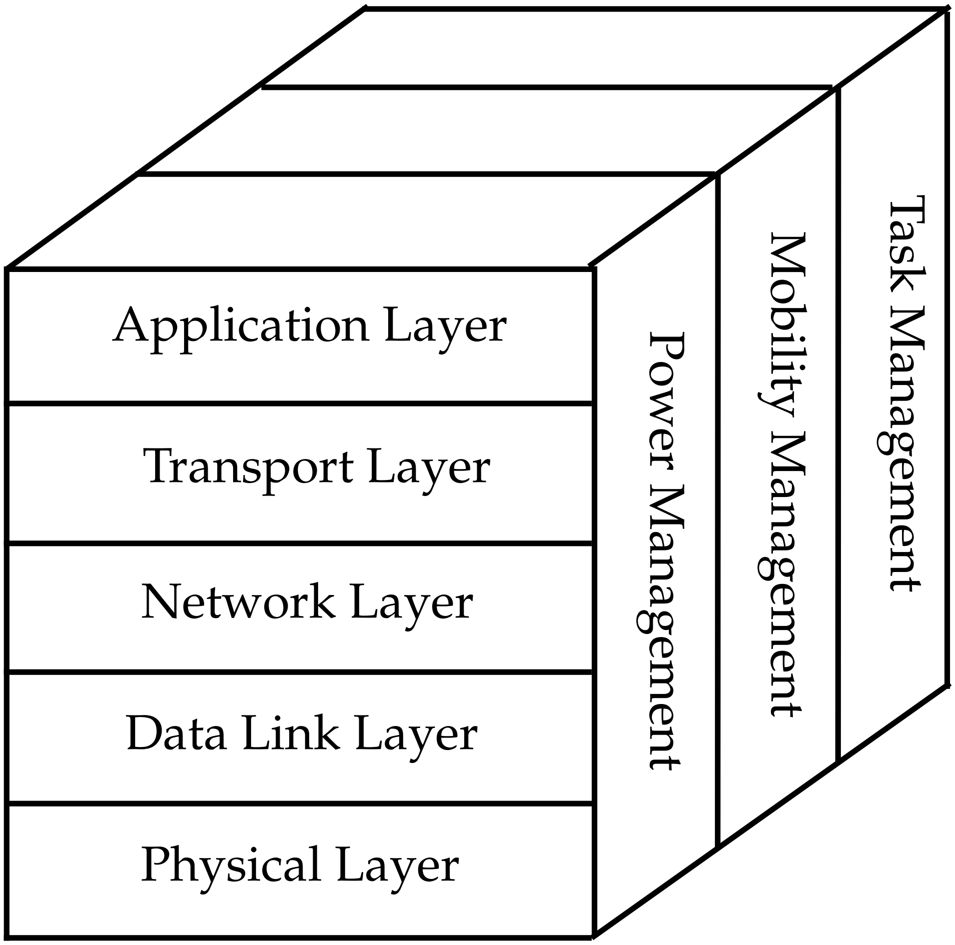

Figure 1.5: Wireless sensor network protocol stack. [1]

- The network layer routes the incoming data to the desired locations.

- The data link layer is responsible for the multiplexing of data streams, data frame detection, medium access and error control.

- The physical layer is responsible for frequency selection, carrier frequency generation, signal detection, modulation, and data encryption.

Meanwhile, the power, mobility and task management planes in Figure 1.5 monitor the power, movement and task among the nodes. The planes reduce the overall energy consumption and coordinate different tasks.

1.1.3 WSNs Applications

With topology and communication protocols in mind, the natural question is what kind of applications we can apply with WSNs. In fact, WSNs are widely employed in many applications such as environmental monitoring, cyber-physical systems, healthcare, diagnostics of complex systems and military applications and so on. We summarize the main applications as the following categories.

- Environmental monitoring includes temperature monitoring (forest fire detection), flood detection, geophysical research and so on [69, 76, 91]. Take forest fire detection as an example, where a large number of sensors are randomly and densely deployed to a forest in order to collect data on weather conditions including temperature, wind speed, rain and relative humidity. These sensors need to be durable in that they are often exposed to harsh environments. They send the compressed outputs to the FC through the wireless communication module, then the FC combines the information collected by local nodes and makes the final decision whether there is a fire or not in that forest [7].

- A cyber-phyiscal system is defined as the system where physical and software components are deeply intertwined, each operating on different spatial and temporal scales, exhibiting multiple and distinct behavioral modalities, and interacting with each other in a myriad of ways that change with context [40, 61]. WSNs play an important role in sensing and providing information for such systems including smart grids and nuclear power plants [54, 95]. For smart grids, a large number of sensors are distributed for monitoring long range power transmission lines in order to improve transmission efficiency, reliability and sustainability [26, 29].

- Health related applications include tracking patients' physiological conditions, movements and behaviors. For this purpose, patients usually wear different types of wireless sensors to collect data on body conditions [31, 102]. For real-time applications such as telemonitoring of patients, sensors transmit to the FC securely in real-time. For offline decision-making, such as future medical diagnostics, drug administration in hospitals and so on, sensors collect data for a long time and then securely transmit the data to the FC.

- For complex systems like vehicles, airplanes or nuclear plants, WSNs keep monitoring the conditions of the parts, the environment or the function units, once the fault or anomaly occurs, the sensors would report to the FC [46, 50, 64]. One example for this application is a WSN deployed in an airplane cabin to monitor particulate matter, carbon dioxide, pressure and humidity to make sure the environment is suitable for passengers [35].

- Military applications include battlefield surveillance and reconnaissance of opposing forces. WSNs can be deployed to detect, localize and track targets, moreover, they can be used to assess damage conditions, monitoring equipment and ammunition [62, 80].

From the above categories, we can see that WSNs have already changed our lives in many aspects, more importantly, WSNs also have the potential for many future applications, one of which is intelligent transportation systems, where sensors mounted on vehicles wirelessly communicate with other vehicles or infrastructure sensor nodes in order to improve the overall quality of road transportations including safety, congestion, emissions, and traffic waiting time [23, 42, 51]. The Internet of things is another application of WSNs which aims at connecting home appliances, smart phones and other Internet connected devices [4] in order to improve energy-efficiency, convenience, safety and so on.

1.2 Cyber Attacks in WSNs

Wireless communication makes the aforementioned applications possible, on the other hand, wireless communication allows local nodes to broadcast and all of their wireless packets are potentially available to any other listeners. It also means WSNs are vulnerable to all kinds of attacks. As the data collected by the aforementioned applications could be extremely sensitive, care must be taken to prevent the collected information from being leaked to any malicious third parties. Thus, we need to understand the potential strategies of attackers against WSNs in order to defend them effectively.

In [92], Wang et al. surveyed cyber threats in sensor networks, which is summarized in Table 1.2. According to the security requirements in WSNs, these attacks can categorized as [73]:

- Secrecy and authentication attacks where eavesdroppers either passively listen to packets or modify packets in order to gain certain advantages.
- Network availability attacks where attackers keep communication channels busy so that transmitters won't be able to send anything through the channel to receivers, e.g., jamming attackers, denial-of-service attackers (Table 1.2).

- Stealthy attacks against service integrity where attackers falsify the data and make the network accept it so that the decision center is confused, e.g., Byzantine attackers.

Among all of these attacks, this dissertation concentrates on eavesdropping attacks in that it forms the basis or starting point for a large number of different, more malicious attack strategies. For example, if Byzantine attackers, jammers or intruders have reliable information provided by the eavesdropper, their subsequent attacks could be more efficient [65]. There are two types of eavesdropping attacks, passive and active. Passive eavesdroppers detect the information by tapping the data transmissions between the local sensors and the legitimate user; active eavesdroppers, however, send queries to some local sensors by disguising themselves as friendly nodes [20]. In this dissertation, we consider the general problem of passive eavesdropping because it is the foundation of active eavesdropping and it is difficult to detect and defend.

1.3 Cyber Defense Mechanisms

Since this dissertation focuses on eavesdropping attacks against WSNs, we need to survey the available defending mechanisms against eavesdroppers and evaluate the feasibility of applying the algorithms to WSNs. In [92], authors mentioned that the standard cryptography algorithms may prevent eavesdroppers and protect the secrecy of the system. To further investigate this issue, we discuss whether the current research on cryptography for WSNs is fully practical due to the constraints on WSNs.

Table 1.2: Denial-of-service Attacks in WSNs [92]

Layer	Attacks	Defense
Physical	Jamming	Spread-spectrum, priority messages, lower duty cycle, region mapping, mode change
	Tampering	Tampering-proofing, hiding
Link	Collision	Error-correcting code
	Exhaustion	Rate limitation
	Unfairness	Small frames
Network	Spoofed, altered	Egress filtering, authentication, monitoring
	Selective forwarding	Redundancy, probing
	Sinkhole	Authentication, monitoring, redundancy
	Sybil	Authentication, probing
	Wormholes	Authentication, packet leashes by using geographic and temporal information
	Hello food attacks	Authentication, verify the bidirectional link
	Acknowledgement spoofing	Authentication
Transport	Flooding	Client puzzles
	Desynchronization	Authentication

1.3.1 Cryptography in WSNs

Cryptographic algorithms, which includes public key and symmetric key [37], have been widely used for computer networks where the nodes (computers) are powerful enough to implement the algorithms. However, due to the constraints in WSNs, where sensors are often operated on a limited battery power and limited computational power, many existing algorithms are not practical for use. Next, we discuss the feasibility of implementing the recent research on public key and symmetric key in WSNs.

Public Key Cryptography in WSNs

This asymmetric cryptography scheme uses pairs of keys: public keys which can be distributed widely, and private keys which are known only to the owner. Anyone can encrypt messages using the public key, however, only the owner of that paired private key can decrypt the message.

There are several popular public key algorithms such as the Diffie-Hellman key agreement protocol [21] or RSA signatures [67], however they are undesirable for WSNs due to the computational intensity and power consumption. One possible solution is elliptic curve cryptography algorithm [43, 57], which appears to offer equal security for a far smaller key size, thereby reducing processing and communication overhead. Some of the researchers implemented different ECC cryptography algorithms on microprocessors such as Atmel ATmega128 [30, 49, 90], however, the public key operations are still expensive for these processors, not to mention less powerful devices as the nodes of WSNs.

Symmetric Key Cryptography in WSNs

Unlike public key cryptography, symmetric key scheme uses the same cryptography keys for both encryption and decryption. The keys can be identical or a simple transformation applied between the two keys. This scheme consumes much less computational energy. In order to investigate the feasibility of symmetric key for WSNs, several popular algorithms including RC4 [56], RC5 [68], IDEA [56], SHA-1 [25] and MD5 [56, 66] were implemented on different microprocessors ranging in word size from 8-bit to 16-bit. The researchers compared the operation time and energy with these algorithms and they concluded that symmetric key cryptography is preferred in a WSN. The measurements on average execution time and energy consumption with different algorithms on Atmel ATmega128 processor are summarized in Table 1.3 and Table 1.4, respectively.

Table 1.3: Symmetric Key Cryptography: Average Execution Times on Atmel ATmega128 [38]

Algorithm	Operation Time (ms)
RC5 [68]	0.26ms
Skipjack [59]	0.38ms

Table 1.4: Symmetric Key Cryptography: Average Energy Consumption on Atmel ATmega128 [90]

Algorithm	Energy
SHA-1 [25]	5.9 mJ/byte
AES-128 Enc/Dec [19]	1.62/2.49 mJ/byte

1.3.2 Physical-Layer Security Approach

Even though the symmetric key cryptography algorithms presented in section 1.3.1 consume low-power, they may not be low enough for long term WSNs operations, furthermore, they do require the devices to have the computational capability to perform the required tasks which may not be true for some of the nodes [98]. Therefore, it is not always possible to completely rely on cryptographic techniques. Besides, key distribution brings another problem to WSNs, especially for dense WSNs. To address these issues, information-theoretic (physical-layer) security approaches, utilizing the characteristics of the physical layer, including transmission channels noises, and the information of the source, have gained considerable attention on this method to enhance the security, secrecy and privacy of WSNs [3, 55, 78, 103]. Additionally, physical-layer security for distributed detection is scalable due to the low computational complexity [74]. Physical-layer security approaches can be used along with cryptosystems to further enhance WSNs and make the systems even more secure.

Based on a physical-layer security approach, several attempts were made to combat eavesdropping attacks for WSNs under the assumption that an attacker has partial or full access to sensor outputs [36]. Nadenla et al. considered the secrecy problem in distributed detection against eavesdropping attacks for WSNs in parallel networks, where the goal was to maximize Kullback-Leibler Divergence (KLD) at the FC for one sensor, D_F, under the constraint that KLD at an eavesdropper for one sensor, D_E, is no more than a pre-specified threshold T_E [58]. For a two-sensor network, where the attacker has the access to one of the sensors output, Li et al. jointly designed sensor decision rules and fusion rules to maximize the FC detection probability by constraining both the FC's probability of false alarm and eavesdropper's detection probability [48]. For privacy issues, Li and Oechtering formulated privacy-constrained and privacy-concerned optimization problems under Bayesian framework and derived the optimal privacy detection rule under a privacy guarantee constraint [47].

As for the secrecy constrained distributed estimation in WSNs, Aysal et al. proposed to solve the problem by adding a stochastic cipher as a security module, to randomly change the sensor outputs and disguise them from the eavesdropper [2]. Guo et al. considered using multiple-input multiple-output beamforming strategies to combat eavesdroppers, where local sensors use the analog amplify and forward scheme to communicate with the FC over a slow-fading orthogonal multiple access channel [28]. In [41], Khan and Stanković proposed to securely estimate distributed data in cyber-physical systems by verifying statistical consistency on the nodal, local information and physical layer feedback.

Notice that the aforementioned efforts did not focus on the maximum achievable inference performance trade-off, nor did they explore the possibility of asymptotic perfect secrecy. For a specific channel model, Marano et al. designed sensor rules for WSNs under the perfect secrecy constraint such that the eavesdropper gains no information from the observations on the channel activities without direct access to sensors outputs [53]. However, the channel model considered in that paper is constrained and cannot be directly employed under more general noisy channel models.

1.4 Contributions and Overview of the Thesis

This dissertation focuses on using a physical-layer security approach to address the distributed inference problems with secrecy constraints in a sense that a WSN with parallel topology is eavesdropped by a global, greedy and informed eavesdropper, which has access to all the sensors outputs. The reason we consider parallel structure for WSNs lies in that it is simple and robust to sensor failures, when a small portion of sensor dies, the performance of the network would not be deteriorated. As a malicious user, this eavesdropper passively listens to the sensor outputs and aims at making informative decisions. However, the data collected by sensors are extremely sensitive, our goal is to prevent a malicious third party (eavesdropper) from stealing information from local nodes. Therefore, the ideal design for a sensor network is perfect secrecy where an eavesdropper does not obtain any useful information. We will discuss the possibility of (asymptotic) perfect secrecy. Moreover, we investigate performance trade-offs between the FC and eavesdropper, where the performance of the attacker is constrained to a level such that she could not make an informative inference; meanwhile, the performance of the legitimate user (FC) is guaranteed to perform well at the desired level. Utilizing the metrics of measuring secrecy in both detection and estimation problems, we provide results on the maximum achievable inference performance trade-off between the FC and eavesdropper.

This dissertation is organized as follows,

Chapter 2: Background and Fundamental Concepts

In this chapter, we introduce the key concepts in understanding the materials in the following chapters. The fundamental concepts include distributed detection, distributed estimation and secrecy metrics under different frameworks.

Chapter 3: Secrecy Constrained Distributed Detection in WSNs

We first investigate detection problems under secrecy constraints. The main contributions of secrecy constrained distributed detection in WSNs are summarized as follows:

1. Analyze the detection performance at the FC and eavesdropper, respectively. For the case where the sensor outputs are binary, we evaluate the quality of the received sensor decisions when the sensors employ likelihood ratio quantizer (LRQ) close to the extreme points on receiver operating characteristics (ROC) curve.

2. Utilizing performance analysis, we propose a novel approach of analyzing the performance trade-off between the FC and eavesdropper using the maximum achievable detection performance ratio between the FC and eavesdropper, given both a noise free and noisy FC channel. Additionally, we show that both asymptotic perfect secrecy and asymptotic perfect detection are possible by increasing the number of sensors when the FC has noiseless channels under the Neyman-Pearson framework.

3. Under the Bayesian framework, we analyze the performance in terms of probability of error, where the detectability of an eavesdropper can be limited to a level where she can only rely on her prior information. The limit of optimal FC detection performance is derived for the performance trade-off analysis. Using the approximated asymptotic error exponent we obtained for both the FC and eavesdropper, we show that both asymptotic perfect secrecy and asymptotic

perfect detection are possible. The results contradict the idea that network security tends to decrease as the number of sensors increases.

Chapter 4: Secrecy Constrained Distributed Estimation in WSNs

In this chapter, we investigate the secrecy constrained distributed estimation problem and the main contributions are summarized as follows:

1. Under classical settings, where the parameter to be estimated is fixed but unknown, we analyze the estimation performance at the FC and eavesdropper using Fisher information, respectively. In order to investigate the possibility of perfect secrecy, we propose the Fisher information ratio between the FC and eavesdropper. Furthermore, for Gaussian noise, we show how to design the threshold in order to achieve asymptotic perfect secrecy and asymptotic perfect estimation.

2. Under the Bayesian framework, where the parameter is a random variable, we analyze the performance trade-off between the FC and eavesdropper using Fisher information and show that the secrecy constraints can be satisfied for both the FC and eavesdropper under Gaussian noise case.

Chapter 5: Secrecy Constrained Distributed Inference with Parallel Fading Binary Symmetric Channel Models

In this chapter, we consider secrecy constrained detection and estimation problem with binary phase-shifting keying modulation in parallel Rayleigh fading binary symmetric channels. Similarly, we investigate the performance ratio between the FC and eavesdropper. We analyze the maximum achievable performance ratio and show that the number of sensors does not affect this ratio.

Chapter 6: We conclude in this chapter and discuss future research related to secrecy constrained distributed inference.

CHAPTER 2

BACKGROUND AND BASIC CONCEPTS

2.1 Statistical Inference

In the classical statistical inference, all the data is collected and processed in a centralized fashion. Distributed inference, however, detects signal presence, estimates parameters and tracks targets based on distributed data from local sensors [83, 86]. It has been the focus of multiple disciplinary research in the past several decades [6, 10–12, 81, 85, 89]. One of the essential problems in distributed inference is to optimize decision-making at the information center by the design of local decision sensor rules for each sensor and global decision rules at an information center [83]. Without constraints on “distributed” settings, the problems of inference share much in common with many centralized statistical inference and learning problems such as signal detection and estimation, dimension reduction and feature extraction [60]. Due to the additional condition on “distributed”, the complexity of the inference problem is increased significantly [82].

Distributed inference includes distributed detection, distributed estimation and tracking. One of the main differences between detection and estimation is the phenomenon to be inferred by sensors. In distributed detection, the phenomenon observed by sensors is discrete, e.g., binary hypothesis testing, where one aims to decide between two potential hypotheses, $H \in (H_0, H_1)$. In distributed estimation, the phenomenon is often a parameter in a continuous set [36]. In the following sections, we introduce the basic settings in distributed detection and distributed estimation.

2.2 Distributed Detection

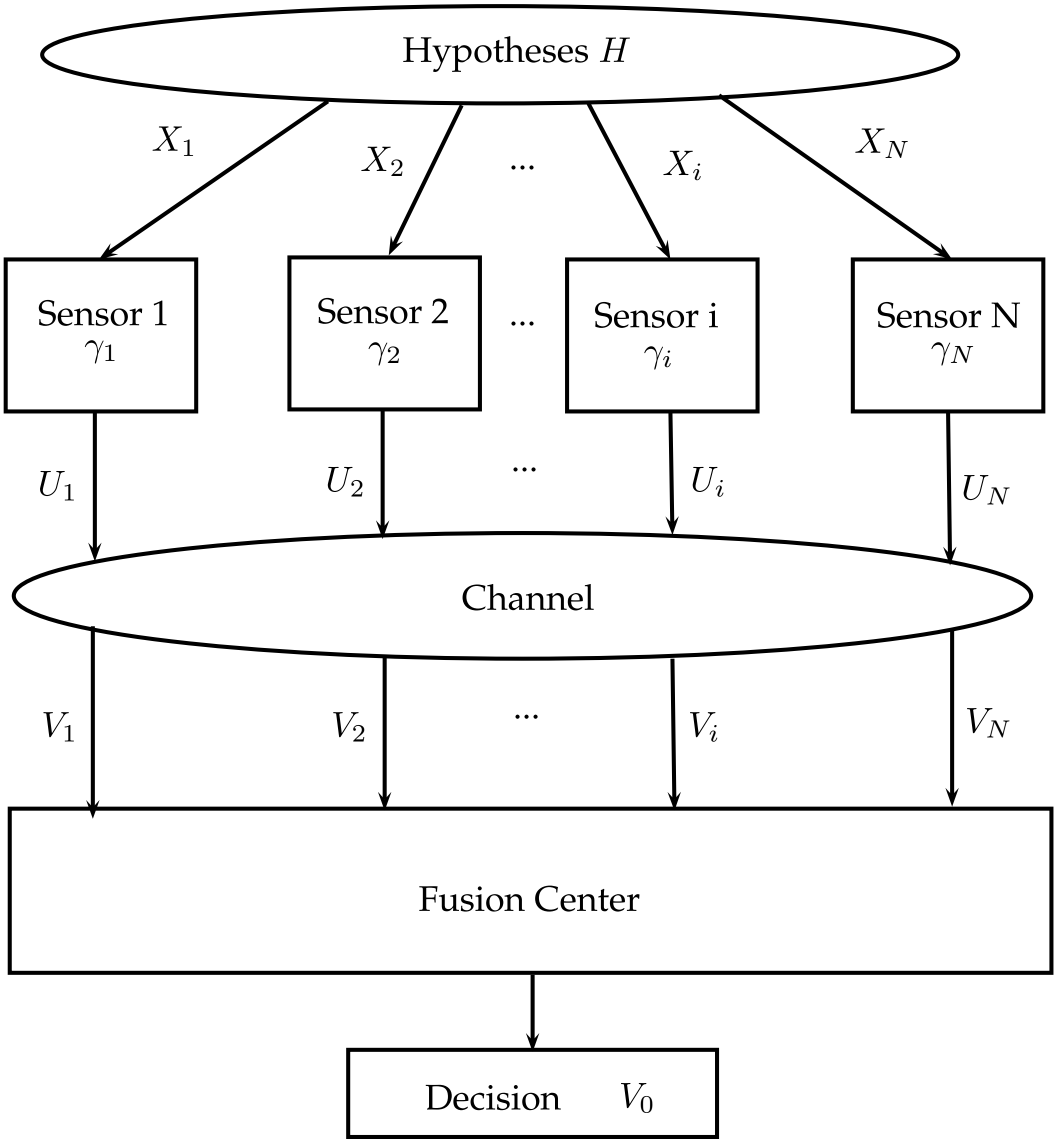

Figure 2.1: Distributed Detection

As one of the essential aspects of distributed inference, distributed detection is often the initial goal of a pattern recognition system and aims at detecting signals or events as accurately as possible [86] with the distributed data collected by various sensors, where the data can be generated from the underlying binary or M-ary hypotheses. Distributed detection can be widely used for both military and civilian applications including distributed array radar, intruder detection, anomaly detection and intelligent transportation system where the infrastructure sensors detect pedestrians, vehicles and anomaly events [23, 42, 51]. For instance, N sensors are densely deployed in forests to observe the temperatures, and through a channel, these nodes send the quantized outputs to the FC where the final decision is made about whether there is forest fire or not [69]. For WSNs, detecting the presence of an event is the priority of all the other tasks including estimation, tracking and learning [11]. Hence, as a key function in WSNs, distributed detection has been an important and active research area over the past several decades [6, 10–12, 14, 77, 81, 85, 89].

In Figure 2.1, we show the structure of distributed detection in a parallel WSN, where local sensors observe the hypotheses H and obtain their data X_i, $(i = 1 \ldots N)$. With the decision rules for each sensor, γ_i, sensor i compresses the data to the outputs U_i, which is transmitted across a channel. In the end, the FC makes the decision V_0 based on the received V_is, the output of channels from the input U_is.

2.3 Distributed Estimation

If the presence of an object, a signal or an event is determined by the detection function in WSNs, more complicated tasks such as estimation and tracking can be performed. For instance, if a vehicle is detected by an intelligent transportation

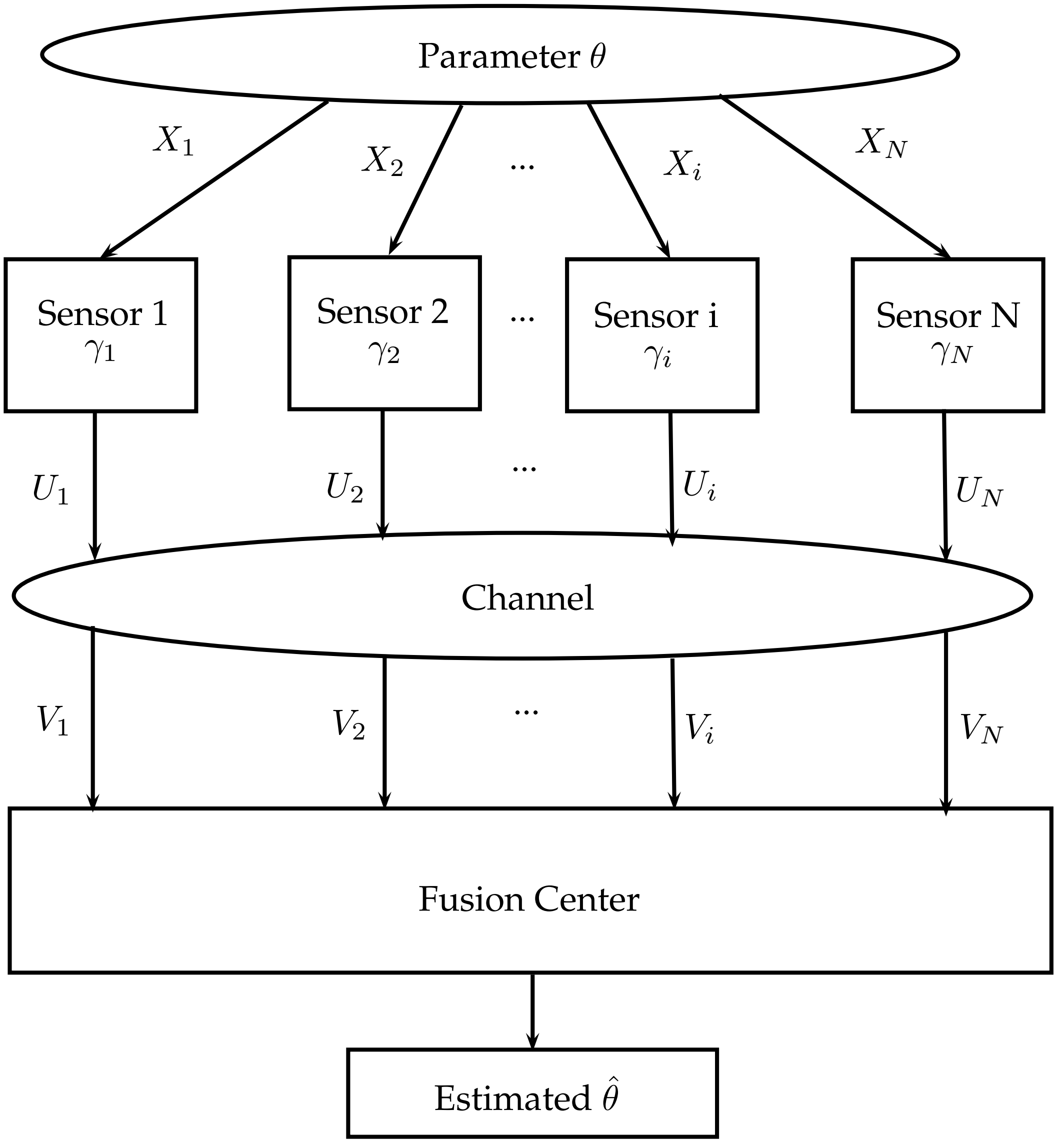

Figure 2.2: Distributed Estimation

system, the following task would be estimating how fast the vehicle is moving and where it is moving.

Aiming to estimate the values of a group of parameters based on a network of collaborating sensors, distributed estimation has been an important and active research area over the past several decades [9, 13, 27, 94].

Similar to the distributed detection setting, in Figure 2.2, we present the structure of distributed estimation, where sensors observe a scalar or vector parameter θ, sensors quantized outputs U_i $(i = 1 \dots N)$ are sent to the FC through a channel. Then the parameter $\hat{\theta}$ is estimated at the FC based upon received V_i.

2.4 Performance and Secrecy Metrics

From the aforementioned applications about distributed detection and distributed estimation, we can see that the information collected by the systems is very sensitive and care must be taken to prevent them from being leaked to any malicious third parties. Hence, we focus on secrecy constrained inference in WSNs where the ultimate goals are restricting the ability of eavesdropping from attackers and maintaining high performance at the FC. Hence, we first introduce secrecy.

Secrecy in WSNs against eavesdropping attacks means that any malicious listeners should not be able to make informative decisions based on messages from local sensors that are supposed to go to the FC. In other words, in distributed inference, secrecy measures the inference performance at the FC and eavesdropper respectively. For instance, if the inference performance at the FC is higher than the specified level while eavesdropper's performance is lower than a random guess, the WSN is considered as secure in terms of secrecy. For this purpose, we introduce the performance metrics for distributed detection and distributed estimation, respectively, in this section.

2.4.1 Distributed Detection under Neyman-Pearson Framework: Information Divergence

For secrecy constrained distributed detection under Neyman-Pearson framework, we consider information divergence as the performance metric. Information divergence maps the dissimilarity between two probability distributions to nonnegative values. It is also extended to machine learning problems where the goal is to minimize the approximation error between the observed data and the approximated model [22]. There are several information divergences and they are summarized in Table 2.1, where $\mathbf{x} > 0$ is the observed data and $\boldsymbol{\mu}$ is the approximation given by the model. For γ-divergence and Rényi-divergence, the input data needs to be normalized, where $\tilde{x}_i = x_i / \sum_j x_j$ and $\tilde{\mu}_i = \mu_i / \sum_j \mu_j$.

Since there are so many choices of information divergence, which one should we consider? According to Stein's lemma [12] and large deviation theory [8, 16], when the decision center observations are i.i.d., the error exponent of probability of missed detection (P_m) is bounded, a special case of γ-divergence, Kullback-Leibler divergence (KLD), $D(p_0(\cdot)||p_1(\cdot))$, where p_0, p_1 are the pdf under H_0 and H_1 hypotheses, respectively. Specifically, $-\lim\limits_{N\to\infty} \frac{1}{N} \log P_m \leq D(p_0(\cdot)||p_1(\cdot))$ (N is the number of sensors in a WSN) when the false alarm probability (P_f) is constrained to be less than a constant, and the equality can be achieved by the optimal LRT or other asymptotic optimal detectors such as type based detectors so that [18],

$$P_m \approx e^{-ND(p_0(\cdot)||p_1(\cdot))}. \tag{2.1}$$

For binary sensor decisions with $P(U_i = 1; H_0) = \alpha$ and $P(U_i = 1; H_1) = \beta$, we have $P(U_i = 0; H_0) = 1 - \alpha$ and $P(U_i = 0; H_1) = 1 - \beta$, the KLD [44] for each sensor is

$$D\left(p_0||p_1\right) = \alpha \log \frac{\alpha}{\beta} + (1-\alpha) \log \frac{(1-\alpha)}{(1-\beta)} = D\left(\alpha, \beta\right). \tag{2.2}$$

It is also true when P_m is constrained to be a constant,

$$P_f \approx e^{-ND(p_1(\cdot)||p_0(\cdot))}. \tag{2.3}$$

The corresponding KLD is

$$D\left(p_1||p_0\right) = \beta \log \frac{\beta}{\alpha} + (1-\beta)\log\frac{(1-\beta)}{(1-\alpha)} = D\left(\beta,\alpha\right). \tag{2.4}$$

2.4.2 Distributed Detection under Bayesian Framework: Probability of Error

Under Bayesian framework, prior information needs to be taken into consideration. Let the risk function $\lambda(a_i|H_j)$ be the risk or loss incurred for taking action a_i when the actual hypothesis is H_j, where $i \in [0, \ldots, N]$, and N indicate the number of possible actions, and $j \in [0, \ldots, C]$, C is the number of states of nature (categories) [24]. The overall risk is

$$r = \sum_{i=0}^{N}\sum_{j=0}^{C} \lambda\left(a_i|H_j\right) P(a_i|H_j)P(H_j),$$

where $P(a_i|H_j)$ is the probability of action i given the state of nature H_j, $P(H_j)$ is the probability of category H_j. For $C = 1$, two-category case, to simplify the notation, let $\lambda_{ij} = \lambda\left(a_i|H_j\right)$, $P(H_j) = P_j$ and the actions be,

$$a_0 : \text{decide } \ H_0$$
$$a_1 : \text{decide } \ H_1$$

The overall risk is

Table 2.1: Information Divergence [22]

Name	Definition	Special Cases
α-divergence $D_\alpha\left(\mathbf{x}\|\|\boldsymbol{\mu}\right)$	$\frac{\sum_i x_i^\alpha \mu_i^{1-\alpha} - \alpha x_i + (\alpha-1)\mu_i}{\alpha(\alpha-1)}$	$D_{\alpha=2}\left(\mathbf{x}\|\|\boldsymbol{\mu}\right) = \frac{1}{2}\sum_i \frac{(x_i-\mu_i)^2}{\mu_i}$
		$D_{\alpha\to 1}\left(\mathbf{x}\|\|\boldsymbol{\mu}\right) = \sum_i \left(x_i \ln \frac{x_i}{\mu_i} - x_i + \mu_i\right)$
		$D_{\alpha=\frac{1}{2}}\left(\mathbf{x}\|\|\boldsymbol{\mu}\right) = 2\sum_i \left(\sqrt{x_i} - \sqrt{\mu_i}\right)^2$
		$D_{\alpha\to 0}\left(\mathbf{x}\|\|\boldsymbol{\mu}\right) = \sum_i \left(\mu_i \ln \frac{\mu_i}{x_i} - \mu_i + x_i\right)$
		$D_{\alpha=-1}\left(\mathbf{x}\|\|\boldsymbol{\mu}\right) = \frac{1}{2}\sum_i \frac{(x_i-\mu_i)^2}{x_i}$
β-divergence $D_\beta\left(\mathbf{x}\|\|\boldsymbol{\mu}\right)$	$\frac{\sum_i x_i^{\beta+1} + \beta\mu^{\beta+1} - (\beta+1)x_i\mu_i^\beta}{\beta(\beta+1)}$	$D_{\beta=1}\left(\mathbf{x}\|\|\boldsymbol{\mu}\right) = \frac{1}{2}\sum_i \left(x_i - \mu_i\right)^2$
		$D_{\beta\to 0}\left(\mathbf{x}\|\|\boldsymbol{\mu}\right) = \sum_i \left(x_i \ln \frac{x_i}{\mu_i} - x_i + \mu_i\right)$
		$D_{\beta\to -1}\left(\mathbf{x}\|\|\boldsymbol{\mu}\right) = \sum_i \left(\frac{x_i}{\mu_i} - \ln \frac{xi_i}{\mu_i} - 1\right)$
		$D_{\beta=-2}\left(\mathbf{x}\|\|\boldsymbol{\mu}\right) = \sum_i \left(\frac{x_i}{2\mu_i^2} - \frac{1}{\mu_i} + \frac{1}{2x_i}\right)$
γ-divergence $D_\gamma\left(\mathbf{x}\|\|\boldsymbol{\mu}\right)$	$\frac{1}{\gamma(1+\gamma)} \ln \left(\sum_i x_i^{\gamma+1}\right)$ $+ \frac{1}{(1+\gamma)} \ln \left(\sum_i \mu_i^{\gamma+1}\right)$ $- \frac{1}{\gamma} \ln \left(\sum_i x_i \mu_i^\gamma\right)$	$D_{\gamma\to 0}\left(\tilde{\mathbf{x}}\|\|\tilde{\boldsymbol{\mu}}\right) = \sum_i \tilde{x}_i \ln \frac{\tilde{x}_i}{\tilde{\mu}_i}$
Rényi-divergence $D_\rho\left(\mathbf{x}\|\|\boldsymbol{\mu}\right)$	$\frac{1}{\rho-1} \ln \left(\sum_i \tilde{x}_i^p \tilde{\mu}_i^{1-p}\right)$ where $\tilde{x}_i = x_i / \sum_j x_j$, $\tilde{\mu}_i = \mu_i / \sum_j \mu_j$	

$$\begin{aligned} r = {} & \lambda_{00} P\left(\text{decide }\ H_0|H_0\right) P_0 + \lambda_{01} P\left(\text{decide }\ H_0|H_1\right) P_1 + \\ & \lambda_{10} P\left(\text{decide }\ H_1|H_0\right) P_0 + \lambda_{11} P\left(\text{decide }\ H_1|H_1\right) P_1 \\ & = \lambda_{00}\left(1 - P_f\right) P_0 + \lambda_{01} P_m P_1 + \lambda_{10} P_f P_0 + \lambda_{11}\left(1 - P_m\right) P_1 \\ & = \lambda_{00} P_0 + \lambda_{11} P_1 + \left(\lambda_{10} - \lambda_{00}\right) P_f P_0 + \left(\lambda_{01} - \lambda_{11}\right) P_m P_1 \end{aligned}$$

Since $\lambda_{00} P_1$ and $\lambda_{11} P_0$ are constant, we can put them aside. Therefore, the overall risk function is reduced to

$$r = \left(\lambda_{10} - \lambda_{00}\right) P_f P_0 + \left(\lambda_{01} - \lambda_{11}\right) P_m P_1$$

Then we normalize r by

$$\begin{aligned} \frac{r}{\left(\lambda_{10} - \lambda_{00}\right) P_0 + \left(\lambda_{01} - \lambda_{11}\right) P_1} = {} & \frac{\left(\lambda_{10} - \lambda_{00}\right) P_0}{\left(\lambda_{10} - \lambda_{00}\right) P_0 + \left(\lambda_{01} - \lambda_{11}\right) P_1} P_f \\ & + \frac{\left(\lambda_{01} - \lambda_{11}\right) P_1}{\left(\lambda_{10} - \lambda_{00}\right) P_0 + \left(\lambda_{01} - \lambda_{11}\right) P_1} P_m \end{aligned}$$

Let

$$\begin{aligned} \pi_0 &= \frac{\left(\lambda_{10} - \lambda_{00}\right) P_0}{\left(\lambda_{10} - \lambda_{00}\right) P_0 + \left(\lambda_{01} - \lambda_{11}\right) P_1} \\ \pi_1 &= \frac{\left(\lambda_{01} - \lambda_{11}\right) P_1}{\left(\lambda_{10} - \lambda_{00}\right) P_0 + \left(\lambda_{01} - \lambda_{11}\right) P_1}, \end{aligned}$$

we have

$$P_e = \pi_0 P_f + \pi_1 P_m \tag{2.5}$$

With the detection performance KLD and probability of error, we can investigate secrecy constrained detection in Chapter 3.

2.4.3 Distributed Estimation under Classical Setting: Mean Squared Error and Fisher Information

For estimation problems under a classical setting, a natural criterion for evaluating the performance is the mean squared error (MSE), which is defined in Equation (2.6).

$$\text{MSE} = E\left(\hat{\theta} - \theta\right)^2 \tag{2.6}$$

where, θ is a scalar parameter and $\hat{\theta}$ is the estimated parameter. We will use MSE evaluation as the performance metric when feasible.

However, sometimes computing MSE is not straightforward and even intractable for some cases. Instead, Cramér-Rao lower bound (CRLB) [39, 79] is often used which is equivalent to evaluating the Fisher information (FI),

$$\mathbf{I}(\mathbf{V};\theta) \triangleq E_{\mathbf{V}}\left(\frac{\partial^2 \log p(\mathbf{V};\theta)}{\partial^2\theta}\right) \tag{2.7}$$

where $\mathbf{V}$ is the data transmitted from local sensors across the channel (Figure 2.2), $p(\mathbf{V};\theta)$ is probability density function (PDF) of parameter θ given $\mathbf{V}$ [39]. And the MSE is bounded away from CRLB for scalar parameter θ is,

$$\text{MSE} \geq \text{CRLB}(\mathbf{V};\theta) = \frac{1}{\mathbf{I}(\mathbf{V};\theta)}.$$

2.4.4 Distributed Estimation under Bayesian Setting: Bayesian Cramér-Rao Lower Bound

Under Bayesian framework, however, we have to take the prior information into consideration, therefore, Bayesian CRLB is defined as,

$$\text{MSE} \geq \text{BCRLB}_F\left(\mathbf{V};\theta\right) = \left(\mathbf{I}\left(\mathbf{V};\theta\right)\right)^{-1} = \left(\int_{-\infty}^{\infty} NI(\eta,\theta,\rho_F)p\left(\theta\right)d\theta + I(\lambda)\right)^{-1} \tag{2.8}$$

where

$$I(\lambda) = \int \left(\frac{\partial \log p\left(\theta\right)}{\partial\theta}\right)^2 p(\theta)d\theta,$$

and $\mathbf{V}$ is the same with the one defined in (2.7) and $p\left(\theta\right)$ is the prior density about the random variable θ.

Similarly, using estimation performance metrics, we can study the performance trade-off, asymptotic perfect secrecy and asymptotic perfect estimation in Chapter 4.

CHAPTER 3

SECRECY CONSTRAINED DISTRIBUTED DETECTION IN WSNS

This chapter is organized as follows. In Section 3.1 and Section 3.2, we introduce the system model, detection performance metric, and set the secrecy constraints under both frameworks in WSNs, respectively. In Section 3.3, we solve the secrecy constrained problem under Neyman-Pearson framework and explain how to achieve asymptotic perfect secrecy in detection. We then analyze the secrecy constrained distributed detection problem under the Bayesian framework in Section 3.4. In Section 3.5, we provide simulation results to further support our proofs.

3.1 Distributed Detection in Sensor Networks

3.1.1 WSN Model

We consider a distributed detection problem with binary hypotheses, H_0 and H_1, in a parallel WSN as shown in Figure 3.1. The key components of our research problem are described as follows:

1. **Network topology**. The SN consists of N sensors connected in parallel to a FC via a set of parallel accessible channels. Instead of considering a multi-hop network, we adopt a parallel structure because even for a multi-hop network,

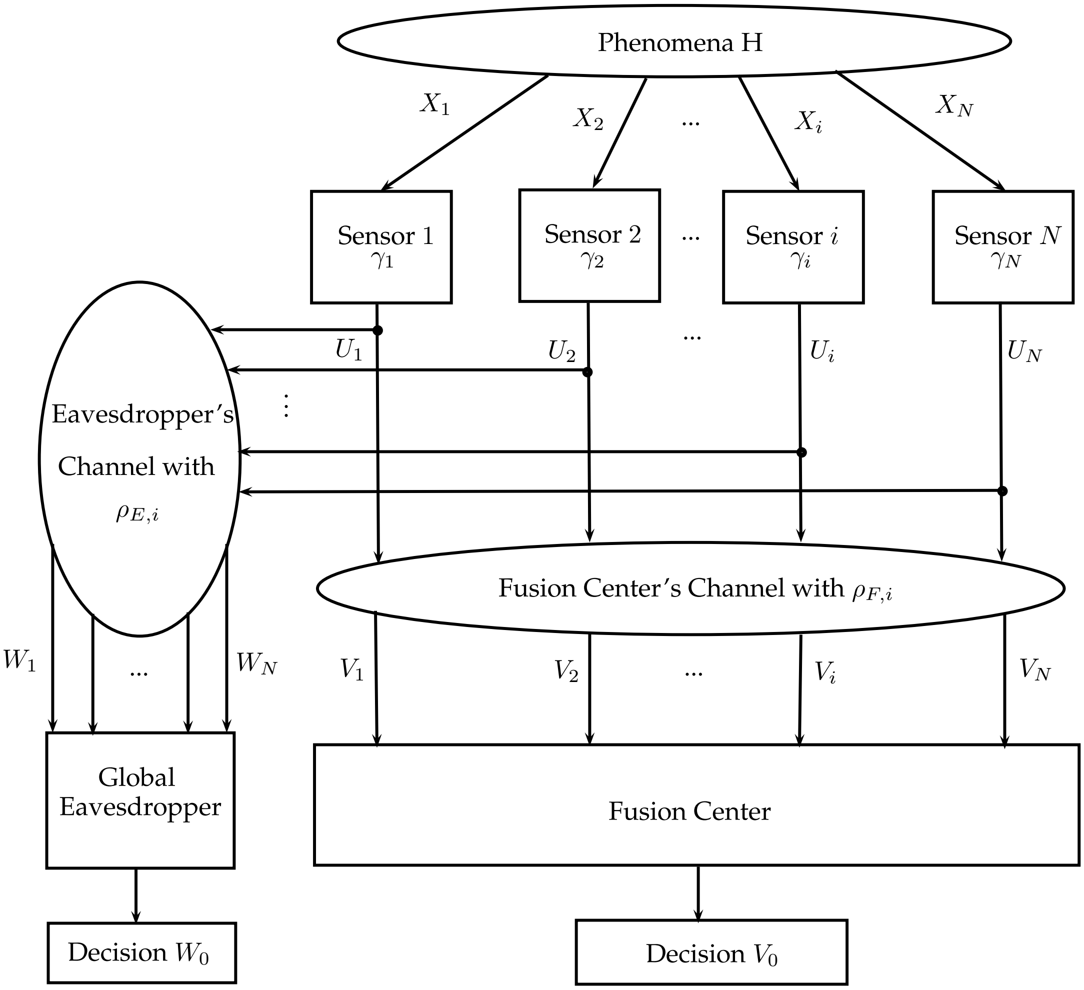

Figure 3.1: The model of a parallel sensor network under the attacks of an informed and greedy eavesdropper who eavesdrops on all the sensors decisions $(i = 1 \dots N)$ that are transmitted wirelessly via a binary symmetric channel with bit error rate $\rho_{E,i}$. The legitimate user receives sensor i data through another binary symmetric channel with bit error rate $\rho_{F,i} < \rho_{E,i}$.

the parallel structure can still be carried out virtually by leaving relay nodes to forward all sensor outputs.

2. **Sensor observations and sensor outputs**. $\mathbf{X} = [X_1, \dots, X_N]$ are the sensor observation, for each X_i, can be a random variable or a random vector. Next, $p_k(X_i) = p(X_i; H_k)$ is the probability density function (pdf) under H_k

at sensor i, respectively, where $k = 0, 1$ and $i = 1, 2, \ldots, N$. We assume that $p_0(X_i)$ and $p_1(X_i)$ are continuous pdfs with no point mass. The log-likelihood ratio $\ln\left(p_1(X_i)/p_0(X_i)\right)$ is assumed to be unbounded. Sensor i makes a binary decision $U_i \in \{0, 1\}$ based on its decision rule γ_i, such that $P(U_i = 1|X_i) = \gamma_i(X_i) \in [0, 1]\,,\ \forall i$.

3. **Channel model**. The communication channel between sensor i and its target receiver is assumed to be a binary symmetric channel (BSC), a channel model widely employed in SN communications for binary coding schemes such as binary phase-shift keying (BPSK) [45, 71, 101]. This model also serves as a good starting point to study other more complicated channel models. Sensor i sends its quantized output U_i to the FC over a BSC with bit error rate (BER) $\rho_{F,i} < \frac{1}{2}$, with a received decision, V_i.

4. **Attack model**. All of the sensors outputs are eavesdropped by Eve via a set of parallel wiretapping channels. Eve receives $W_i \ \ (i = 1, \ldots, N)$, from sensor i as an output of a BSC channel with BER $\rho_{E,i} < \frac{1}{2}$. We assume that Eve's channel is noisier than the FC's such that $\rho_{E,i} > \rho_{F,i}$, which can be achieved by using directional antennas to improve the FC SNR, resulting in a lower BER [52, 75, 99]. Note, an analysis similar to what follows can be employed for different channel models. Other than receiving a set of different observations $\mathbf{W} = [W_1, \ldots, W_N]$, Eve is assumed to have the *same information* about the detection algorithm as the FC does, including the sensor observation model, sensor decision rule, channel status and the prior probabilities of hypotheses, $P(H_0) = \pi_0$ and $P(H_1) = \pi_1$.

5. **Identical and conditional independence assumption**. We assume that

sensors observations and the communication channels are conditionally independent and identically distributed (i.i.d.). Specifically,

$$p(X_1, X_2, \ldots, X_N; H_j) = \prod_{i=1}^{N} p(X_i; H_j), \ \ j = 0, 1,$$

for the sensor observations, and $\rho_E = \rho_{E,i}$ with $\rho_F = \rho_{F,i}$ for all i for the communication channels.

3.1.2 Received Decision Qualities

For its simplicity and robustness, we employ identical sensor design in this chapter. That is, the decision rule $\gamma_i(\cdot)$ at sensor i is a likelihood ratio quantizer such that

$$\gamma_i(x) = \gamma(x) = \begin{cases} 1 & \frac{p_1(x)}{p_0(x)} \geq \eta \\ \\ 0 & \frac{p_1(x)}{p_0(x)} < \eta. \end{cases} \tag{3.1}$$

Under the conditional i.i.d. assumption, it has been shown that the identical sensor decision rule design, where each sensor uses the same likelihood ratio test (LRT) with the same threshold, is at least asymptotically optimal at the FC (i.e., no eavesdropper) [17, 81, 93].

At the ith local sensor, the resulting probability of false alarm α_i, and the probability of detection β_i are given by [84]

$$\alpha_i = P\left(U_i = 1 | H_0\right) = P\left(\frac{p_1\left(X_i\right)}{p_0\left(X_i\right)} \geq \eta | H_0\right),$$

$$\beta_i = P(U_i = 1|H_1) = P\left(\frac{p_1(X_i)}{p_0(X_i)} \geq \eta|H_1\right).$$

and

$$\frac{d\beta_i}{d\alpha_i} = \eta. \tag{3.2}$$

Therefore, since $\ln(p_1(X_i)/p_0(X_i))$ is unbounded, then $\eta \to \infty$ as $\alpha_i, \beta_i \to 0$, or $\eta \to 0$ as $\alpha_i, \beta_i \to 1$. Because of the i.i.d. assumption on the observations, decision rules and channels, we have

$$\begin{aligned}\alpha &= \alpha_1 = \alpha_2 \cdots = \alpha_N \\ \beta &= \beta_1 = \beta_2 \cdots = \beta_N.\end{aligned}$$

Due to the binary symmetric channel between the local sensors and the FC, the received decision, V_i, from sensor i at the FC, has the following performance,

$$\begin{aligned}P(V_i = 1|H_0) = \alpha_F &= \alpha(1-\rho_F) + (1-\alpha)\rho_F \\ &= \rho_F + (1-2\rho_F)\alpha, \\ P(V_i = 1|H_1) = \beta_F &= \beta(1-\rho_F) + (1-\beta)\rho_F \\ &= \rho_F + (1-2\rho_F)\beta.\end{aligned} \tag{3.3}$$

Similarly, the received decision, W_i, at eavesdropper, has the following performance,

$$\begin{aligned} P(W_i = 1|H_0) = \alpha_E = \rho_E + (1 - 2\rho_E)\alpha, \\ P(W_i = 1|H_1) = \beta_E = \rho_E + (1 - 2\rho_E)\beta. \end{aligned} \tag{3.4}$$

3.2 Secrecy in Distributed Detection

With the model of the WSN, we introduce the performance metrics that lead to secrecy constraints in distributed detection under both frameworks. The first performance metric applicable under the Neyman-Pearson framework is the KLD.

3.2.1 Performance Metric and Secrecy Constraints under Neyman-Pearson framework

When the decision center's observations are i.i.d. and the probability of false alarm, $P_f =$ (decide $H_1|H_0$), is constrained to be no greater than a fixed constant, it is known that the error exponent of the probability of missed-detection, $P_m =$ (decide $H_0|H_1$), is bounded by the corresponding KLD, $D\left(p_0(\cdot)||p_1(\cdot)\right)$ [44], between the p_0, the pdf under H_0, and p_1, the pdf under H_1, such that [8, 12, 16]

$$-\lim_{N\to\infty}\frac{1}{N}\ln P_m \le D\left(p_0(\cdot)||p_1(\cdot)\right) = E_{p_0(\cdot)}\left(\frac{dp_0(\cdot)}{dp_1(\cdot)}\right) \tag{3.5}$$

Notice that equality in (3.5) can be achieved via optimal LRT detectors or other asymptotically optimal detectors such as type based detectors [18]. Similarly, $D\left(p_1(\cdot)||p_0(\cdot)\right)$ is the error exponent rate for P_f when P_m is constrained to be no more than a certain threshold.

For binary sensor decisions with $P(U_i = 1|H_0) = \bar{\alpha}$ and $P(U_i = 1|H_1) = \bar{\beta}$, we have $P(U_i = 0|H_0) = 1 - \bar{\alpha}$ and $P(U_i = 0|H_1) = 1 - \bar{\beta}$, the KLD at the decision center for one sensor is

$$D\left(\bar{\alpha}, \bar{\beta}\right) \triangleq \bar{\alpha} \ln \frac{\bar{\alpha}}{\bar{\beta}} + (1-\bar{\alpha}) \ln \frac{1-\bar{\alpha}}{1-\bar{\beta}}, \tag{3.6}$$

where $\bar{\alpha}$ and $\bar{\beta}$ are generic notations of probability of false alarm and probability of detection, respectively, for both the FC and eavesdropper.

The KLD is always non-negative and equals 0 if and only if $\bar{\alpha} = \bar{\beta}$. Similarly, for a bounded P_m, the error exponent of P_f decays exponentialy in the number of sensors at the rate of $D\left(\bar{\beta}, \bar{\alpha}\right)$ such that $P_f \propto e^{-ND(\bar{\beta},\bar{\alpha})}$, where,

$$D\left(\bar{\beta}, \bar{\alpha}\right) \triangleq \bar{\beta} \ln \frac{\bar{\beta}}{\bar{\alpha}} + \left(1-\bar{\beta}\right) \ln \frac{1-\bar{\beta}}{1-\bar{\alpha}}. \tag{3.7}$$

For example, the KLD of each received sensor decision V_i at the FC is $D_i\left(\alpha_F, \beta_F\right)$ with α_F, β_F defined in Equation (3.3) and KLD of each received sensor decisions W_i at the eavesdropper is $D_i\left(\alpha_E, \beta_E\right)$. Owing to i.i.d. condition, $D_i\left(\alpha_F, \beta_F\right) = D\left(\alpha_F, \beta_F\right)$, $D_i\left(\alpha_E, \beta_E\right) = D\left(\alpha_E, \beta_E\right)$, and the KLD at the FC and at eavesdropper for all N are,

$$\begin{aligned} \mathbb{D}_F &= \sum_{i=1}^{N} D_i\left(\alpha_F, \beta_F\right) = ND\left(\alpha_F, \beta_F\right), \\ \mathbb{D}_E &= \sum_{i=1}^{N} D_i\left(\alpha_E, \beta_E\right) = ND\left(\alpha_E, \beta_E\right), \end{aligned} \tag{3.8}$$

respectively.

The detection performance in terms of the probability of missed-detection at the FC and at eavesdropper decays exponentially such that

$$P_{m,F} \propto e^{-\mathbb{D}_F},$$

and

$$P_{m,E} \propto e^{-\mathbb{D}_E}.$$

Therefore, to limit eavesdropper's detectability, one needs to make $\mathbb{D}_E$ as small as possible, and $\mathbb{D}_F$ as large as possible, to maximize the FC detection performance, which leads to the following secrecy constraints.

Secrecy Constraints under the Neyman-Pearson framework

$$\begin{cases} \mathbb{D}_E = ND\left(\alpha_E, \beta_E\right) \leq T_E, \\ \mathbb{D}_F = ND\left(\alpha_F, \beta_F\right) \geq T_F, \end{cases} \tag{3.9}$$

where T_E and T_F are the KLD thresholds for eavesdropper and the FC, respectively, and $\mathbb{D}_E$ and $\mathbb{D}_F$ are defined in Equation (3.8).

- Feasibility: Is it possible to design a sensor network for the targeted T_E and T_F?

- Secrecy and detection trade-off: minimize T_E under fixed T_F or maximize T_F under fixed T_E. For non-asymptotic cases, we want the detectability at eavesdropper to be as low as possible and the detection performance at the FC to be as high as possible. However, in practice, a performance trade-off between the FC and eavesdropper must be considered.

- Asymptotic perfect secrecy: $T_E \to 0$ as the number of sensors, $N \to \infty$, for example, $T_E \propto N^{-\mu}$, $0 < \mu < 1$. In this case, eavesdropper's detection capability diminishes as N increases.

- Asymptotic perfect detection: $T_F \to \infty$ as the number of sensors, $N \to \infty$.

3.2.2 Performance Metric and Secrecy Constraints under Bayesian framework

To measure the detection performance under Bayesian framework, we consider the overall probability of error, P_e,

$$P_e = \pi_0 P_f + \pi_1 P_m, \tag{3.10}$$

where π_0 and $\pi_1 = 1 - \pi_0$ are known to both the FC and an informed and greedy eavesdropper. Without loss of generality, we assume $\pi_1 \leq \frac{1}{2} \leq \pi_0$, which is known by both the FC and eavesdropper. Note, π_0 and $\pi_1 = 1 - \pi_0$ are the prior probabilities of H_0 and H_1, respectively.

Thus, for the binary hypotheses testing problem secrecy constraint, the goal is to minimize the probability of error at the FC and to increase the P_e at eavesdropper as much as possible. We formulate the optimization problem as follows:

Secrecy Constraints under Bayesian framework

$$\begin{cases} P_{e,E} \geq \Theta_E \\ P_{e,F} \leq \Theta_F, \end{cases} \tag{3.11}$$

where $P_{e,E}$ and $P_{e,F}$ are the probability of error for eavesdropper and the FC respectively, and Θ_E and Θ_F are the probability of error thresholds for eavesdropper and the FC, respectively.

- Secrecy and detection trade-off: $\Theta_E = \min(\pi_0, \pi_1) = \pi_1$ and $\min\{\Theta_F\}$. Here, we desire to constrain the detection performance at eavesdropper to a level where she can only use the prior information, and maximize the performance at the FC.

- Asymptotic perfect secrecy: $\Theta_E \to \min(\pi_0, \pi_1) = \pi_1$ and $P_{e,E} \to \min(\pi_0, \pi_1) = \pi_1$ as $N \to \infty$. In this case, observations do not provide any useful or critical information and all that eavesdropper can do is rely on the prior information and decide H_0 regardless of any W_i. Similar to the perfect secrecy constraint, as the number of sensors increases, eavesdropper receives vanishingly useful information from the observations.

- Asymptotic perfect detection: $\Theta_F \to 0$ as $N \to \infty$.

Knowing the secrecy definition and constraints, we will now solve the optimization problems in the following sections.

3.3 Performance Analysis Under Neyman-Pearson Framework

3.3.1 Maximum Achievable Performance

In order to analyze the detection performance at both the FC and eavesdropper, as well as the performance trade-offs, we derive the following approximated KLD at the receiver with BER, $\bar{\rho}$ (a generic notation of BER for both the FC and eavesdropper), when the sensors operate in the vicinity of the extreme points, i.e., if the local sensors log-likelihood ratio $\ln((p_1(x)/p_0(x))$ is unbounded, $(\alpha, \beta) \approx (0, 0)$ or $(\alpha, \beta) \approx (1, 1)$ in Table 3.1. Detailed proofs and analysis are shown in Appendix A.

Table 3.1: Approximated KLD

		$(\alpha, \beta) \approx (0, 0)$	$(\alpha, \beta) \approx (1, 1)$
$\bar{\rho} = 0$	$D(\bar{\alpha}, \bar{\beta})$	β	$(1-\alpha)\left(\ln \frac{1-\alpha}{1-\beta} - 1\right)$
	$D(\bar{\beta}, \bar{\alpha})$	$\beta\left(\ln \frac{\beta}{\alpha} - 1\right)$	$1 - \alpha$
$\bar{\rho} > 0$	$D(\bar{\alpha}, \bar{\beta})$	$\frac{1}{2}\frac{\beta^2(1-2\bar{\rho})^2}{(1-\bar{\rho})\bar{\rho}}$	$\frac{1}{2}\frac{(1-\alpha)^2(1-2\bar{\rho})^2}{(1-\bar{\rho})\bar{\rho}}$
	$D(\bar{\beta}, \bar{\alpha})$	$\frac{1}{2}\frac{\beta^2(1-2\bar{\rho})^2}{(1-\bar{\rho})\bar{\rho}}$	$\frac{1}{2}\frac{(1-\alpha)^2(1-2\bar{\rho})^2}{(1-\bar{\rho})\bar{\rho}}$

3.3.2 Noiseless Channel at the FC, where, $\rho_E > 0$ and $\rho_F = 0$

Based on the defined secrecy constraints and Table 3.1, we investigate two different scenarios for the FC; one is when the channel is perfect, the other is with an imperfect channel.

For N i.i.d sensors with total KLD at eavesdropper is constrained at T_E by,

$$\mathbb{D}_E = ND\left(\alpha_E, \beta_E\right) = T_E. \tag{3.12}$$

Since $\rho_E \neq 0$ and from Equation (A.3), we approximate the threshold at eavesdropper,

$$\frac{N}{2}\frac{\beta^2\left(1-2\rho_E\right)^2}{\left(1-\rho_E\right)\rho_E} \approx T_E.$$

Therefore, at all the sensors, the operating point should be

$$\beta \approx \sqrt{\frac{2T_E\left(1-\rho_E\right)\rho_E}{N\left(1-2\rho_E\right)^2}}, \tag{3.13}$$

which indeed goes to 0 as $T_E/N \to 0$. Because $\rho_F = 0$, from Equation (3.3) and (A.4), it can be shown that the per sensor KLD is

$$D\left(\alpha_F, \beta_F\right) \approx \beta \approx \sqrt{\frac{2T_E\left(1-\rho_E\right)\rho_E}{N\left(1-2\rho_E\right)^2}},$$

and the total KLD is

$$\begin{aligned} \mathbb{D}_F &= ND\left(\alpha_F, \beta_F\right) \approx N\beta \\ &\approx \sqrt{\frac{2NT_E\left(1-\rho_E\right)\rho_E}{\left(1-2\rho_E\right)^2}}. \end{aligned} \tag{3.14}$$

This can be utilized to design the secrecy against eavesdropper and the detection performance at the FC. For example, if we let T_E be $N^{-\mu}$, $(0 < \mu < 1)$, which results in $\beta \approx \sqrt{\frac{2N^{-\mu}(1-\rho_E)\rho_E}{N(1-2\rho_E)^2}}$, then

$$\mathbb{D}_F \approx N^{\frac{1-\mu}{2}}\sqrt{\frac{2\left(1-\rho_E\right)\rho_E}{\left(1-2\rho_E\right)^2}}.$$

Therefore, the performance and secrecy of the SN improves as the increment of the number of sensors, N, such that,

$$\begin{cases} \mathbb{D}_E \propto N^{-\mu} \\ \mathbb{D}_F \propto N^{\frac{1-\mu}{2}}, \end{cases} \tag{3.15}$$

when $\mu \approx 0$, eavesdropper's performance is constant, the performance at the FC improves at the order of $\sqrt{N}$; when $\mu \approx 1$, the FC has a guaranteed performance, eavesdropper's performance diminishes at the order of $1/N$. when $N \to \infty$, $\mathbb{D}_F \to \infty$, which results in asymptotic perfect detection [18] at the FC and $\mathbb{D}_E \to 0$, asymptotic perfect secrecy. We summarize the findings in the following theorem.

Theorem 1. ***Asymptotic Perfect Secrecy and Asymptotic Perfect Detection under Neyman-Pearson Framework:*** *When eavesdropper has a noisy channel, $\rho_E > 0$, and the FC has a noiseless channel, $\rho_F = 0$, the secrecy constraints ($\mathbb{D}_E \leq T_E$; $\mathbb{D}_F \geq T_F$) can be satisfied for any arbitrary constants T_E and T_F, given a sufficiently large number of sensors, N.*

3.3.3 Noisy Channel, where, $\rho_F > 0$ and $\rho_E > 0$

Rarely does a perfect communication channel exist in practice, so we investigate the case where the FC has a noisy channel. Since $\rho_F \neq 0$, from Table 3.1, we know that

$$D\left(\alpha_F, \beta_F\right) \approx \frac{1}{2}\frac{\beta^2\left(1-2\rho_E\right)^2}{\left(1-\rho_E\right)\rho_E}.$$

Under the secrecy constraint in Equation (3.9), after applying β from Equation (3.13), we obtain

$$\mathbb{D}_F \approx \frac{T_E\rho_E(1-\rho_E)(1-2\rho_F)^2}{\rho_F(1-\rho_F)(1-2\rho_E)^2}. \tag{3.16}$$

To measure the performance trade-off, we define the KLD ratio between the FC and eavesdropper as,

$$R = \frac{\mathbb{D}_F}{\mathbb{D}_E} = \frac{D\left(\alpha_F, \beta_F\right)}{D\left(\alpha_E, \beta_E\right)} = \frac{D\left(\beta_F, \alpha_F\right)}{D\left(\beta_E, \alpha_E\right)} \tag{3.17}$$

Therefore, if we plug Equation (3.16) into Equation (3.17), we have the following result for the performance trade-off between the FC and eavesdropper.

Theorem 2. ***Maximum Achievable Performance Trade-off under Neyman-Pearson Framework*** *When both eavesdropper and the FC have noisy channels,*

$\rho_F > 0$ and $\rho_E > 0$, the secrecy constraints can only be achieved for certain T_E and T_F such that T_F/T_E is no more than the ratio, $R = \frac{(1-\rho_E)\rho_E}{(1-\rho_F)\rho_F}\left(\frac{1-2\rho_F}{1-2\rho_E}\right)^2$.

(See Appendix B for the detailed proof).

For example, if $\rho_F = 0.1$, $\rho_E = 0.3$, and the required $\mathbb{D}_F > 10$, the resulting information leakage is $\mathbb{D}_E > 1.07$. In other words, the information leakage is inevitable, no matter how one increases the number of sensors in the network. On the other hand, when ρ_F is much smaller than ρ_E by using the techniques mentioned in [32–34, 99], then the performance ratio can still be large enough to maintain high detectability at the FC and poor performance at eavesdropper. This point is expanded upon in Section 3.5. The ratio can also serve as a performance design protocol for the SN, for instance, the desired performance at the FC and at eavesdropper are T_F and T_E, we can compute the corresponding ρ_F when ρ_E is fixed or the other way round.

3.4 Performance Analysis Under Bayesian Framework

Recall that the goal under Bayesian framework is to minimize the probability of error in Equation (3.10) at the FC and constrain that at eavesdropper at a certain level.

3.4.1 Detection Performance Trade-off under Perfect Secrecy Constraint

Since both the FC and eavesdropper know the exact prior probabilities and $\pi_0 \geq \pi_1$, the detection eavesdroppers probability of error bound is π_1 achieved by accepting H_0 regardless of the observations. The constraints on eavesdropper are that observations should not be of any help in her decision making ability and the $P_{e,E}$ at the eavesdropper should remain at π_1. When this is true, eavesdropper still makes the decision H_0 regardless of the observations W_i, $i = 1, 2, \ldots, N$. That means,

$$\begin{aligned} & P(H_1|W) \leq P(H_0|W), \;\; \forall W_i \\ \Longrightarrow\ & P(W|H_1)\pi_1 \leq P(W|H_0)\pi_0, \;\; \forall W_i \\ \Longrightarrow\ & \frac{p(W|H_1)}{p(W|H_0)} \leq \frac{\pi_0}{\pi_1} \;\; \forall W_i \\ \Longrightarrow\ & \operatorname*{argmax}_{W} \left(\frac{p(W|H_1)}{p(W|H_0)} \right) \leq \frac{\pi_0}{\pi_1} \;\; \forall W_i. \end{aligned}$$

The maximum of the LRT is achieved when $W_1 = W_2 = \cdots = W_N = 1$ such that $\max\left(\frac{p(W|H_1)}{p(W|H_0)}\right) = (\beta_E/\alpha_E)^N$ with α_E and β_E defined in Equation (3.4). That is, in order to limit eavesdropper's detectability to the prior information,

$$\left(\frac{\beta_E}{\alpha_E}\right)^N \leq \frac{\pi_0}{\pi_1}. \tag{3.18}$$

In this case, the wirelessly tapped sensors observations can provide some information, but not enough to overcome the prior information to make any difference in the final decision making.

Meanwhile, for the performance at the FC, we derive the following theorem,

Theorem 3. ***Maximum Achievable Performance Trade-off under Bayesian Framework*** *When the FC has a noiseless channel and eavesdropper has a noisy channel,* $0 < \rho_E < 0.5$*, the minimum achievable* $P_{e,F}$ *at the FC is given by ,* $\lim_{N\to\infty} P_{e,F} = P_f\pi_0 = \pi_1 \left(\frac{\pi_0}{\pi_1}\right)^{-\frac{\rho_E}{1-2\rho_E}}$.

$P_{e,F}$ is a function of prior probabilities and the eavesdropper's channel qualities ρ_E, and it is also strictly greater than 0 for any $\rho_E < 0.5$. The details of proofs are shown in Appendix C.

Remark 1. When $\pi_0 = \pi_1 = 0.5$, then $P_{e,F} = 0.5$, that means, it is impossible to achieve perfect secrecy, while providing the FC with any useful information.

For the case that the FC does not have a perfect channel, where $\rho_F > 0$, the detection performance becomes worse, and the corresponding probability of error at the FC increases as well.

3.4.2 Asymptotic Perfect Secrecy and Asymptotic Perfect Detection

We know that asymptotic perfect secrecy and asymptotic perfect detection can be achieved under N-P framework from Theorem 1, here we investigate the same problem under the Bayesian framework, requiring

$$P_{e,E} \to \min\left(\pi_0, \pi_1\right), \;\; N \to \infty.$$

To evaluate the asymptotic error rate, we need to establish the error decay rate bound for the FC and eavesdropper respectively. First, from large deviation theory, for any decision center with conditionally i.i.d., Bernoulli observations Y_i with $P(Y_i = 1|H_0) = \bar{\alpha}$ and $P(Y_i = 1|H_1) = \bar{\beta}$, $i = 1, 2, \ldots, N$, the decision rule is

$$\frac{\sum_{i=1}^{N} Y_i}{N} \underset{H_0}{\overset{H_1}{\gtrless}} T.$$

Based on the work in [16] and the Chernoff inequality

$$P_f \approx e^{-ND(T,\bar{\alpha})},$$
$$P_m \approx e^{-ND(T,\bar{\beta})},$$

where T is the decision rule threshold. Notice that

$$\begin{aligned}\max\left(\ln P_f + \ln \pi_0,\ \ln P_m + \ln \pi_1\right) &\le \ln P_e, \\ \ln P_e &\le \max\left(\ln P_f + \ln \pi_0,\ \ln P_m + \ln \pi_1\right) + \ln 2.\end{aligned} \tag{3.19}$$

Hence, for a sufficiently large N,

$$\begin{aligned}-\frac{\ln P_e}{N} &\approx \min\left(-\frac{\ln P_f}{N}, -\frac{\ln P_m}{N}\right), \\ &\approx \min\left(D(T,\bar{\alpha}), D(T,\bar{\beta})\right).\end{aligned} \tag{3.20}$$

Therefore, the optimal T for large N is chosen such that $D(T,\bar{\alpha}) = D(T,\bar{\beta})$. In Appendix D, we show that

$$T = \frac{D(\bar{\alpha},\bar{\beta})\bar{\beta} + D(\bar{\beta},\bar{\alpha})\bar{\alpha}}{D(\bar{\beta},\bar{\alpha}) + D(\bar{\alpha},\bar{\beta})}, \tag{3.21}$$

which reveals the relationship between T and the KLD distances $D(\bar{\alpha},\bar{\beta})$ and $D(\bar{\beta},\bar{\alpha})$.

Table 3.2: Decision Rule Threshold *T*

	$(\alpha,\beta) \approx (0,0)$	$(\alpha,\beta) \approx (1,1)$
$\bar{\rho} = 0$	$\frac{\beta-\alpha}{\ln\frac{\beta}{\alpha}} + \alpha$	$\beta + \frac{\alpha-\beta}{\ln\frac{1-\alpha}{1-\beta}}$
$\bar{\rho} > 0$	$\frac{\bar{\alpha}+\bar{\beta}}{2} = \bar{\rho} + \frac{(1-\bar{\rho})(\alpha+\beta)}{2}$	$\frac{\bar{\alpha}+\bar{\beta}}{2} = \bar{\rho} + \frac{(1-\bar{\rho})(\alpha+\beta)}{2}$

By plugging the decision rule threshold into the approximated KLD in Table 3.1, we summarize the approximated asymptotic error exponent in Table 3.3.

Table 3.3: Approximated Asymptotic Error Exponent

	$(\alpha, \beta) \approx (0,0)$	$(\alpha, \beta) \approx (1,1)$
$\bar{\rho} = 0$	$\left(\frac{\beta-\alpha}{\ln\frac{\beta}{\alpha}} + \alpha\right) \ln\left(\frac{\beta-\alpha}{\alpha \ln\frac{\beta}{\alpha}}\right)$	$\left(1 - \beta - \frac{\alpha-\beta}{\ln\frac{1-\alpha}{1-\beta}}\right) \ln\left(\frac{\alpha-\beta}{(1-\beta)\ln\frac{1-\alpha}{1-\beta}}\right)$
$\bar{\rho} > 0$	$\frac{1}{8}\frac{\beta^2(1-2\bar{\rho})^2}{(1-\bar{\rho})\bar{\rho}}$	$\frac{1}{8}\frac{(1-\alpha)^2(1-2\bar{\rho})^2}{(1-\bar{\rho})\bar{\rho}}$

According to Table 3.3, when $(\alpha, \beta) \to (0,0)$ and $\rho_F > 0$, the asymptotic error rate at FC can be approximated as $\frac{\beta^2(1-2\rho_F)^2}{8\rho_F(1-\rho_F)}$. Similarly, for eavesdropper with $\rho_E > 0$, the error exponent $D\left(T_E, \alpha_E\right) = D\left(T_E, \beta_E\right) \approx \frac{\beta^2(1-2\rho_E)^2}{8\rho_E(1-\rho_E)}$. Since $P_{e,E} \propto e^{-ND(T_E,\beta_E)}$, in order to achieve the asymptotic perfect secrecy under the Bayesian settings, it is required that

$$N\beta^2(1-2\rho_E)^2 \to 0, N \to \infty \implies \beta = o\left(\sqrt{N^{-1}}\right), \tag{3.22}$$

where $f = o(g)$ denotes that function f grows strictly slower than function g, whereas $f = O(g)$ means f grows slower than or equal to g.

Meanwhile, when the FC channel is noise free and $(\alpha, \beta) \to (0,0)$, we have $T_F = \frac{\beta-\alpha}{\ln\frac{\beta}{\alpha}} + \alpha$ and the resulting asymptotic error exponent $D(T_F, \alpha) = D(T_F, \beta) \approx \left(\frac{\beta-\alpha}{\ln\frac{\beta}{\alpha}} + \alpha\right) \ln\left(\frac{\beta-\alpha}{\alpha\ln\frac{\beta}{\alpha}}\right) = o(\beta)$. Hence, $P_{e,F} \propto e^{-N\beta}$. With the constraint of asymptotic perfect secrecy (3.22), we know that $N\beta = O\left(\sqrt{N}\right)$, i.e., the probability of error at the FC decays no faster than $e^{-\sqrt{N}}$. In practical applications, β can be chosen as $c(1/(\sqrt{N}\ln N))$, where c is a constant. This result leads to Theorem 4.

Theorem 4. ***Asymptotic Perfect Secrecy and Asymptotic Perfect Detection under Bayesian Framework*** *When eavesdropper has a noisy channel, $\rho_E > 0$ and the FC has a noiseless channel, $\rho_F = 0$, the secrecy constraint ($P_{e,E} \geq \Theta_E$, $P_{e,F} \leq \Theta_F$,) can be satisfied for any arbitrary constants $\Theta_E < \min\left(P(H_0), P(H_1)\right)$ and $\Theta_F > 0$ given a sufficiently large number of sensors, N.*

In summary, we showed that under Bayesian framework, eavesdropper's detectability can be limited to the level where she can only rely on the prior information, but this induces a performance cost at the FC. Additionally, it was shown that both asymptotic perfect secrecy and asymptotic perfect detection are possible.

3.5 Experimental Results

In this section, we compare the detection performance at eavesdropper and the FC via the canonical distributed detection problem of a constant signal with zero mean additive white Gaussian noise. Specifically, the sensor observations are given by

$$\begin{cases} H_1: & X_i = A + Z_i \\ H_0: & X_i = Z_i, \end{cases}$$

where $Z_i \sim \mathcal{N}(0,1)$ is the normalized observation noise following a standard Gaussian distribution, $A > 0$ is a fixed constant signal to be detected with signal-to-noise ratio, SNR $= 20\log_{10} A$ dB. In this setting, the sensor log-likelihood ratio $\ln\left(p_1(x_i)/p_0(x_i)\right) = Ax - \frac{A^2}{2}$, is unbounded from above and below, and the detection probability is given by $\beta(\alpha) = Q(Q^{-1}(\alpha) - A)$, where $Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty \exp(-\frac{u^2}{2})du$ is the tail probability of a standard Gaussian distribution.

3.5.1 Simulations under Neyman-Pearson Framework

We first examine the system secrecy when the FC has a non-perfect channel, $\rho_F > 0$ and $\rho_E > 0$. The upper figures of Figure 3.2 and Figure 3.3, show the performance comparison between the FC and eavesdropper in terms of KLD for both cases, $D(\bar{\alpha}, \bar{\beta})$ and $D(\bar{\beta}, \bar{\alpha})$ when $N = 1$, SNR $= 0$ dB ($A = 1$), $\rho_F = 0.01$, and $\rho_E = 0.35$. The bottom figures show the ratio between KLD at the FC and KLD at eavesdropper for one sensor. In this case, the maximum achievable KLD ratio defined in Equation (3.17), $R = 250$ and the marker star in the bottom figure indicates the actual maximum ratio. We can see that when probability of false alarm, α, is close to 0 or 1, the ratio $D(\alpha_F, \beta_F)/D(\alpha_E, \beta_E) = D(\beta_F, \alpha_F)/D(\beta_E, \alpha_E)$ and they are close to the theoretical value in Theorem 2, which is represented by the horizontal line in each figure. Given the secrecy constraint such that $\mathbb{D}_E \leq T_E$ is bounded, then the maximum achievable $\mathbb{D}_F \leq RT_E$ is also bounded, however, since ρ_F is small here, the ratio R is still a large number, which reflects the detection performance gap between eavesdropper and the FC. In other words, this ratio can be utilized in sensor network design to improve secrecy at the physical-layer.

In order to show the achievable asymptotic perfect detection at the FC and asymptotic zero detection at eavesdropper, we compare the detection performance at eavesdropper and the FC in terms of their KLDs for all N sensors, $\mathbb{D}_E$ and $\mathbb{D}_F$, which are under the conditions that SNR $= 0$ dB, $\rho_F = 0$, $\rho_E = 0.35$ and we set $\mathbb{D}_E \to \frac{0.1}{\sqrt{N}}$, which is monotone decreasing in N. From Figure 3.4 and Figure 3.5, the trends of KLDs at the FC and KLD at eavesdropper show that both asymptotic

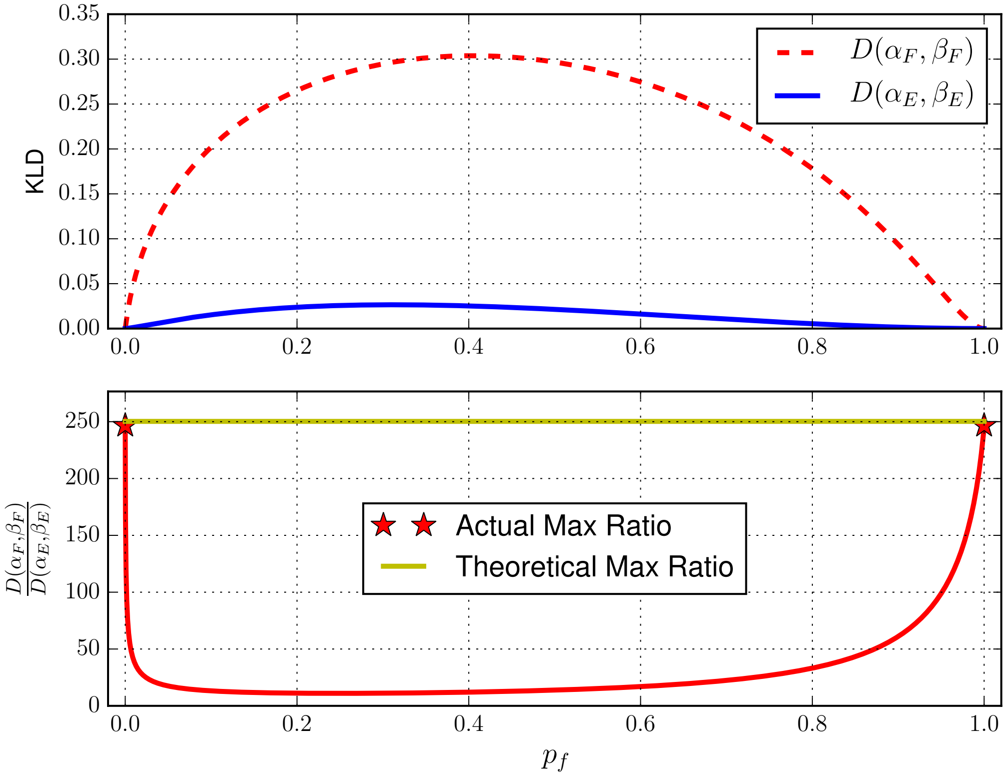

Figure 3.2: The maximum achievable detection performance trade-off under Neyman-Pearson framework using KLD for one sensor.

perfect detection and asymptotic perfect secrecy are possible by increasing the number of sensors. In the figure, $\widetilde{\mathbb{D}_F}$ and $\widetilde{\mathbb{D}_E}$ denote the approximated KLDs at the FC and at eavesdropper, respectively, where $\widetilde{\mathbb{D}_F}$ is computed using Equation (3.14). We can see that the approximated KLDs approach to the actual KLDs for both cases when $(\alpha, \beta) \to (0, 0)$ and $(\alpha, \beta) \to (1, 1)$.

ROC curves shown in Figure 3.6, are obtained under the same settings in Figure 3.4 with 10 and 50 sensors. The corresponding eavesdropper curves are approaching a diagonal line which implies no detectability. When the number of sensors is 50, the corresponding detection performance of the FC is almost perfect. The ROC curves

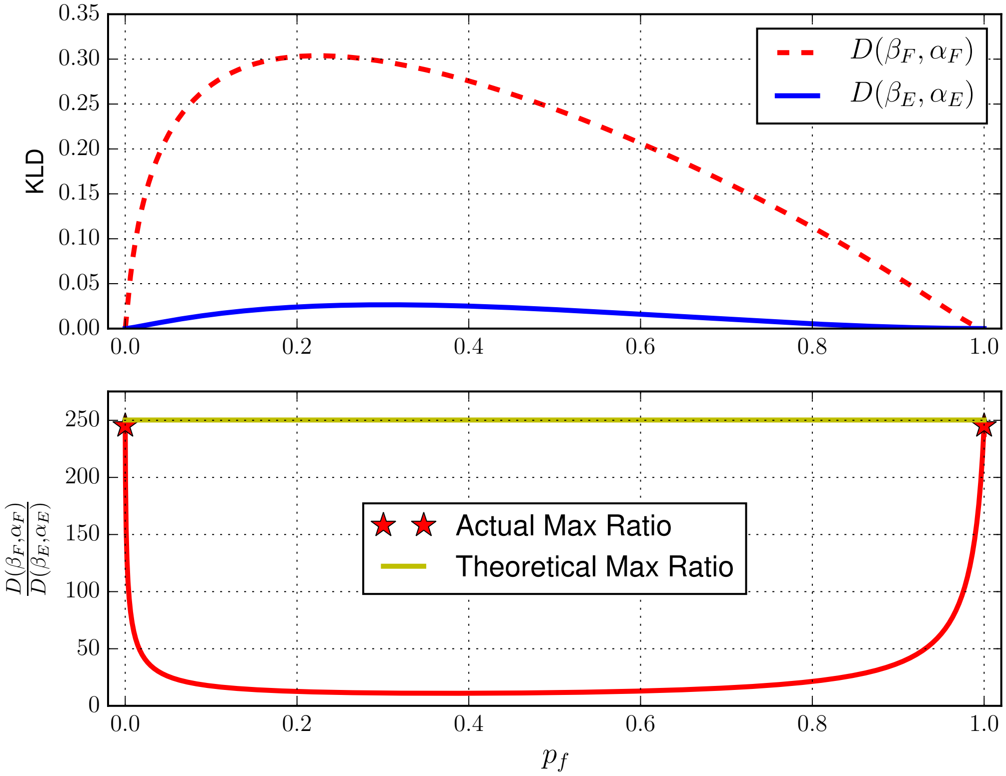

Figure 3.3: The maximum achievable detection performance trade-off under Neyman-Pearson framework using KLD for one sensor.

show again that asymptotic perfect secrecy can be achieved by increasing the number of sensors and adjusting sensor optimality points accordingly.

3.5.2 Simulations under Bayesian Framework

Under Bayesian framework, we illustrate the secrecy and detection performance trade-off with relative to the total number of sensors, N, in Figure 3.7. We set π_0 and π_1 as 0.7 and 0.3 respectively, SNR $=$ 3 dB, and $\rho_E = 0.3$, $\rho_F = 0$. From the figure, we can see that the simulated data approaches the FC theoretical probability of error value of 0.159, computed using Theorem 3. Meanwhile, the probability of

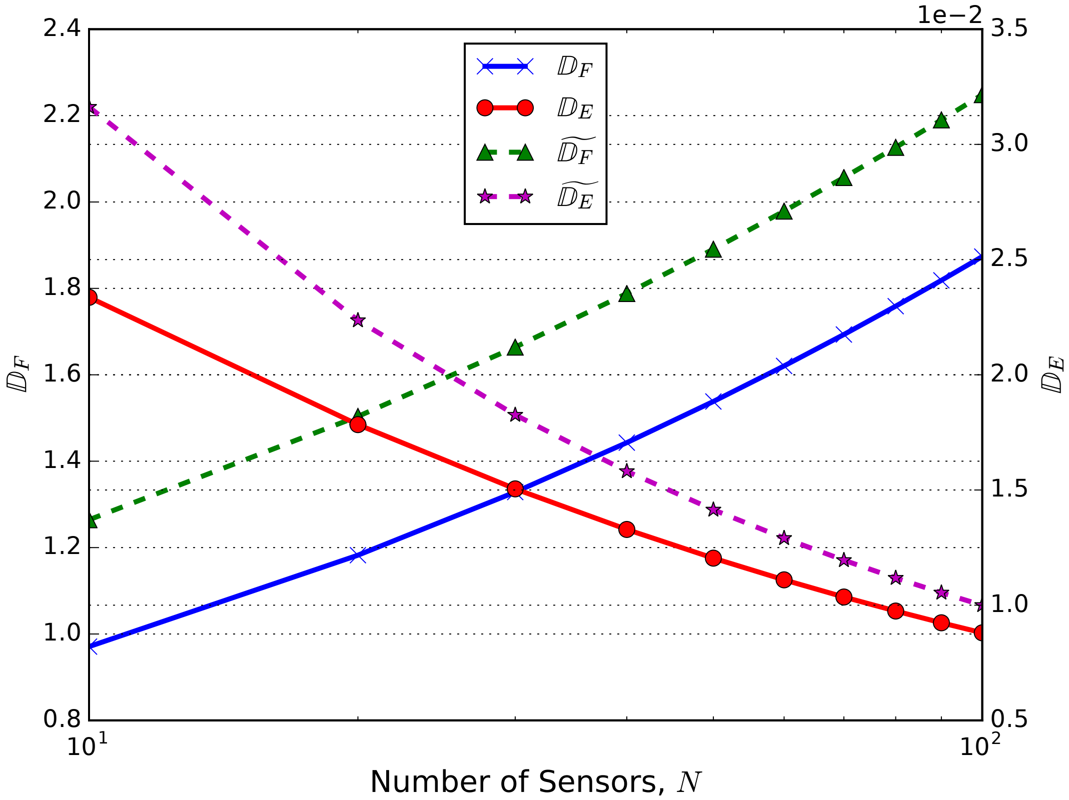

Figure 3.4: When $(\alpha, \beta) \approx (\mathbf{0}, \mathbf{0})$, the asymptotic secrecy and detection performance using approximated and actual KLD $(D(\bar{\alpha}, \bar{\beta}))$ at the FC and at eavesdropper for N sensors. $\mathbb{D}_F$ and $\mathbb{D}_E$ denote the actual KLDs at the FC and at eavesdropper, respectively. $\widetilde{\mathbb{D}_F}$ and $\widetilde{\mathbb{D}_E}$ denote the approximated KLDs at the FC and at eavesdropper, respectively.

error at eavesdropper remains fixed at $\pi_1 = 0.3$. In other words, the detectability of eavesdropper is constrained at her prior information and observations do not improve her decision-making. Meanwhile, the detection performance at the FC does not exceed the bound derived in Theorem 3.

For asymptotic performance analysis, we first plot the asymptotic error exponent for the FC with noiseless channels in Figure 3.8, the SNR and prior probabilities are the same with the ones in Figure 3.7. Meanwhile, the sensor's probability of

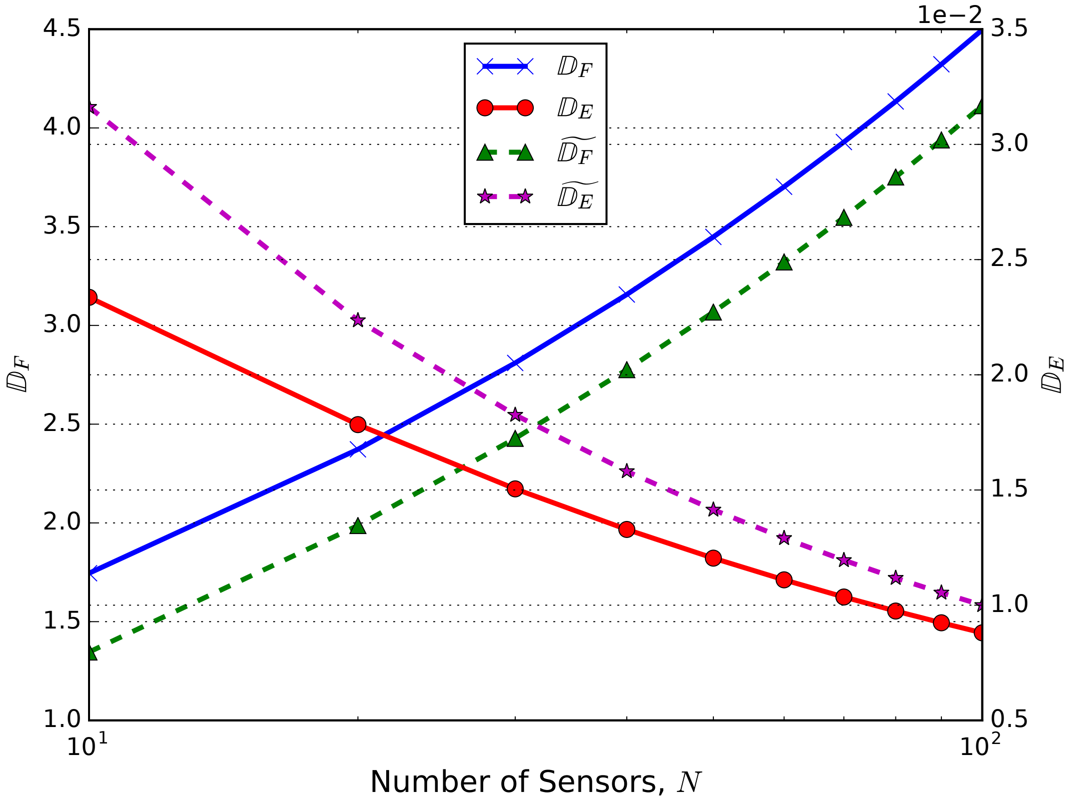

Figure 3.5: When $(\alpha, \beta) \approx (\mathbf{1}, \mathbf{1})$, the asymptotic secrecy and detection performance using approximated and actual KLD $(D(\bar{\alpha}, \bar{\beta}))$ at the FC and at eavesdropper for N sensors. $\mathbb{D}_F$ and $\mathbb{D}_E$ denote the actual KLDs at the FC and at eavesdropper, respectively. $\widetilde{\mathbb{D}_F}$ and $\widetilde{\mathbb{D}_E}$ denote the approximated KLDs at the FC and at eavesdropper, respectively.

detection is selected as $\beta = \frac{1.5}{\sqrt{N \ln N}}$, and the corresponding probability of false alarm is $\alpha = Q\left(Q^{-1}(\beta) + A\right)$. We can see the estimated error exponent of the FC in Figure 3.8, approaches the actual error exponent when the number of sensors is 100.

We then plot the probability of error for the FC and eavesdropper in Figure 3.9 under the above conditions and $\rho_E = 0.35$. As the number of sensors increase, probability of error for eavesdropper stays at π_1, which means the reported sensor observations do not improve eavesdroppers detection ability. As for the FC, the

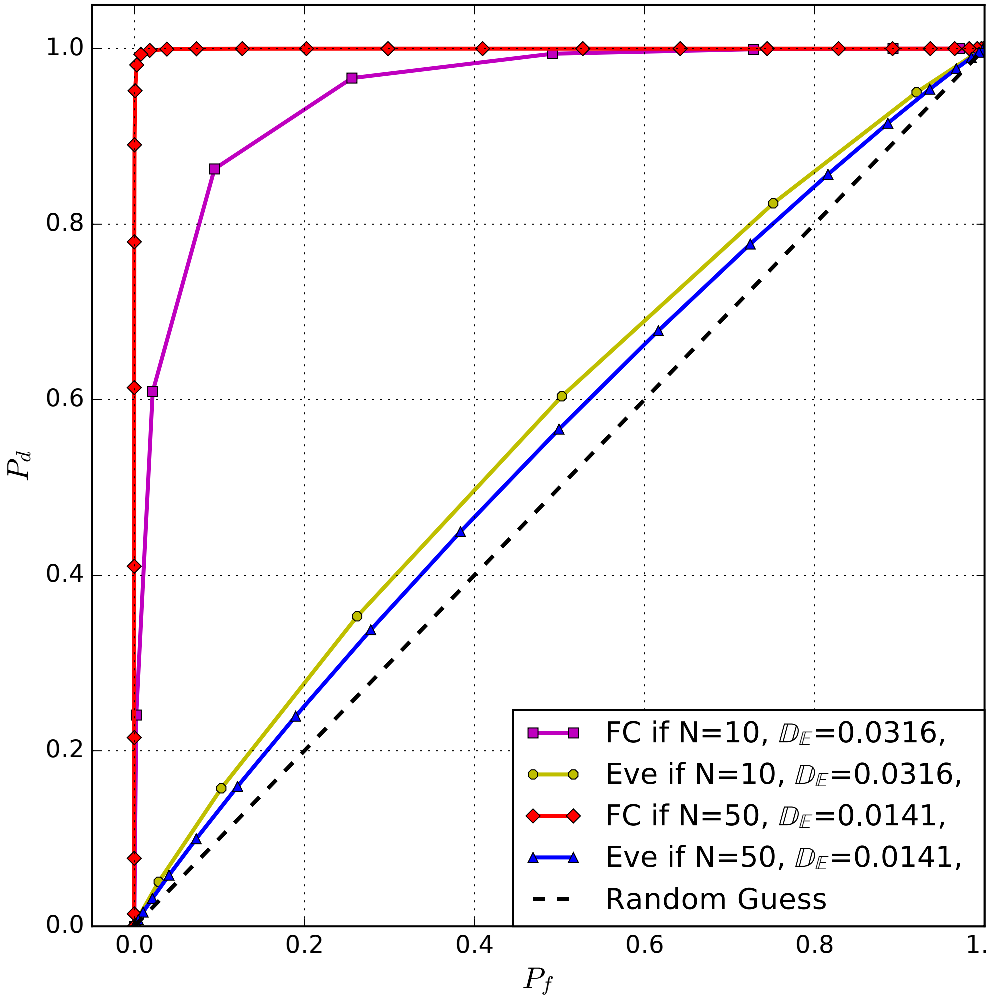

Figure 3.6: ROC curves for the FC and eavesdropper with different number of sensors in N-P. The random guess line implies zero detectability. eavesdropper's detectability is getting closer and closer to the diagonal line.

probability of error quickly diminishes to zero with a few hundred sensors. In other words, asymptotic perfect secrecy is possible under Bayesian framework.

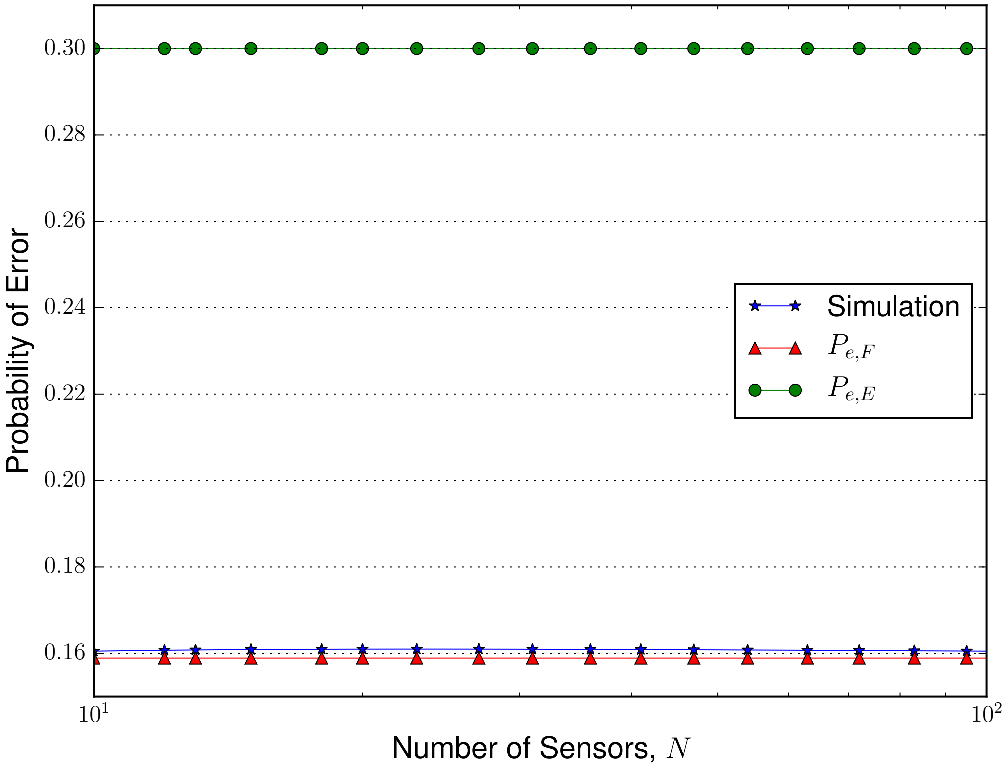

Figure 3.7: The maximum achievable detection performance trade-off under Bayesian framework, where eavesdropper's detection performance is constrained to the prior information.

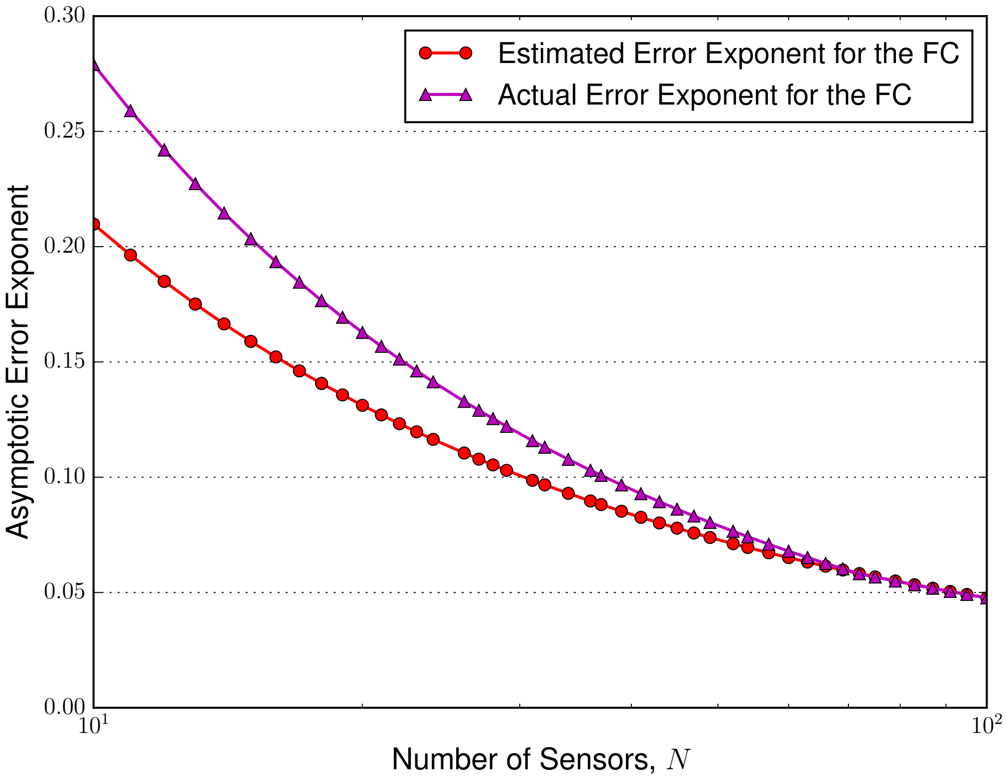

Figure 3.8: The asymptotic error exponent for the FC with $\rho_F = 0$ when $(\alpha, \beta) \to (0, 0)$. As the number of sensor increases, the estimated error exponent approaches to the actual error exponent.

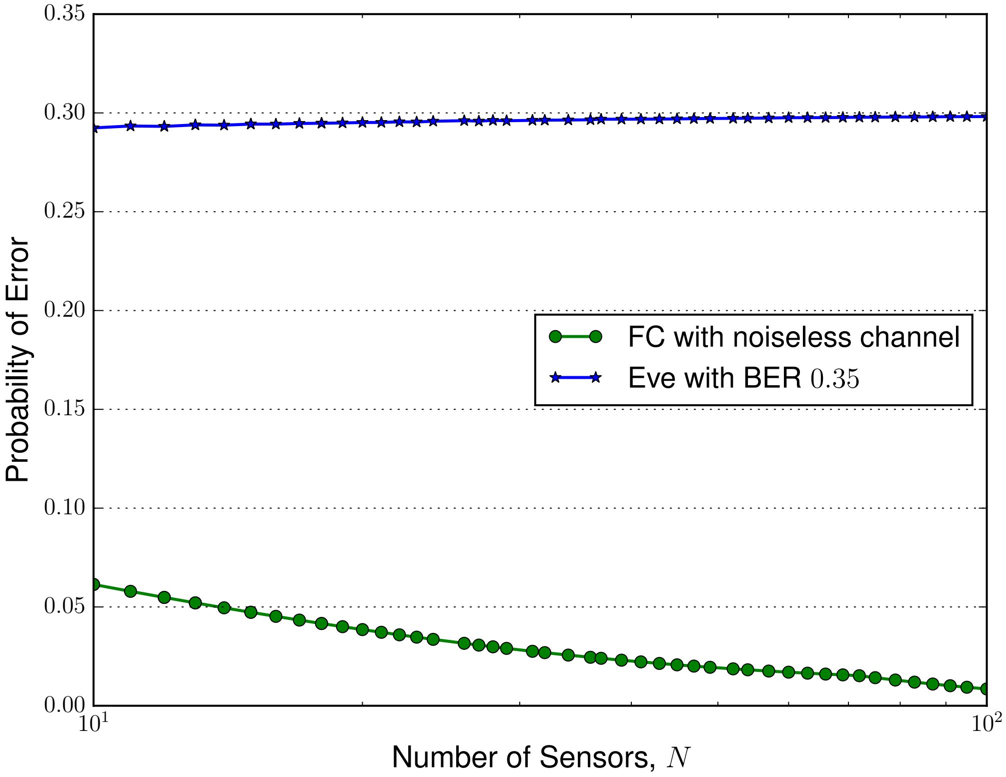

Figure 3.9: The asymptotic detection performance under Bayesian framework where, $\pi_1 = 0.3$ and $\beta = \frac{1.5}{\sqrt{N \ln N}}$. $P_{e,F}$ decreases as the number sensors increases, while $P_{e,E}$ at eavesdropper is constrained at π_1.

CHAPTER 4

SECRECY CONSTRAINED DISTRIBUTED ESTIMATION IN WSNS

This chapter is organized as follows. In Section 4.1, we introduce the estimation system model with secrecy constraints and in 4.2, we introduce estimation performance metric, FI and explain how to achieve asymptotic perfect secrecy. We continue analyzing the case where the parameter is fixed but unknown and the observation noise follows Gaussian distribution in Section 4.3. For Bayesian cases where the parameter is a random variable, we show that the secrecy constraints for the FC and eavesdropper can be satisfied as well in Section 4.4. Simulation results are provided to further support our proofs in Section 4.5.

4.1 Distributed Estimation Model

The parallel WSN model with a global and greedy eavesdropper who has access to all sensors outputs is shown in Figure 4.1, where sensors observe parameter θ, quantize their observations, then send them to the FC across channels. In many applications, the sensors are deployed to monitor the environment. In such scenarios, the sensor observations can often be assumed to be conditionally independent and identically distributed (i.i.d) given the underlying parameter θ. Under this assumption, the sensor observations $\mathbf{X} = [X_1, X_2, \ldots, X_N]$ can be written as follows,

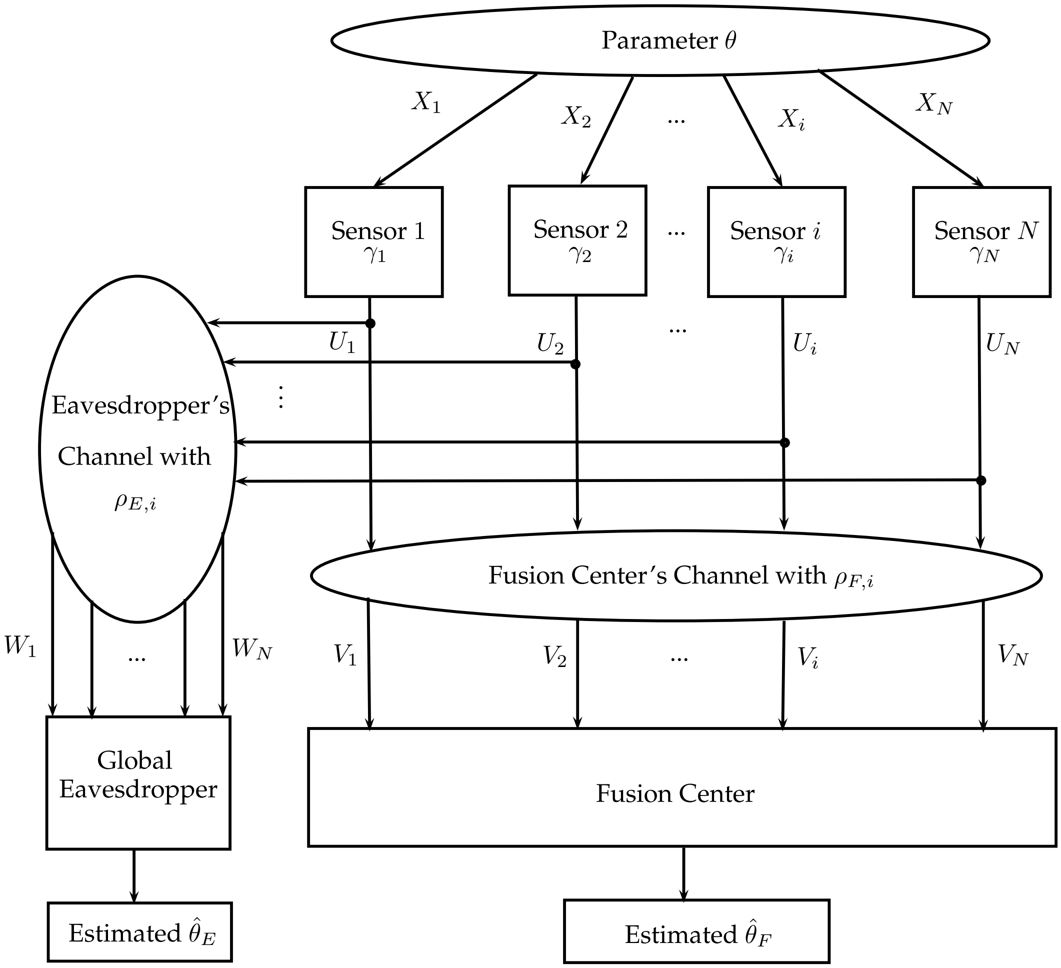

Figure 4.1: Parallel sensor network model under eavesdropper attack, who eavesdrops on the output of sensor i, transmitted wirelessly via a binary symmetric channel with bit error rate $\rho_{E,i}$. The FC receives sensor i data through another binary symmetric channel with bit error rate $\rho_{F,i} < \rho_{E,i}$.

$$f(\mathbf{X}|\theta) = \prod_{i=1}^{N} f(X_i|\theta),$$

where $f(\mathbf{X}|\theta)$ and $f(X_i|\theta)$ are known probability density functions (pdfs) and X_i is the observation of sensor i.

We consider the estimation problem, where the ith sensor observation X_i is,

$$X_i = \theta + Z_i, \;\; i = 1, 2, \ldots, N, \tag{4.1}$$

where Z_i is an additive i.i.d zero mean observation noise with pdf $f(\cdot)$. Due to the bandwidth constraint between local sensors and the FC, we assume the X_i are quantized to a single bit of compressed data, U_i, via the quantization rule

$$U_i = \begin{cases} 1, & X_i > \eta_i \\ & \qquad\qquad \forall i, \\ 0, & X_i \leq \eta_i \end{cases} \tag{4.2}$$

where the threshold, η_i, is fixed and known to both the FC and eavesdropper. To reduce the system complexity and improve system robustness, we assume that the sensors employ identical quantization rules such that $\eta_1 = \eta_2 = \cdots = \eta_N = \eta$. Because the sensors observations are conditionally i.i.d., we have

$$\begin{aligned} &\Pr(U_i = 1|\theta) = \beta = \Pr(\theta + Z_i > \eta) = Q(\eta - \theta), \\ &\Pr(U_i = 0|\theta) = 1 - \beta = 1 - \Pr(U_i = 1|\theta), \end{aligned}$$

where $Q(t) = \int_t^\infty f_Z(x)dx$ is the complementary distribution function of Z.

The communication channels between sensors and the receivers are assumed to be BSCs. Sensor i sends decision U_i to the FC over a BSC with BER $\rho_{F,i} < \frac{1}{2}$, with the received decision V_i. All of the sensors outputs are eavesdropped by eavesdropper via a set of parallel wiretapping channels. eavesdropper receives W_i, from sensor i as an output of a separate BSC channel with BER $\rho_{E,i} < \frac{1}{2}$. We assume that eavesdropper's channel is noisier than the FC's such that $\rho_{E,i} > \rho_{F,i}$ [32, 99]. Assuming that the sensors are within similar distances to eavesdropper and the FC, then the channels can be assumed to be independent and identical, i.e., $\rho_F = \rho_{F,1} = \cdots = \rho_{F,N}$ and $\rho_E = \rho_{E,1} = \cdots = \rho_{E,N}$.

As a result, the observations at the FC and eavesdropper possess the following quality,

$$\begin{aligned}\Pr(V_i = 1|\theta) &= (1 - 2\rho_F)\Pr(U_i = 1|\theta) + \rho_F,\\ \Pr(W_i = 1|\theta) &= (1 - 2\rho_E)\Pr(U_i = 1|\theta) + \rho_E.\end{aligned}$$

For the purposes of this chapter we analyze identical channels, although non-identical channels can be treated in a similar fashion.

4.2 Estimation Performance and Asymptotic Perfect Secrecy under Classical Setting

We now evaluate the estimation performance at the FC and eavesdropper under a classical setting, using the widely employed Mean Squared Error (MSE) metric under some wild conditions [15, 39]. The Cramér-Rao inequality given observations $\mathbf{V} = [V_1, \ldots, V_N]^T$ and known quantization rules [84, 87], establishes a MSE lower bound for any unbiased estimator of $\hat{\theta}_F$, ϵ_F such that ,

$$\epsilon_F \triangleq E\left(\hat{\theta}_F - \theta\right)^2 \geq \text{CRLB}(\mathbf{V};\theta) = \frac{1}{\mathbf{I}(\mathbf{V};\theta)}, \tag{4.3}$$

where CRLB is Cramér-Rao lower bound [39] and $\mathbf{I}(\mathbf{V};\theta)$ is the FI, given by,

$$\begin{aligned}\mathbf{I}(\mathbf{V};\theta) &\triangleq E_{\mathbf{V}}\left[\left(\frac{\partial \log p(\mathbf{V};\theta)}{\partial\theta}\right)^2\right]\\ &\overset{(a)}{=} \sum_{i=1}^{N} E_{V_i}\left[\left(\frac{\partial \log p(V_i;\theta)}{\partial\theta}\right)^2\right]\\ &\overset{(b)}{=} NI(\eta,\theta,\rho_F),\end{aligned}$$

where $p(\mathbf{V};\theta)$ is the pdf of parameter θ given $\mathbf{V}$ [39]. Note, (a) and (b) follow from the sensors observations conditionally i.i.d property, the identical channels assumption, and

$$I(\eta,\theta,\rho) = \frac{f^2(\eta-\theta)(1-2\rho)^2}{(\rho+(1-2\rho)Q(\eta-\theta))(1-\rho-(1-2\rho)Q(\eta-\theta))} \tag{4.4}$$

is the per sensor FI when the sensor observation is received over a BSC with BER ρ.

Similarly, at eavesdropper with $\mathbf{W} = [W_1,\ldots,W_N]^T$, the MSE lower bound, ϵ_E, for any unbiased estimator $\hat{\theta}_E$ is

$$\begin{aligned}\epsilon_E \triangleq E\left(\hat{\theta}_E-\theta\right)^2 &\geq \text{CRLB}(\mathbf{W};\theta)\\ = \frac{1}{\mathbf{I}(\mathbf{W};\theta)} &= \frac{1}{NI(\eta,\theta,\rho_E)}.\end{aligned}$$

4.2.1 Fisher Information Ratio

Based on the CRLB, the secrecy design problems can be framed as maximizing the FI at the FC while minimizing the FI at eavesdropper. Therefore, we introduce the FI ratio R as an intermediate step to achieve these secrecy requirements, with a higher R indicating a better secrecy. The FI ratio is defined as follows,

$$
\begin{aligned}
R(\eta,\theta) &\triangleq \frac{\mathbf{I}(\eta,\theta,\rho_F)}{\mathbf{I}(\eta,\theta,\rho_E)} \\
&= \frac{(1-2\rho_F)^2(\rho_E+(1-2\rho_E)Q(\eta-\theta))}{(1-2\rho_E)^2(\rho_F+(1-2\rho_F)Q(\eta-\theta))} \\
&\times \frac{(1-\rho_E-(1-2\rho_E)Q(\eta-\theta))}{(1-\rho_F-(1-2\rho_F)Q(\eta-\theta))} \\
&= \frac{\left(\frac{\rho_E}{1-2\rho_E}+Q(\eta-\theta)\right)\left(\frac{1-\rho_E}{1-2\rho_E}-Q(\eta-\theta)\right)}{\left(\frac{\rho_F}{1-2\rho_F}+Q(\eta-\theta)\right)\left(\frac{1-\rho_F}{1-2\rho_F}-Q(\eta-\theta)\right)} \\
&= \frac{-Q^2(\eta-\theta)+Q(\eta-\theta)+\frac{\rho_E(1-\rho_E)}{(1-2\rho_E)^2}}{-Q^2(\eta-\theta)+Q(\eta-\theta)+\frac{\rho_F(1-\rho_F)}{(1-2\rho_F)^2}} \\
&= 1+\frac{\frac{\rho_E(1-\rho_E)}{(1-2\rho_E)^2}-\frac{\rho_F(1-\rho_F)}{(1-2\rho_F)^2}}{-Q^2(\eta-\theta)+Q(\eta-\theta)+\frac{\rho_F(1-\rho_F)}{(1-2\rho_F)^2}} \\
&= 1+\frac{\frac{\rho_E(1-\rho_E)}{(1-2\rho_E)^2}-\frac{\rho_F(1-\rho_F)}{(1-2\rho_F)^2}}{-\left(Q(\eta-\theta)-\frac{1}{2}\right)^2+\frac{1}{4}+\frac{\rho_F(1-\rho_F)}{(1-2\rho_F)^2}}.
\end{aligned}
\tag{4.5}
$$

Notice that the function $\frac{\rho(1-\rho)}{(1-2\rho)^2}$ is a monotone increasing function for $\rho < 0.5$, and since $\rho_F < \rho_E < \frac{1}{2}$, then $\frac{\rho_E(1-\rho_E)}{(1-2\rho_E)^2} - \frac{\rho_F(1-\rho_F)}{(1-2\rho_F)^2} > 0$. Therefore, $R(\eta,\theta)$ is a decreasing function of $Q(\eta-\theta)$ when $Q(\eta-\theta) \in (0, 0.5]$ and increasing function of $Q(\eta-\theta)$ when $Q(\eta-\theta) \in [0.5, 1)$. The supremum of the FI ratio,

$$
\sup(R) = \frac{\rho_E(1-\rho_E)(1-2\rho_F)^2}{\rho_F(1-\rho_F)(1-2\rho_E)^2},
\tag{4.6}
$$

is achieved when $Q(\eta-\theta)$ approaches to 0 or 1. However, such choices of Q are not desirable in that they result in $f(\eta-\theta) = -\frac{dQ(\eta-\theta)}{d\eta} = 0$ and further the FI at the FC, $NI(\theta,\eta,\rho_F)$=0, indicating that the FC does not obtain any useful information for estimation either. Nevertheless, as R is a continuous function of Q, to achieve the design goal, we can choose $Q(\eta-\theta)$ close to 0 or 1 and increase the number of sensors N. In other words, we need to design η and N jointly to realize maximum achievable performance at the FC and secrecy against eavesdropper. Meanwhile, we discuss the minimized FI ratio in Appendix E.

Also notice that the FI ratio limit in Equation (4.6) is exactly the same with the KLD ratio derived in distributed detection in Section 3.3.3.

4.2.2 Asymptotic Perfect Secrecy

In order to achieve APS against eavesdropper, we require

$$\mathbf{I}(\mathbf{W};\theta) = NI(\eta,\theta,\rho_E) \to 0, \ \ N \to \infty. \tag{4.7}$$

Naturally, we also require the WSN to have an asymptotic perfect estimation at the FC, i.e.,

$$\mathbf{I}(\mathbf{V};\theta) = NI(\eta,\theta,\rho_E) \to \infty, \ \ N \to \infty. \tag{4.8}$$

Notice that when the FC has noiseless channels such that $\rho_F = 0$, the maximum FI ratio $\sup(R) = \infty$, indicating it is possible to simultaneously achieve both asymptotic perfect estimation and asymptotic perfect secrecy by choosing the appropriate η as a function of N. Next, we demonstrate how to do so for the case where the observation noises are Gaussian distributed, with similar design approaches employed for other noise distributions.

4.3 Estimation in Gaussian Noise under Classical Settings

We now consider the case where the observation noise, Z_i, follows the standard Gaussian distribution with zero mean and unit variance, where $f(x) = \frac{1}{\sqrt{2\pi}}e^{\frac{-x}{2}}$.

According to the bounds on Mills ratio [70], $x+1 > \frac{f(x)}{Q(x)} > x$, for $x > 0$, we have

$$\frac{f(x)}{xQ(x)} \to 1, \text{ as } x \to \infty. \tag{4.9}$$

For the FC with noiseless channel, the total FI is

$$\begin{aligned}\mathbf{I}(\mathbf{V};\theta) &\propto N(\eta-\theta)f(\eta-\theta)\\ &= N(\eta-\theta)\frac{1}{\sqrt{2\pi}}e^{\frac{-(\eta-\theta)^2}{2}}.\end{aligned}$$

For the noisy eavesdropper, the corresponding FI is

$$\begin{aligned}\mathbf{I}(\mathbf{W};\theta) &\propto Nf^2(\eta-A)\\ &= N\left(\frac{1}{2\pi}e^{\frac{-(\eta-\theta)^2}{2}}\right)^2.\end{aligned}$$

By choosing $\eta = \sqrt{(1+\mu)\log N}$, where $\eta \gg \theta$ and $\mu \in [0,1]$, for a fixed but unknown θ, and N sufficiently large, $e^{\frac{-(\eta-\theta)^2}{2}} \propto N^{-\frac{1+\mu}{2}}$.

For the FC,

$$\begin{aligned}\mathbf{I}(\mathbf{V};\theta) &\propto N\sqrt{(1+\mu)\log N}\,N^{-\frac{1+\mu}{2}}\\ &= N^{\frac{1-\mu}{2}}\sqrt{(1+\mu)\log N}.\end{aligned} \tag{4.10}$$

For eavesdropper,

$$\begin{aligned}\mathbf{I}\left(\mathbf{W};\theta\right) &\propto N\left(N^{-\frac{1+\mu}{2}}\right)^2 \\ &= N^{-\mu}.\end{aligned} \tag{4.11}$$

Now, the problem becomes how to choose μ such that $\mathbf{I}(\mathbf{V};\theta) \to \infty$ for asymptotic perfect estimation and $\mathbf{I}(\mathbf{W};\theta) \to 0$ for asymptotic perfect secrecy.

If $\mu = 0$,

$$\begin{aligned}&\mathbf{I}\left(\mathbf{V};\theta\right) \propto N^{\frac{1}{2}}\sqrt{\log N} \to \infty, \text{ when } N \to \infty. \\ &\mathbf{I}\left(\mathbf{W};\theta\right) \propto N\left(N^{-\frac{1}{2}}\right)^2 \to \mathcal{O}(1), \text{ when } N \to \infty.\end{aligned}$$

In this case, the secrecy at eavesdropper can be constrained to a constant, however, the performance of the FC can still be guaranteed to be asymptotic perfect at the rate of $\mathcal{O}\left(\sqrt{N}\right)$.

If we choose $\mu = \frac{1}{3}$,

$$\begin{aligned}&\mathbf{I}\left(\mathbf{V};\theta\right) \propto N^{\frac{1}{3}}\sqrt{\frac{4}{3}\log N} \to \infty, \text{ when } N \to \infty. \\ &\mathbf{I}\left(\mathbf{W};\theta\right) \propto N^{\frac{-1}{3}} \to 0, \text{ when } N \to \infty.\end{aligned}$$

Hence, both asymptotic perfect secrecy and asymptotic perfect estimation can be achieved under standard Gaussian observation noise. Similarly, when $\mu = 2$, the constraints can be satisfied as well.

When $\mu = 1$, the performance of the FC is guaranteed, however, eavesdropper's performance diminishes at the rate of $1/N$.

4.3.1 Performance Comparison in Detection and Estimation

In Section 3.3.2, we summarized the detection performance for the FC and eavesdropper, as

$$\begin{cases} \mathbb{D}_E \propto N^{-\mu}, \\ \mathbb{D}_F \propto N^{\frac{1-\mu}{2}}. \end{cases}$$

where $(0 < \mu < 1)$.

And in this section, we derived the estimation performance as,

$$\begin{cases} \mathbf{I}\left(\mathbf{W};\theta\right) \propto N^{-\mu}, \\ \mathbf{I}\left(\mathbf{V};\theta\right) \propto N^{\frac{1-\mu}{2}}\sqrt{(1+\mu)\log N}. \end{cases}$$

where $(0 \leq \mu \leq 1)$. We can see that the inference performance trade-off for the FC and eavesdropper under detection and estimation are almost the same.

4.4 Estimation under Bayesian Framework

We analyzed the case where the parameter θ is fixed but unknown, now we continue to investigate the case where θ is a random variable. In this case, the prior density about parameter θ should be considered in estimation. Similar to the classical case, the Bayesian CRLB (BCRLB) at the FC and at eavesdropper are defined, respectively, as

$$\begin{aligned}\mathrm{BCRLB}_F\left(\mathbf{V};\theta\right) &= \left(\mathbf{I}\left(\mathbf{V};\theta\right)\right)^{-1} = \left(\int_{-\infty}^{\infty} NI(\eta,\theta,\rho_F)p\left(\theta\right)d\theta + I(\lambda)\right)^{-1} \\ \mathrm{BCRLB}_E\left(\mathbf{W};\theta\right) &= \left(\mathbf{I}\left(\mathbf{W};\theta\right)\right)^{-1} = \left(\int_{-\infty}^{\infty} NI(\eta,\theta,\rho_E)p\left(\theta\right)d\theta + I(\lambda)\right)^{-1}\end{aligned} \tag{4.12}$$

where

$$I(\lambda) = \int \left(\frac{\partial \log p\left(\theta\right)}{\partial\theta}\right)^2 p(\theta)d\theta,$$

and $p\left(\theta\right)$ is the prior density of the parameter θ and $I\left(\eta,\theta,\rho\right)$ is from Equation (4.4).

4.4.1 $\theta \sim \mathcal{N}(0,1)$

For the case that $\theta \sim \mathcal{N}(0,1)$ such that $p\left(\theta\right) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{\theta^2}{2}\right)$ and because the channel of the FC is noiseless, we can simplify the FI at the FC for one sensor as follows,

$$I\left(\eta,\theta,\rho_F\right) = \frac{f^2\left(\eta-\theta\right)}{Q\left(\eta-\theta\right)\left(1-Q\left(\eta-\theta\right)\right)}$$

Assume the observation noise has unit variance, $I\left(\eta,\theta,\rho_F\right)$ is maximized when $\eta-\theta=0$, therefore, $Q\left(\eta-\theta\right)=\frac{1}{2}$.

$$\begin{aligned}I\left(\eta,\theta,\rho_F\right) &= \frac{\left(\frac{1}{\sqrt{2\pi}}\right)^2}{\frac{1}{2}\frac{1}{2}} \\ &= \frac{2}{\pi}.\end{aligned}$$

Therefore, the total FI at the FC is $NI\left(\eta,\theta,\rho_F\right) = \frac{2N}{\pi}$, which indicates that the total FI at the FC is $O(N)$ at most.

However, what if $Q(\eta-\theta) \neq \frac{1}{2}$? How should we choose η to satisfy the secrecy constraints under Bayesian framework? Similarly, utilize Mills ratio in Equation (4.9) and set η as $\sqrt{(1+\mu)\log N}$, where $0 \leq \mu \leq 1$. Hence, the total FI at the FC,

$$\begin{aligned}\mathbf{I}\left(\mathbf{V};\theta\right) &= \int_{-\infty}^{\infty} N\left(\eta-\theta\right) f\left(\eta-\theta\right) p\left(\theta\right) d\theta \\ &\propto \int_{-\infty}^{\infty} N\sqrt{(1+\mu)\log N} N^{-\frac{1+\mu}{2}} p\left(\theta\right) d\theta \\ &= N^{\frac{1-\mu}{2}}\sqrt{(1+\mu)\log N} \rightarrow \infty, \text{ as } N \rightarrow \infty.\end{aligned}$$

For eavesdropper, the total FI,

$$\begin{aligned}\mathbf{I}\left(\mathbf{W};\theta\right) &= \int_{-\infty}^{\infty} N f^{2}\left(\eta-\theta\right) p\left(\theta\right) d\theta \\ &\propto \int_{-\infty}^{\infty} N\left(N^{-\frac{1+\mu}{2}}\right)^{2} p\left(\theta\right) d\theta \\ &= N^{-\mu}.\end{aligned}$$

Therefore,

$$\begin{cases}\mathbf{I}\left(\mathbf{W};\theta\right) \rightarrow 0; & 0<\mu\leq 1 \\ \mathbf{I}\left(\mathbf{W};\theta\right) \rightarrow \mathcal{O}\left(1\right); & \mu=0\end{cases}$$

Similar to the classical setting case, the equations show that the secrecy constrains can be satisfied by choosing the appropriate threshold under Bayesian framework.

4.5 Experimental Results

In this section, we compare the estimation performance under two settings, one is a classical setting where the parameter is fixed but unknown; the other is a Bayesian setting where θ is a random variable.

4.5.1 Fixed but Unknown Parameter in Gaussian Noise

We first compare the estimation performance at eavesdropper and at the FC via the distributed estimation of a fixed but unknown parameter with zero mean additive white Gaussian noise. Specifically, the sensor observations are given in Equation (4.1), where $Z_i \sim \mathcal{N}(0,1)$ is the normalized observation noise following a standard Gaussian distribution. Both the FC and eavesdropper employ Maximum Likelihood Estimation (MLE) to obtain $\hat{\theta}_F$ and $\hat{\theta}_E$ based on $\mathbf{V}$ and $\mathbf{W}$, respectively. The two MSE estimates are

$$\begin{aligned} \hat{\theta}_F &= \left(\eta - Q^{-1}\left(\frac{\bar{V}-\rho_F}{1-2\rho_F}\right)\right) \\ \hat{\theta}_E &= \left(\eta - Q^{-1}\left(\frac{\bar{W}-\rho_E}{1-2\rho_E}\right)\right), \end{aligned} \tag{4.13}$$

where $\bar{V}$, $\bar{W}$ are the mean of received outputs for the FC and eavesdropper, respectively.

We first examine the system secrecy when the FC has a perfect channel, $\rho_F = 0$, eavesdropper has a noisy channel, $\rho_E = 0.40$, and the threshold $\eta = \sqrt{\frac{4}{3}\log N}$. First, the FI as a function of N for $\theta = 1$ is displayed in Figure 4.2. We see that the FI at the FC is increasing with the number of sensors, while the FI at eavesdropper is close to zero, consistent with the proofs for Equation (4.10) and (4.11).

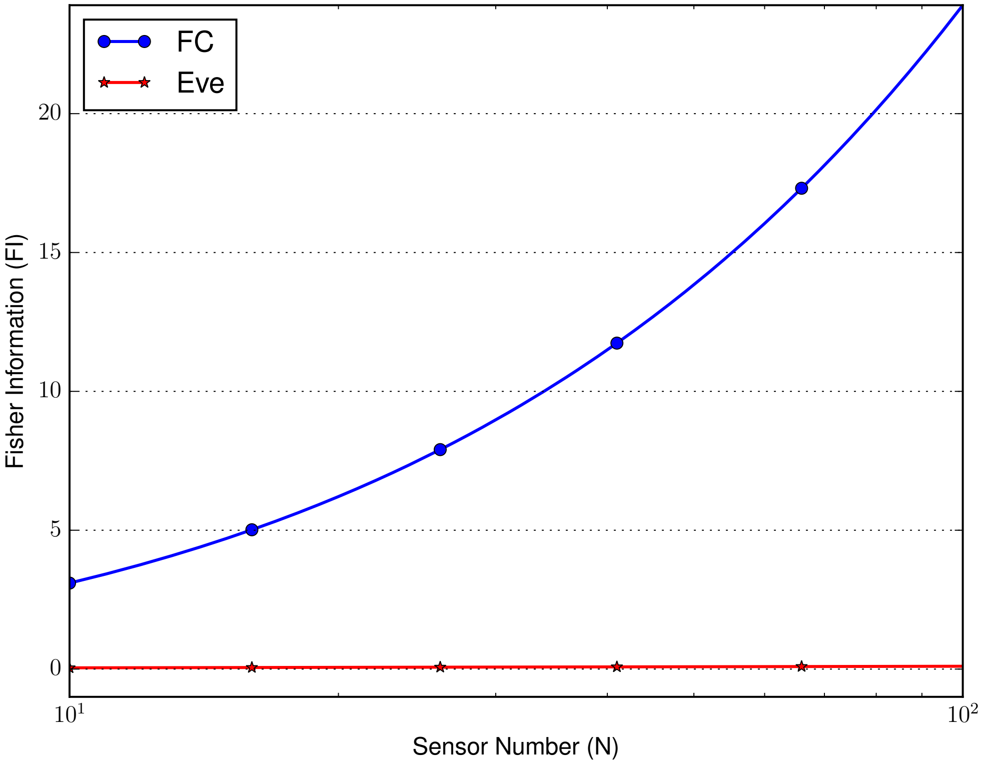

Figure 4.2: Total Fisher Information for the FC and eavesdropper with different number of sensors given $\theta = 1$, $\eta = \sqrt{\frac{4}{3}\log N}$ under classical setting. As the number of sensors grow, the FI at the FC increases significantly while the FI at eavesdropper is close to zero.

Under the same conditions of η, ρ_E and ρ_F, via Monte-Carlo simulation with 1000 trials, we plot the resulting mean and MSE of the estimated parameters θ_F and θ_E by the FC and eavesdropper in Figure 4.3 and Figure 4.4, respectively, where $\theta \in [0, 1.4]$, and the number of sensors is fixed at $N = 100$. In both figures, the trends show that eavesdropper, with a larger BSC BER, cannot accurately estimate θ. Meanwhile, the FC can almost perfectly estimate the parameter, where the estimated parameter mean is close to the ground truth in Figure 4.3 and the MSE is close to zero in Figure

4.4.

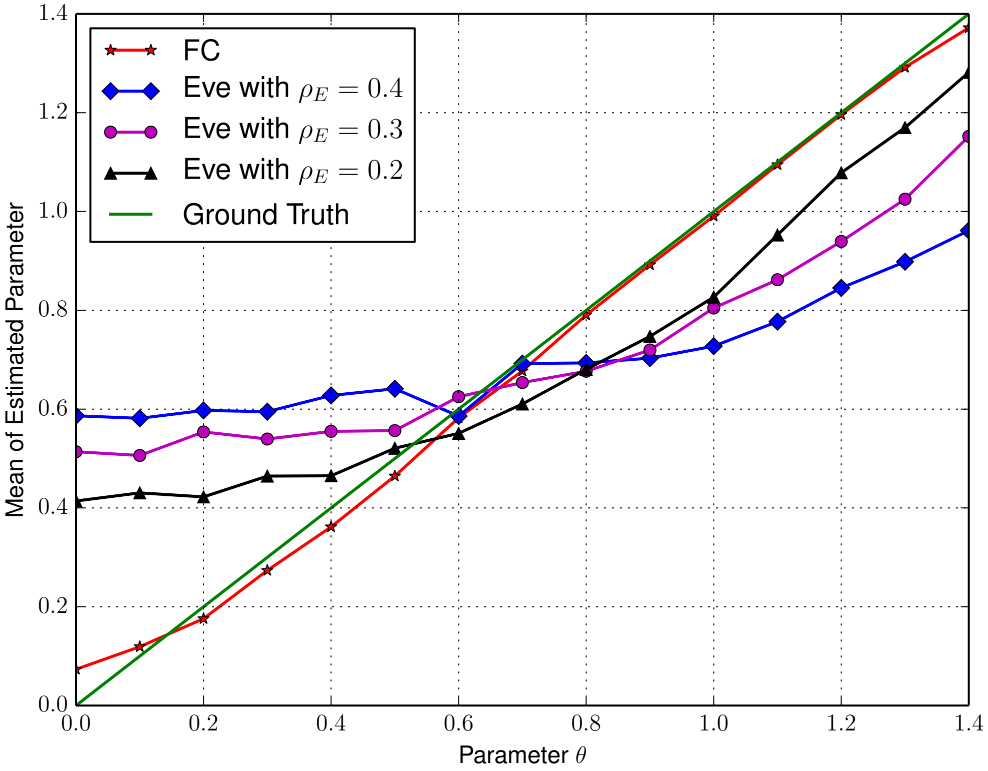

Figure 4.3: Mean of estimated signal by the FC and eavesdropper with different bit error rates under classical setting, where $\eta = \sqrt{\frac{4}{3}\log N}$, $N = 100$. The FC's estimation is close to the ground truth, while for eavesdropper, the estimations are off even the one with small bit error rate.

In Figure 4.5, we plot the exact MSE against the number of sensors for both the FC and eavesdropper, where the channel of the FC is noiseless, $\rho_F = 0$, and eavesdropper's channel is noisy, $\rho_E = 0.4$. Meanwhile, SNR = 0 dB, the threshold η is set as $\sqrt{\log N}$. We can see that the MSE of the FC diminishes close to zero while the MSE of eavesdropper are much higher than the FC's as the number of sensors increases from 10 to 100. As the performance metric, the MSEs of the FC and eavesdropper in

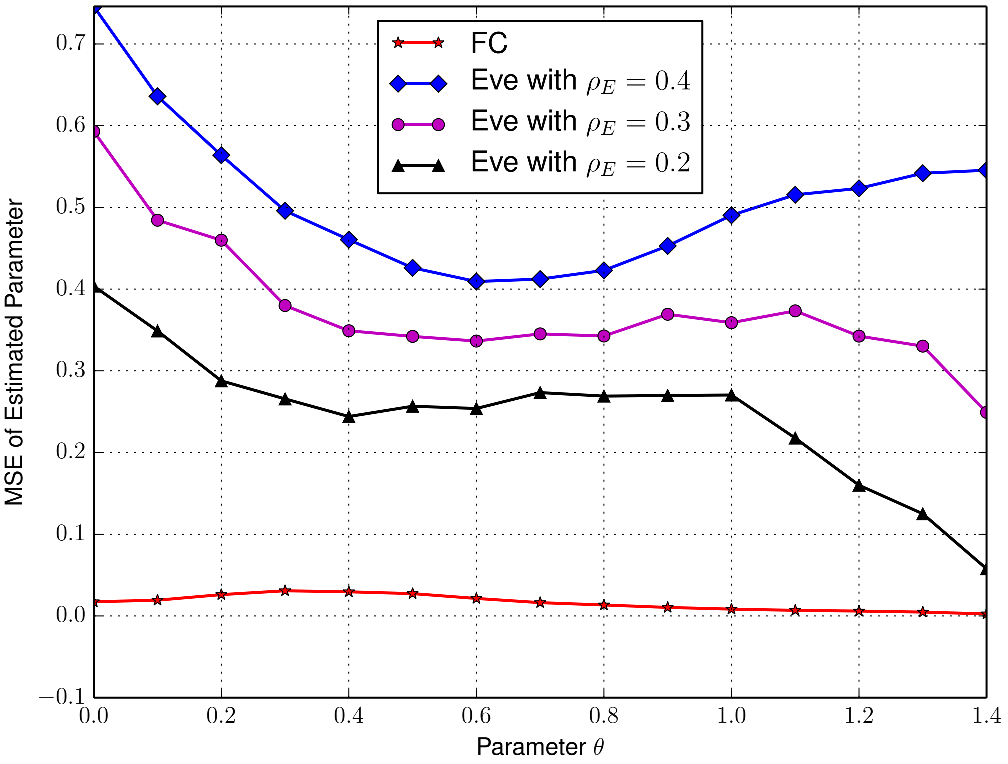

Figure 4.4: MSE of estimated signals by the FC and eavesdroppers with different bit error rates under classical setting, where the threshold, $\eta = \sqrt{\frac{4}{3}\log N}$, $N = 100$.

this figure indicate that the FC significantly outperforms eavesdropper in estimating the parameter θ with merely 10 to 100 sensors.

4.5.2 Simulations under Bayesian Framework

We continue comparing the performance at the FC and eavesdropper under Bayesian setting where the parameter θ is a random variable. Here, we assume that the parameter, θ, follows standard Gaussian distribution, $\mathcal{N}(0,1)$, and the observation noise also follows Gaussian distribution, $\mathcal{N}\left(0,\sqrt{2}\right)$. The FC channel is set to noiseless,

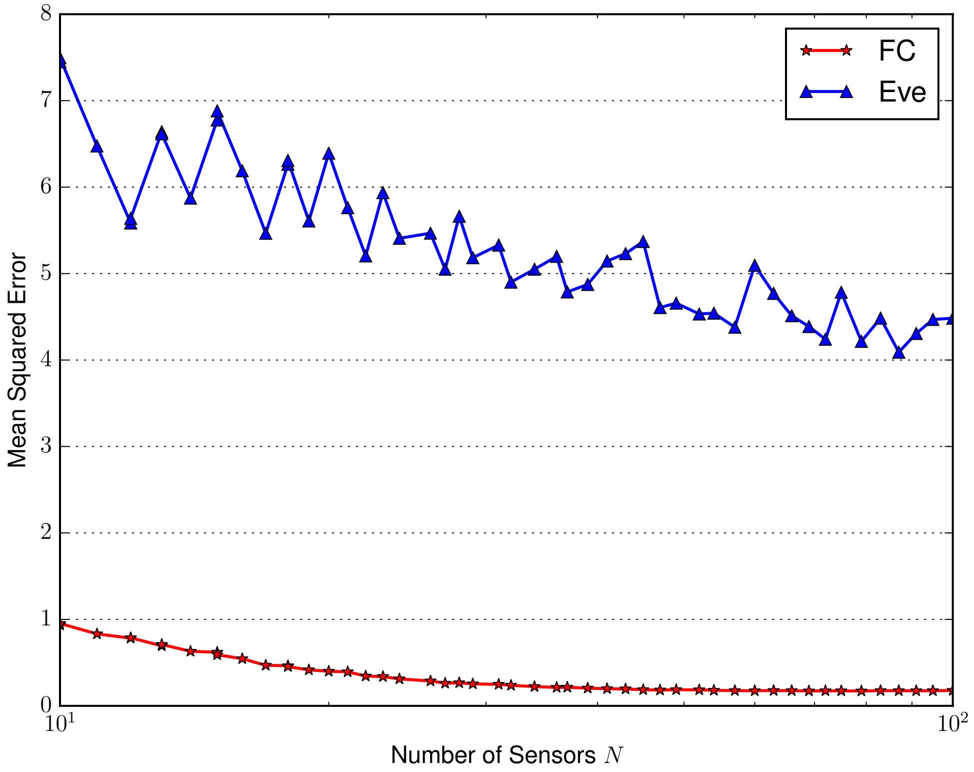

Figure 4.5: Exact Mean squared error for both the FC and eavesdropper with different number of sensors under classical setting, where the parameter, $\theta = 1$ and the threshold, $\eta = \sqrt{\log N}$.

$\rho_F = 0$ and $\rho_E = 0.4$, the threshold $\eta = \sqrt{\log N}$.

In Figure 4.6, we plot Fisher information with respect to the number of sensors for the FC and eavesdropper. Similar to the case of classical setting, the FI at the FC is increasing as the number of sensors increases from 10 to 100, meanwhile the FI at eavesdropper is close to zero. The growth rate at the FC is proportional to $\sqrt{N}$, which is consistent with the proof in Section 4.4.1.

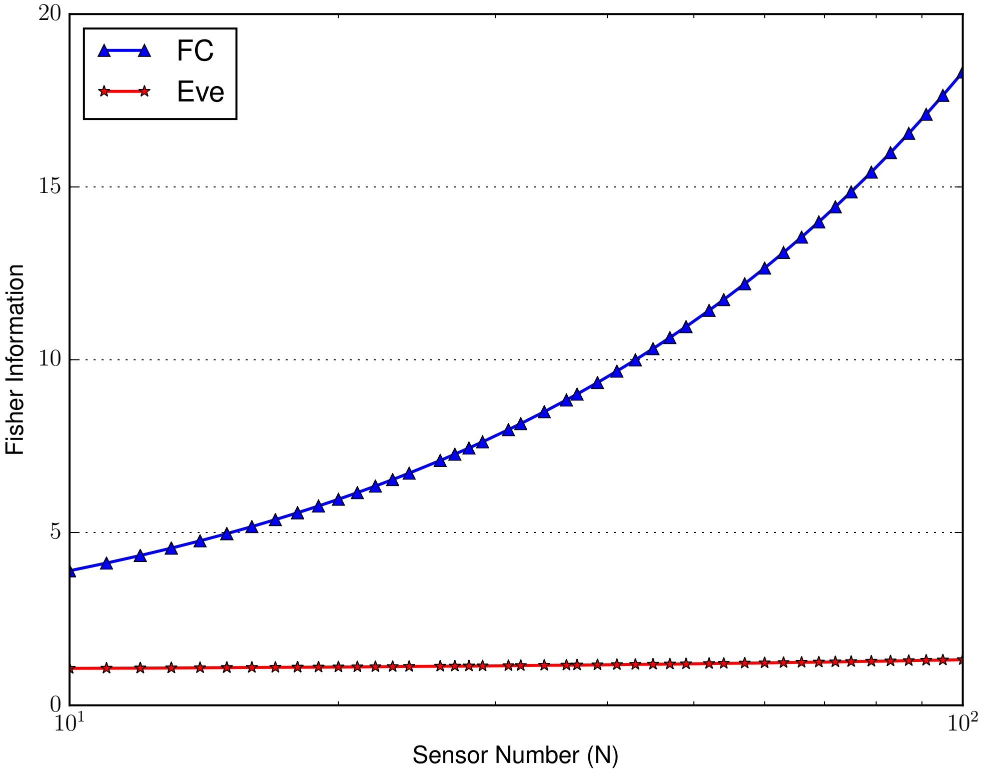

Figure 4.6: Total Fisher Information for both the FC and eavesdropper with different number of sensors under Bayesian framework, where $\eta = \sqrt{\log N}$.

CHAPTER 5

SECRECY CONSTRAINED DISTRIBUTED INFERENCE WITH FADING CHANNEL MODELS

In Chapter 3 and 4, we consider the secrecy constrained distributed detection and distributed estimation, respectively, where local sensors and the FC are connected through a parallel binary symmetric channel. In this chapter, we take fading channel into consideration because fading channels are also widely used in wireless communications for modeling scattered signals that reach a receiver by multiple paths. Hence, in this chapter, we consider the distributed inference problems under secrecy constraints with Rayleigh fading with BPSK signaling. This chapter is organized as the follows. In Section 5.1, we consider the secrecy constrained distributed detection with Rayleigh fading with BPSK signaling. Similar analysis are given in Section 5.2 for distributed estimation. We present simulations results in Section 5.3.

5.1 Secrecy Constrained Distributed Detection with Parallel Rayleigh Fading Binary Symmetric Channel

Similar to the case in Chapter 3, the network consists of N sensors connected in parallel to a FC via a set of parallel accessible channels. X_i and $p_k\left(X_i\right) = p\left(X_i; H_k\right)$ are the sensor observation and the pdf under hypothesis H_k at sensor i, respectively, where $k = 0, 1$ and $i = 1, 2, \ldots, N$. Sensors employ the same quantization rule as set in Equation (4.2). The wireless communication channels between each sensor and the FC are assumed to be Rayleigh fading BSCs, where the channel gain for the FC and eavesdropper are $h_{F,i}$ and $h_{E,i}$, respectively, and the corresponding BERs are $\rho_{F,i}$ and $\rho_{E,i}$. The sensors observations and the channels are assumed to be conditionally i.i.d.. For each sensor, the probability of false alarm is α and the probability of detection is β, where

$$\alpha = P\left(U_i|H_0\right) = P\left(\frac{p_1\left(X_i\right)}{p_0\left(X_i\right)} \geq \eta|H_0\right)$$
$$\beta = P\left(U_i|H_1\right) = P\left(\frac{p_1\left(X_i\right)}{p_0\left(X_i\right)} \geq \eta|H_1\right)$$

At the FC level, the received decisions are,

$$P\left(V_i = 1|H_0\right) = \alpha_F = \rho_F + \left(1 - 2\rho_F\right)\alpha$$
$$P\left(V_i = 1|H_1\right) = \beta_F = \rho_F + \left(1 - 2\rho_F\right)\beta$$

where h_F is the channel gain for the FC and Z_i is the observation noise. Similarly for eavesdropper,

$$P\left(W_i = 1|H_0\right) = \alpha_E = \rho_E + \left(1 - 2\rho_E\right)\alpha$$
$$P\left(W_i = 1|H_1\right) = \beta_E = \rho_E + \left(1 - 2\rho_E\right)\beta$$

Assume the average SNR between local sensors and the FC to be ξ_F and the average SNR between sensors and eavesdropper be ξ_E, according to the derivation in [63] and since the channels are independent and identical, the bit error rate of the fading channel with binary phase-shift keying (BPSK) modulation (with coherent detection) is

$$\begin{aligned} \rho_E = \rho_{E,i} &= \frac{1-\Psi_E}{2} \\ \rho_F = \rho_{F,i} &= \frac{1-\Psi_F}{2}, \end{aligned} \tag{5.1}$$

where,

$$\begin{aligned} \Psi_E &= \sqrt{\frac{\xi_E}{1+\xi_E}} \\ \Psi_F &= \sqrt{\frac{\xi_F}{1+\xi_F}} \end{aligned} \tag{5.2}$$

Recall from Chapter 3.3, we analyze the maximum achievable performance tradeoff between the FC and eavesdropper using the KLD ratio under Neyman-Pearson framework. Here, we use the same technique with Rayleigh fading channel model.

According to Appendix B, the KLD ratio for parallel channel is

$$\begin{aligned} R &= \frac{\mathbb{D}(\alpha_F, \beta_F)}{\mathbb{D}(\alpha_E, \beta_E)} \\ &= \frac{(1-2\rho_F)^2(1-\rho_E)\rho_E}{(1-2\rho_E)^2(1-\rho_F)\rho_F} \end{aligned} \tag{5.3}$$

Plugging in ρ_F and ρ_E in Equation (5.1), we have

$$
\begin{aligned}
R &= \frac{\left(1-2\tfrac{1-\Psi_F}{2}\right)^2\left(1-\tfrac{1-\Psi_E}{2}\right)\tfrac{1-\Psi_E}{2}}{\left(1-2\tfrac{1-\Psi_E}{2}\right)^2\left(1-\tfrac{1-\Psi_F}{2}\right)\tfrac{1-\Psi_F}{2}} \\
&= \frac{\Psi_F^2(1-\Psi_E)^2}{\Psi_E^2(1-\Psi_F)^2} \\
&= \frac{\frac{\xi_F}{1+\xi_F}\left(1-\frac{\xi_E}{1+\xi_E}\right)}{\frac{\xi_E}{1+\xi_E}\left(1-\frac{\xi_F}{1+\xi_F}\right)} \\
&= \frac{\frac{\xi_F}{1+\xi_F}\left(\frac{1}{1+\xi_E}\right)}{\frac{\xi_E}{1+\xi_E}\left(\frac{1}{1+\xi_F}\right)} \\
&= \frac{\xi_F}{\xi_E}.
\end{aligned}
\tag{5.4}
$$

Similar to the case of parallel binary symmetric channel, when $\rho_F \neq 0$ and $\rho_E \neq 0$, shown in Section 3.3.3, we derive the maximum achievable ratio and we summarize it as the following theorem.

Theorem 5. ***Maximum Achievable Performance Trade-off with Rayleigh Fading Channel*** *When both eavesdropper and the FC have noisy Rayleigh fading channels, with the average SNRs ξ_E and ξ_F, respectively, the secrecy constraints can only be achieved for certain thresholds T_E and T_F such that T_F/T_E is no more than the ratio, $R = \frac{\xi_F}{\xi_E}$, regardless of the number of sensors, N.*

Remark 2. To improve the secrecy in WSNs, one needs to either decrease SNR of eavesdropper, ξ_E, by using directional antenna or keys, or one can increase ξ_F. For example, if secrecy requirements for a WSN are $T_F = 10$ and $T_E = 1$, then $\xi_F/\xi_E \geq T_F/T_E = 10$ dB. It means that the SNR of the FC needs to be at least 10 dB better than eavesdropper's, which requires the FC to have high communication quality.

5.2 Secrecy Constrained Distributed Estimation with Parallel Rayleigh Fading Binary Symmetric Channel

Similarly for the distributed estimation with parallel Rayleigh fading binary symmetric channel, sensors observes one parameter θ. Here, we consider the classical setting, where θ is fixed but unknown,

$$X_i = \theta + Z_i, \;\; 1 = 1, 2, \ldots, N,$$

where Z_i is an additive i.i.d zero mean observation noise. We assume the sensor observations to be conditionally i.i.d given the underlying parameter θ, which results in

$$f(\mathbf{X}|\theta) = \prod_{i=1}^{N} f(X_i|\theta),$$

where $f(\mathbf{X}|\theta)$ and $f(X_i|\theta)$ are known probability density functions (pdfs) and X_i is the observation of sensor i.

Same with the assumption in Chapter 4, the sensors employ identical quantization rules such that

$$\eta_1 = \eta_2 = \cdots = \eta_N = \eta,$$

where the sensors observations are quantized to a single bit U_i according to the threshold η_i.

$$U_i = \begin{cases} 1, & X_i > \eta_i \\ & \\ 0, & X_i \le \eta_i \end{cases} \quad \forall i. \tag{5.5}$$

At the local sensor level, the

$$\Pr\left(U_i = 1|\theta\right) = \beta = \Pr\left(\theta + Z_i > \eta\right) = Q\left(\eta - \theta\right)$$
$$\Pr\left(U_i = 0|\theta\right) = 1 - \beta = 1 - \Pr\left(U_i = 1|\theta\right).$$

Meanwhile the received decisions at the receiver level are

$$\Pr(V_i = 1|\theta) = (1 - 2\rho_F)\Pr(U_i = 1|\theta) + \rho_F,$$
$$\Pr(W_i = 1|\theta) = (1 - 2\rho_E)\Pr(U_i = 1|\theta) + \rho_E.$$

Again, we consider the FI as the performance metric for the FC, $\mathbf{I}\left(\mathbf{V};\theta\right) = NI(\eta,\theta,\rho_F)$ and eavesdropper, $\mathbf{I}\left(\mathbf{W};\theta\right) = NI(\eta,\theta,\rho_E)$, where

$$I(\eta,\theta,\rho) = \frac{f^2(\eta-\theta)(1-2\rho)^2}{(\rho + (1-2\rho)Q(\eta-\theta))(1-\rho-(1-2\rho)Q(\eta-\theta))},$$

and we compute the FI ratio . From Section 4.2, we have the FI ratio, which is exactly the same with Equation (5.3). In other words, the supremum of the FI ratio is

$$\begin{aligned}\sup\left(R\right) &= \frac{\rho_E(1-\rho_E)(1-2\rho_F)^2}{\rho_F(1-\rho_F)(1-2\rho_E)^2} \\ &= \frac{\xi_F}{\xi_E}.\end{aligned}$$

Similarly, the secrecy of the network can be improved by either increasing the FC SNR or decreasing eavesdropper SNR.

This ratio is of help in achieving asymptotic perfect secrecy and asymptotic perfect estimation in that the requirements are

$$\begin{aligned}\mathbf{I}\left(\mathbf{W};\theta\right) = NI\left(\eta,\theta,\rho_E\right) \to 0,\ \ N \to \infty, \\ \mathbf{I}\left(\mathbf{V};\theta\right) = NI\left(\eta,\theta,\rho_F\right) \to \infty,\ \ N \to \infty,\end{aligned}$$

If ξ_E is zero, the maximum FI ratio is infinity, which means it is possible to achieve asymptotic perfect secrecy and asymptotic perfect estimation.

5.3 Experimental Results

In this section, we plot and evaluate the performance trade-off between the FC and eavesdropper for both the detection and estimation problems.

5.3.1 Secrecy Constrained Distributed Detection

For distributed detection problems, we consider a constant signal with zero mean additive white Gaussian noise. The sensor observations are given by

$$\begin{cases} H_1: & X_i = A + Z_i \\ H_0: & X_i = Z_i, \end{cases}$$

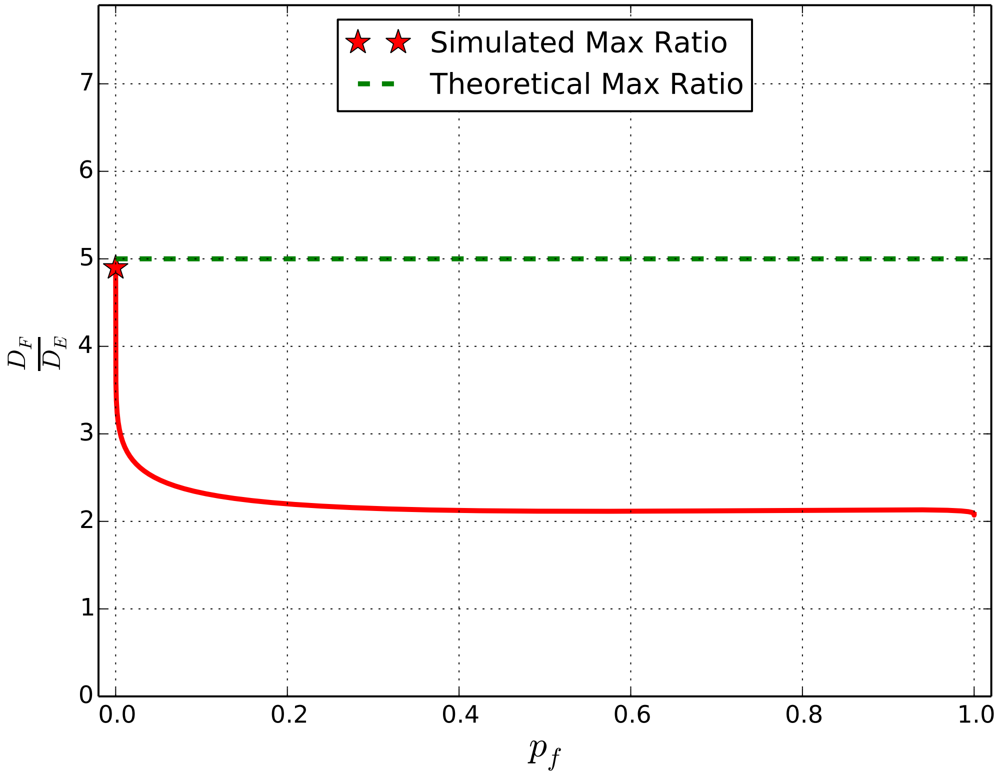

Figure 5.1: The maximum achievable KLD ratio for secrecy constrained distributed detection, $\xi_F = 0$ dB and $\xi_E = -7$ dB.

where $Z_i \sim \mathcal{N}(0,1)$ is the normalized observation noise following a standard Gaussian distribution, $A > 0$ is a fixed constant signal to be detected with signal-to-noise ratio, $\mathrm{SNR} = 20 \log_{10} A$ dB.

We first plot the maximum achievable KLD ratio, D_F/D_E, in Figure 5.1. SNR for the FC and eavesdropper are set as, $\xi_F = 0$ dB and $\xi_E = -7$ dB, respectively. From the figure, we can see that the simulated maximum KLD ratio is very close to the theoretical one derived in Equation (5.4).

In Figure (5.2), the actual KLD is plotted against the number of sensors N. As we expected, as the number of sensors grows, the FI at the FC increases faster than

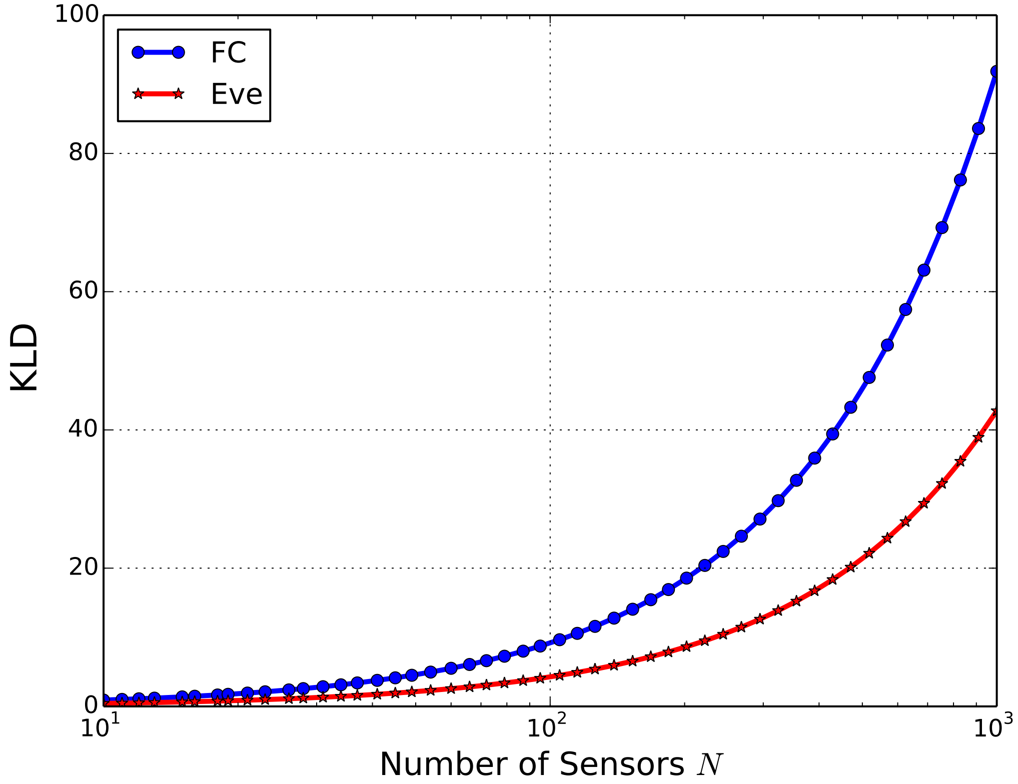

Figure 5.2: The actual KLD for the FC and eavesdropper with different number of sensors, $\xi_F = 0$ dB and $\xi_E = -7$ dB.

eavesdropper FI.

We show the maximum KLD ratio in Figure 5.3. The SNR of eavesdropper is fixed to -7 dB (corresponds to 0.2 in the figure), and the SNR of the FC increases to 20. We can see that it is linear between the maximum KLD ratio and SNR of the FC.

5.3.2 Secrecy Constrained Distributed Estimation

The estimation performance at eavesdropper and at the FC via the distributed estimation of a fixed but unknown signal with zero mean additive white Gaussian noise.

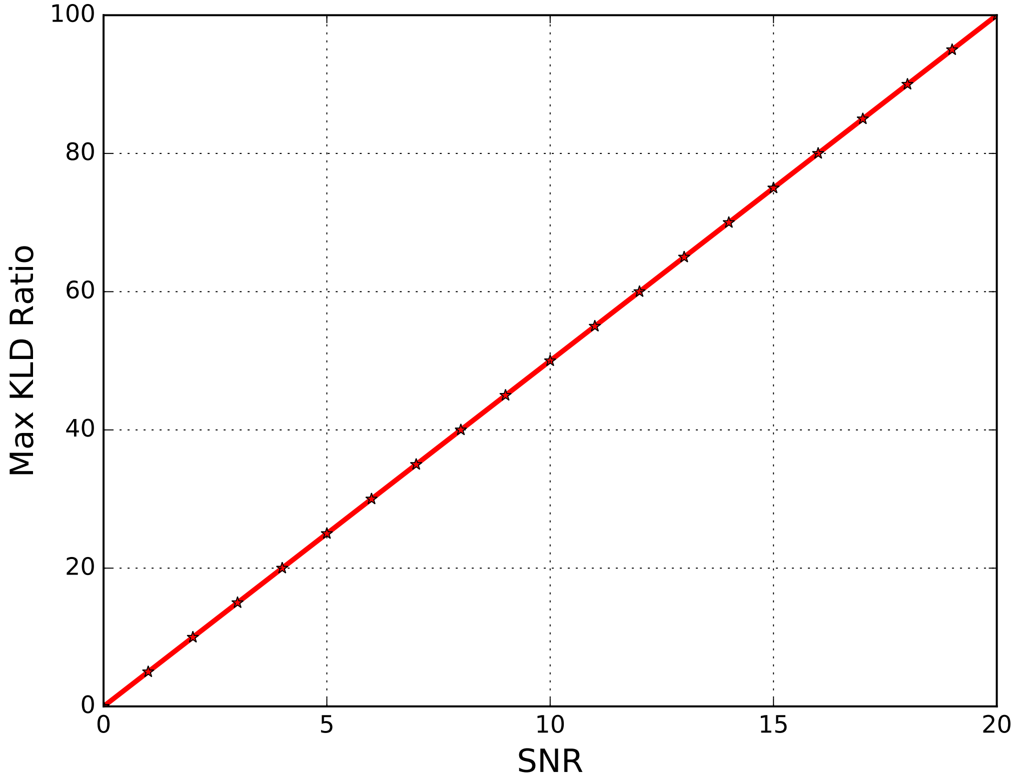

Figure 5.3: The maximum KLD ratio for the FC and eavesdropper with $\xi_E = -7$ dB.

Specifically, the sensor observations are given in Equation (4.1), where $Z_i \sim \mathcal{N}(0, 1)$ is the normalized observation noise following a standard Gaussian distribution. Both the FC and eavesdropper employ Maximum Likelihood Estimation (MLE) to obtain $\hat{\theta}_F$ and $\hat{\theta}_E$ based on $\mathbf{V}$ and $\mathbf{W}$, respectively. The two MSE estimates are

$$\begin{aligned} \hat{\theta}_F &= \left(\eta - Q^{-1}\left(\frac{\bar{V} - \rho_F}{1 - 2\rho_F}\right)\right) \\ \hat{\theta}_E &= \left(\eta - Q^{-1}\left(\frac{\bar{W} - \rho_E}{1 - 2\rho_E}\right)\right), \end{aligned} \tag{5.6}$$

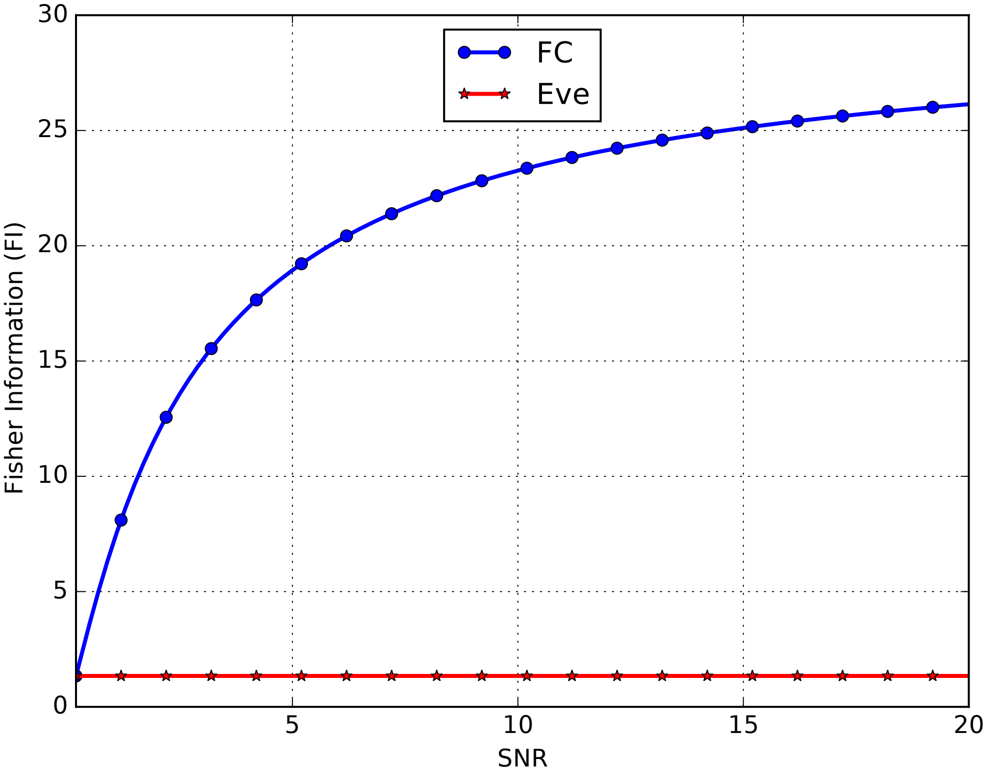

Figure 5.4: The actual FI for secrecy constrained distributed estimation with $\xi_E = -7$ dB, the number of sensors, $N = 100$.

where $\bar{V}$, $\bar{W}$ are the mean of received outputs for the FC and eavesdropper, respectively.

We show the actual estimation performance of the FC and eavesdropper in Figure 5.4. The SNR of eavesdropper is fixed to -7 dB, the number of sensors is 100, the threshold $\eta = \sqrt{\log N}$ and the SNR of the FC increases from 0.2 to 20. We can see that the FI at the FC is increasing with the increment of the SNR, which means the secrecy of the network can be improved by increasing the FC SNR.

CHAPTER 6

CONCLUSION

6.1 Summary

Comprised of a large number of low-cost, low-power, mobile and miniature sensors, WSNs are widely employed in many applications, such as environmental monitoring, health-care, and diagnostics of complex systems. For these applications, the data collected by local sensors are extremely sensitive, and care must be taken to prevent that information from being leaked to any malicious third parties, e.g., eavesdroppers. In WSNs, the sensor outputs are often transmitted across a wireless communication network to legitimate users such as a fusion center for final decision-making. However, because of wireless links across the network, data are vulnerable to security breaches.

Eavesdropping on wireless links between the sensor and the legitimate user by a third party (eavesdropper) is defined as an eavesdropping attacks. The reason we focus on eavesdropping attack is that it forms the basis or starting point for a large number of different, more malicious attack strategies. For example, if Byzantine, jamming attackers or intruders have reliable information provided by the eavesdropper, their subsequent attacks can be more efficient.

Hence, we focus on security issues for WSNs especially on secrecy constrained distributed inference in WSNs. The off-the-shelf solution for eavesdropping attack is cryptography techniques, public-key and symmetric key algorithms. However, due to the constraints on computational power, bandwidth, and time constraints, these algorithms could hardly be implemented for WSNs. Even though the symmetric keys algorithm consume relatively low-power, they do require the nodes to have the computational capability to perform the required tasks which may not be true for some of the nodes. Therefore, we resort to a physical-layer security approach which utilizes the characteristics of the physical layer, including transmission channels noises, and the information of the source. Additionally, physical-layer security for distributed detection is scalable due to the low computational complexity. Physical-layer security approaches can be used along with cryptography techniques to further enhance WSNs and make systems even more secure.

Chapter 3 considered the secrecy constrained distributed detection in WSNs under both the Neyman-Pearson and Bayesian frameworks using physical-layer security approaches by adjusting sensor optimality points. We analyzed the asymptotic detection performance and proposed a novel way of analyzing the maximum performance trade-off using KLD ratio between the FC and eavesdropper. Under the N-P framework, we showed that eavesdropper's detection performance can be limited such that her decision-making is no better than random guessing; meanwhile, the detection performance at the FC is guaranteed at the prespecified level. Similar analyses and proofs are provided under the Bayesian framework, where it was shown that eavesdropper can be constrained to an error probability level equal to her prior information. Additionally, we derived the asymptotic error exponent and showed that asymptotic perfect secrecy and asymptotic perfect detection are possible by increasing the number of sensors under both frameworks if the FC has noiseless channels to the sensors. The numerical results showed that with reasonable number of sensors, we can guarantee the detection performance of the FC to achieve the desired level, while the detectability at eavesdropper deteriorates significantly as the number of sensors increases. These results in this chapter can be applied to improve the secrecy of WSNs against eavesdropping attacks at the physical-layer.

Chapter 4 concentrated on the secrecy constrained distributed estimation in WSNs which are subject to an eavesdropping attack under both classical setting and Bayesian framework. The eavesdropper has access to all sensors outputs instead of partial access. The maximum achievable secrecy performance was derived and it was proved that under the condition that eavesdropper has a noisy channel and the FC has a noiseless channel, both APS and APE can be achieved. The secrecy design method in this dissertation might greatly enhance the secrecy in distributed estimation for large sensor networks.

Chapter 5 considered secrecy constrained distributed inference with Rayleigh fading binary symmetric channel models. We derived the maximum achievable performance ratio and show that the number of sensors does not affect this ratio. We showed that for both detection and estimation problems, asymptotic perfect secrecy cannot be achieved.

6.2 Future Research Topics

We investigated distributed detection and estimation under secrecy constraints for parallel channels and a few other research topics within this framework can be investigated as well.

- The design of distributed inference algorithms depends on the underlying sensor network topology. Different topologies requires different algorithms in designing secrecy rules for sensors and the FC. Therefore we considered parallel topology in this dissertation; however, the results can be extended to distributed inference with generalized topology, such as tree topology shown in Figure 1.3. The structure can be balanced or unbalanced. In a balanced structure, sensors at the same level have equal numbers of children where an unbalanced structure does not have such restriction. For tree-structured systems, each sensor observes the same phenomenon, quantizes their observations and then transmits the decision to their parent node [96], and the parent node keeps doing the same until the decisions are reached by the FC. Then global decision rule is designed so that the FC could make the final decision.

- In this dissertation, we considered the scenario that the eavesdropping attacker has a noisier channel than the FC, and has access to all the sensors outputs. We could also consider situations such that the attacker could have only partial access to the sensor decisions; however, the channel quality is not necessarily worse than the FC's.

- We could also consider the scenario where the FC and eavesdropper may have access to side information that allows them to improve their performance. In such a scenario, the same approach in this dissertation can still be employed. Special treatment utilizing the side information may help further improve the inference performance.

- We did not consider the model uncertainty in this dissertation. Considering real data at the sensors in a real application would make it more applicable. In such case, we may have to divide the data into training and testing sets, and apply machine learning algorithms to the problem. , Machine learning in distributed systems itself is an interesting topic and awaits to be further pursued.

- In this dissertation, we considered eavesdroppers, other more sophisticated cyber threats to WSN, such as intrusion, jamming or Byzantine, should also be investigated in the future. The interaction and relation among cyber attacks bring new dimensions to the distributed inference problem.

REFERENCES

[1] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci. Wireless sensor networks: A survey. *Comput. Netw.*, 38(4):393–422, March 2002.

[2] T.C. Aysal and K.E. Barner. Sensor data cryptography in wireless sensor networks. *Information Forensics and Security, IEEE Transactions on*, 3(2):273–289, June 2008.

[3] M. Baldi, M. Bianchi, N. Maturo, and F. Chiaraluce. A physical layer secured key distribution technique for ieee 802.11g wireless networks. *Wireless Communications Letters, IEEE*, 2(2):183–186, April 2013.

[4] M. J. Beevi. A fair survey on internet of things (iot). In *2016 International Conference on Emerging Trends in Engineering, Technology and Science (ICETETS)*, pages 1–6, Feb 2016.

[5] R. S. Blum, S. A. Kassm, and H. V. Poor. Distributed detection with multiple sensors ii. - advanced topics. *Proceedings of the IEEE*, 85(1):64–79, 1997.

[6] R.S. Blum, S.A. Kassam, and H.V. Poor. Distributed detection with multiple sensors i. advanced topics. *Proceedings of the IEEE*, 85(1):64–79, Jan 1997.

[7] Kechar Bouabdellah, Houache Noureddine, and Sekhri Larbi. Using wireless sensor networks for reliable forest fires detection. *Procedia Computer Science*, 19:794 – 801, 2013.

[8] J.A. Bucklew. *Large deviation techniques in decision, simulation, and estimation*. Wiley-interscience publication. Wiley, 1990.

[9] D. Castanon and D. Teneketzis. Distributed estimation algorithms for nonlinear systems. *IEEE Transactions on Automatic Control*, 30(5):418–425, May 1985.

[10] Z. Chair and P.K. Varshney. Optimal data fusion in multiple sensor detection systems. *Aerospace and Electronic Systems, IEEE Transactions on*, AES-22(1):98–101, Jan 1986.

[11] J. Chamberland and V.V. Veeravalli. Wireless sensors in distributed detection applications. *Signal Processing Magazine, IEEE*, 24(3):16–25, May 2007.

[12] J.-F. Chamberland and V.V. Veeravalli. Decentralized detection in sensor networks. *Signal Processing, IEEE Transactions on*, 51(2):407–416, Feb 2003.

[13] Y. A. Chau and E. Geraniotis. Distributed multisensor parameter estimation in dependent noise. *IEEE Transactions on Communications*, 40(2):373–384, Feb 1992.

[14] B. Chen, L. Tong, and P. K. Varshney. Channel-aware distributed detection in wireless sensor networks. *IEEE Signal Processing Magazine*, 23(4):16–26, July 2006.

[15] H. Chen and P. K. Varshney. Performance limit for distributed estimation systems with identical one-bit quantizers. *IEEE Transactions on Signal Processing*, 58(1):466–471, Jan 2010.

[16] Po-Ning Chen. General formulas for the neyman-pearson type-ii error exponent subject to fixed and exponential type-i error bounds. *Information Theory, IEEE Transactions on*, 42(1):316–323, Jan 1996.

[17] Q. Cheng, B. Chen, and P. K. Varshney. Detection performance limits for distributed sensor networks in the presence of nonideal channels. *IEEE Transactions on Wireless Communications*, 5(11):3034–3038, November 2006.

[18] I. Csiszár and J. Körner. *Information Theory: Coding Theorems for Discrete Memoryless Systems.* Cambridge University Press, 2011.

[19] Joan Daemen and Vincent Rijmen. Aes proposal: Rijndael, 1999.

[20] Hong-Ning Dai, Qiu Wang, Dong Li, and Raymond Chi-Wing Wong. On eavesdropping attacks in wireless sensor networks with directional antennas. *International Journal of Distributed Sensor Networks*, 2013, 2013.

[21] W. Diffie and M. Hellman. New directions in cryptography. *IEEE Trans. Inf. Theor.*, 22(6):644–654, September 2006.

[22] O. Dikmen, Z. Yang, and E. Oja. Learning the information divergence. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 37(7):1442–1454, July 2015.

[23] G. Dimitrakopoulos and P. Demestichas. Intelligent transportation systems. *IEEE Vehicular Technology Magazine*, 5(1):77–84, March 2010.

[24] Richard O. Duda, Peter E. Hart, and David G. Stork. *Pattern Classification, 2nd Ed.* 2001.

[25] D. Eastlake, 3rd and P. Jones. Us secure hash algorithm 1 (sha1), 2001.

[26] Etimad Fadel, V.C. Gungor, Laila Nassef, Nadine Akkari, M.G. Abbas Malik, Suleiman Almasri, and Ian F. Akyildiz. A survey on wireless sensor networks for smart grid. *Computer Communications*, 71:22 – 33, 2015.

[27] J. A. Gubner. Distributed estimation and quantization. *IEEE Transactions on Information Theory*, 39(4):1456–1459, Jul 1993.

[28] X. Guo, A. S. Leong, and S. Dey. Power allocation for distortion minimization in distributed estimation with security constraints. In *2014 IEEE 15th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, pages 299–303, June 2014.

[29] Z. Guo, F. Ye, J. Guo, Y. Liang, G. Xu, X. Zhang, and Y. Qian. A wireless sensor network for monitoring smart grid transmission lines. In *2014 23rd International Conference on Computer Communication and Networks (ICCCN)*, pages 1–6, Aug 2014.

[30] Nils Gura, Arun Patel, Arvinderpal Wander, Hans Eberle, and Sheueling Chang Shantz. *Comparing Elliptic Curve Cryptography and RSA on 8-bit CPUs*, pages 119–132. Springer Berlin Heidelberg, Berlin, Heidelberg, 2004.

[31] A. Hakimi, N. Hassan, K. Anwar, A. Zakaria, and A. Ashraf. Development of real-time patient health (jaundice) monitoring using wireless sensor network. In *2016 3rd International Conference on Electronic Design (ICED)*, pages 404–409, Aug 2016.

[32] Lingxuan Hu and David Evans. Using directional antennas to prevent wormhole attacks. In *Proceedings of the Network and Distributed System Security Symposium, NDSS 2004, San Diego, California, USA*, 2004.

[33] H. Jeon, J. Choi, S. W. McLaughlin, and J. Ha. Channel aware encryption and decision fusion for wireless sensor networks. In *2011 IEEE International Workshop on Information Forensics and Security*, pages 1–6, Nov 2011.

[34] H. Jeon, D. Hwang, J. Choi, H. Lee, and J. Ha. Secure type-based multiple access. *IEEE Transactions on Information Forensics and Security*, 6(3):763–774, Sept 2011.

[35] Joshua Kiepert Jim A. Hall, Michael L. Pook and Sin Ming Loo. Monitoring aircraft cabin particulate matter using a wireless sensor network. 43rd International Conference on Environmental Systems. Vail, 2013.

[36] B. Kailkhura, V. S. Siddhardh Nadendla, and P. K. Varshney. Distributed Inference in the Presence of Eavesdroppers: A Survey. *ArXiv e-prints*, February 2015.

[37] B. Kaliski. A survey of encryption standards. *IEEE Micro*, 13(6):74–81, Dec 1993.

[38] Chris Karlof, Naveen Sastry, and David Wagner. Tinysec: A link layer security architecture for wireless sensor networks. In *Proceedings of the 2Nd International Conference on Embedded Networked Sensor Systems*, SenSys '04, pages 162–175, New York, NY, USA, 2004. ACM.

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

[40] S. K. Khaitan and J. D. McCalley. Design techniques and applications of cyberphysical systems: A survey. *IEEE Systems Journal*, 9(2):350–365, June 2015.

[41] U. A. Khan and A. M. Stankovic. Secure distributed estimation in cyber-physical systems. In *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 5209–5213, May 2013.

[42] M. Khanafer, M. Guennoun, and H. T. Mouftah. Wsn architectures for intelligent transportation systems. In *2009 3rd International Conference on New Technologies, Mobility and Security*, pages 1–8, Dec 2009.

[43] Neal Koblitz. Elliptic curve cryptosystems. *Mathematics of Computation*, 48(177):203–209, January 1987.

[44] S. Kullback and R. A. Leibler. On information and sufficiency. *Ann. Math. Statist.*, 22(1):79–86, 03 1951.

[45] K. Sampath Kumar and H. Li. Distributed estimation in sensor networks over binary symmetric channels. In *2009 Conference Record of the Forty-Third Asilomar Conference on Signals, Systems and Computers*, pages 265–269, Nov 2009.

[46] X. Li, H. Huang, and Y. Sun. Dritri: An in-vehicle wireless sensor network platform for daily health monitoring. In *2016 IEEE SENSORS*, pages 1–3, Oct 2016.

[47] Z. Li and T. J. Oechtering. Privacy-aware distributed bayesian detection. *IEEE Journal of Selected Topics in Signal Processing*, 9(7):1345–1357, Oct 2015.

[48] Zuxing Li, T.J. Oechtering, and J. Jalden. Parallel distributed neyman-pearson detection with privacy constraints. In *Communications Workshops (ICC), 2014 IEEE International Conference on*, pages 765–770, June 2014.

[49] A. Liu and P. Ning. Tinyecc: A configurable library for elliptic curve cryptography in wireless sensor networks. In *2008 International Conference on Information Processing in Sensor Networks (ipsn 2008)*, pages 245–256, April 2008.

[50] Y. Lu, A. Savvaris, A. Tsourdos, and M. Bevilacqua. Vibration energy harvesters for wireless sensor networks for aircraft health monitoring. In *2016 IEEE Metrology for Aerospace (MetroAeroSpace)*, pages 25–32, June 2016.

[51] A. Maimaris and G. Papageorgiou. A review of intelligent transportation systems from a communications technology perspective. In *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*, pages 54–59, Nov 2016.

[52] R. Malekian and A. H. Abdullah. A novel approach to improve signal to noise ratio based on sector antenna in radio networks. In *2012 32nd International Conference on Distributed Computing Systems Workshops*, pages 250–253, June 2012.

[53] S. Marano, V. Matta, and P.K. Willett. Distributed detection with censoring sensors under physical layer secrecy. *Signal Processing, IEEE Transactions on*, 57(5):1976–1986, May 2009.

[54] R. Mariappan, P. V. N. Reddy, and C. Wu. Cyber physical system using intelligent wireless sensor actuator networks for disaster recovery. In *2015 International Conference on Computational Intelligence and Communication Networks (CICN)*, pages 95–99, Dec 2015.

[55] S. Mathur, A. Reznik, Chunxuan Ye, R. Mukherjee, A. Rahman, Y. Shah, W. Trappe, and N. Mandayam. Exploiting the physical layer for enhanced security [security and privacy in emerging wireless networks]. *Wireless Communications, IEEE*, 17(5):63–70, October 2010.

[56] Alfred J. Menezes, Scott A. Vanstone, and Paul C. Van Oorschot. *Handbook of Applied Cryptography*. CRC Press, Inc., Boca Raton, FL, USA, 1st edition, 1996.

[57] Victor S. Miller. Use of elliptic curves in cryptography. In *Advances in Cryptology*, CRYPTO '85, pages 417–426, London, UK, UK, 1986. Springer-Verlag.

[58] V.S.S. Nadendla, Hao Chen, and P.K. Varshney. Secure distributed detection in the presence of eavesdroppers. In *Signals, Systems and Computers (ASILOMAR), 2010 Conference Record of the Forty Fourth Asilomar Conference on*, pages 1437–1441, Nov 2010.

[59] National and Technology (nist). SKIPJACK and KEA Algorithm Specifications Version 2.0. Technical report, National Institute of Standards and Technology (NIST), Mai 1998.

[60] X. Nguyen, M. J. Wainwright, and M. I. Jordan. Nonparametric decentralized detection using kernel methods. *IEEE Transactions on Signal Processing*, 53(11):4053–4066, Nov 2005.

[61] National Science Foundation (NSF). Cyber physical systems nsf10515. [Online]. Available: http://www.nsf.gov/pubs/2010/nsf10515/nsf10515.htm, 2014.

[62] W. Pawgasame. A survey in adaptive hybrid wireless sensor network for military operations. In *2016 Second Asian Conference on Defence Technology (ACDT)*, pages 78–83, Jan 2016.

[63] J.G. Proakis. *Digital Communications*. Electrical engineering series. McGraw-Hill, 2001.

[64] M. L. Rajaram, E. Kougianos, S. P. Mohanty, and P. Sundaravadivel. A wireless sensor network simulation framework for structural health monitoring in smart cities. In *2016 IEEE 6th International Conference on Consumer Electronics - Berlin (ICCE-Berlin)*, pages 78–82, Sept 2016.

[65] A.S. Rawat, P. Anand, H. Chen, and P. K. Varshney. Collaborative spectrum sensing in the presence of byzantine attacks in cognitive radio networks. *IEEE Transactions on Signal Processing*, 59(2):774–786, July 2011.

[66] R. Rivest. The md5 message-digest algorithm, 1992.

[67] R. L. Rivest, A. Shamir, and L. Adleman. A method for obtaining digital signatures and public-key cryptosystems. *Commun. ACM*, 21(2):120–126, February 1978.

[68] Ronald L. Rivest. The rc5 encryption algorithm, 1995.

[69] L. A. Rossi, B. Krishnamachari, and C. C. J. Kuo. Hybrid data and decision fusion techniques for model-based data gathering in wireless sensor networks [environmental monitoring applications]. In *IEEE 60th Vehicular Technology*

Conference, 2004. VTC2004-Fall. 2004, volume 7, pages 4616–4620 Vol. 7, Sept 2004.

[70] M. R. Sampford. Some inequalities on mill's ratio and related functions. *Ann. Math. Statist.*, 24(1):130–132, 03 1953.

[71] I. Shahid and P. Yahampath. Distributed joint source-channel coding of correlated binary sources in wireless sensor networks. In *2011 8th International Symposium on Wireless Communication Systems*, pages 236–240, Nov 2011.

[72] Divya Sharma, Sandeep Verma, and Kanika Sharma. Network topologies in wireless sensor networks: A review 1.

[73] E. Shi and A. Perrig. Designing secure sensor networks. *Wireless Commun.*, 11(6):38–43, December 2004.

[74] R. Soosahabi and M. Naraghi-Pour. Scalable phy-layer security for distributed detection in wireless sensor networks. In *Vehicular Technology Conference (VTC Fall), 2012 IEEE*, pages 1–5, Sept 2012.

[75] Kamil Staniec and Grzegorz Debita. *Novel Modifications in WSN Network Design for Improved SNR and Reliability*, pages 243–259. Springer Berlin Heidelberg, Berlin, Heidelberg, 2012.

[76] G. Sudha, R. Prakash, A. B. Ganesh, and S. V. Girish. Network coding based real time wireless sensor network for environmental monitoring. In *2016 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET)*, pages 1269–1272, March 2016.

[77] A. Swami, Q. Zhao, Hong Y.-W., and L. Tong. *Wireless Sensor Networks:Signal Processing and Communications Perspectives*. John Wiley & Sons, Ltd., 2007.

[78] H. Taha and E. Alsusa. A mimo precoding based physical layer security technique for key exchange encryption. In *Vehicular Technology Conference (VTC Spring), 2015 IEEE 81st*, pages 1–5, May 2015.

[79] H. L. V. Trees. *Detection, Estimation, and Modulation Theory, Part I*. Wiley-Interscience, 2007.

[80] B. Triki, S. Rekhis, and N. Boudriga. An rfid based system for the detection of sybil attack in military wireless sensor networks. In *2014 World Congress on Computer Applications and Information Systems (WCCAIS)*, pages 1–2, Jan 2014.

[81] J. N. Tsitsiklis. Decentralized detection. In H. V. Poor and J. B. Thomas, editors, *Advances in Signal Processing*, volume 2, pages 297–344. JAI Press, 1993.

[82] J. N. Tsitsiklis and M. Athans. On the complexity of decentralized decision making and detection problems. In *The 23rd IEEE Conference on Decision and Control*, pages 1638–1641, Dec 1984.

[83] J.N. Tsitsiklis. Extremal properties of likelihood-ratio quantizers. *Communications, IEEE Transactions on*, 41(4):550–558, Apr 1993.

[84] H.L. Van Trees. *Detection, Estimation, and Modulation Theory*. Detection, Estimation, and Modulation Theory. Wiley, 2004.

[85] Pramod K. Varshney. *Distributed Detection and Data Fusion.* Springer-Verlag New York, Inc., Secaucus, NJ, USA, 1st edition, 1996.

[86] Venugopal V. Veeravalli1 and Pramod K. Varshney. Distributed inference in wireless sensor networks. *Phil Trans R Soc A*, pages 100–117, 2012.

[87] A. Vempaty, H. He, B. Chen, and P. K. Varshney. On quantizer design for distributed bayesian estimation in sensor networks. *IEEE Transactions on Signal Processing*, 62(20):5359–5369, Oct 2014.

[88] R. Viswanathan and P. K. Varshney. Distributed detection with multiple sensors i. - fundamentals. *Proceedings of the IEEE*, 85(1):54 – 63, 1997.

[89] R. Viswanathan and P.K. Varshney. Distributed detection with multiple sensors i. fundamentals. *Proceedings of the IEEE*, 85(1):54–63, Jan 1997.

[90] A. S. Wander, N. Gura, H. Eberle, V. Gupta, and S. C. Shantz. Energy analysis of public-key cryptography for wireless sensor networks. In *Third IEEE International Conference on Pervasive Computing and Communications*, pages 324–328, March 2005.

[91] J. Wang, I. S. Ahn, Y. Lu, and G. Staskevich. A new distributed algorithm for environmental monitoring by wireless sensor networks with limited communication. In *2016 IEEE SENSORS*, pages 1–3, Oct 2016.

[92] Yong Wang, G. Attebury, and B. Ramamurthy. A survey of security issues in wireless sensor networks. *Commun. Surveys Tuts.*, 8(2):2–23, April 2006.

[93] D. Warren and P. Willett. Optimum quantization for detector fusion: some proofs, examples and pathology. *Journal of the Franklin Institute*, 336:323–359, 1999.

[94] A. Willsky, M. Bello, D. Castanon, B. Levy, and G. Verghese. Combining and updating of local estimates and regional maps along sets of one-dimensional tracks. *IEEE Transactions on Automatic Control*, 27(4):799–813, Aug 1982.

[95] M. Won, H. Ra, T. Park, and S. H. Son. Modeling random deployment in wireless sensor networks for infrastructure-less cyber physical systems. In *2014 IEEE International Conference on Cyber-Physical Systems, Networks, and Applications*, pages 81–86, Aug 2014.

[96] Ming Xiang and Chongzhao Han. Optimization of distributed detection networks with tree structures. In *Information Fusion, 2002. Proceedings of the Fifth International Conference on*, volume 1, pages 164–169 vol.1, July 2002.

[97] Elias Yaacoub and Adnan Abu-Dayya. Wireless sensor networks-technology and protocols. 2012-09-06.

[98] A. Yener and S. Ulukus. Wireless physical-layer security: Lessons learned from information theory. *Proceedings of the IEEE*, 103(10):1814–1825, Oct 2015.

[99] Su Yi, Yong Pei, and Shivkumar Kalyanaraman. On the capacity improvement of ad hoc wireless networks using directional antennas. In *Proceedings of the 4th ACM International Symposium on Mobile Ad Hoc Networking &Amp; Computing*, MobiHoc '03, pages 108–116, New York, NY, USA, 2003. ACM.

[100] Jennifer Yick, Biswanath Mukherjee, and Dipak Ghosal. Wireless sensor network survey. *Computer Networks*, 52(12):2292 – 2330, 2008.

[101] L. Ying, R. Srikant, and G. E. Dullerud. Distributed symmetric function computation in noisy wireless sensor networks. *IEEE Transactions on Information Theory*, 53(12):4826–4833, Dec 2007.

[102] R. Zhang, D. Yuan, and Y. Wang. A health monitoring system for wireless sensor networks. In *2007 2nd IEEE Conference on Industrial Electronics and Applications*, pages 1648–1652, May 2007.

[103] Y. Zhang, Y. Shen, H. Wang, J. Yong, and X. Jiang. On secure wireless communications for iot under eavesdropper collusion. *Automation Science and Engineering, IEEE Transactions on*, PP(99):1–13, 2015.

APPENDIX A

KLD ANALYSIS

We first analyze $(\alpha, \beta) \approx (0,0)$,

$$\begin{aligned} D\left(\bar{\alpha}, \bar{\beta}\right) &= \bar{\alpha} \ln \frac{\bar{\alpha}}{\bar{\beta}} + (1-\bar{\alpha}) \ln \frac{1-\bar{\alpha}}{1-\bar{\beta}} \\ &= \bar{\alpha} \ln \left(1 + \frac{\bar{\alpha}-\bar{\beta}}{\bar{\beta}}\right) + (1-\bar{\alpha}) \ln \left(1 + \frac{\bar{\beta}-\bar{\alpha}}{1-\bar{\beta}}\right). \end{aligned}$$

Because of the properties of binary symmetric channel, we have

$$\begin{aligned} \bar{\alpha} &= \bar{\rho} + (1-2\bar{\rho})\alpha \\ \bar{\beta} &= \bar{\rho} + (1-2\bar{\rho})\beta. \end{aligned}$$

Given the following Taylor series expansion

$$\ln(1+x) = \sum_{n=1}^{\infty} (-1)^{n-1} \frac{x^n}{n},$$

and setting $x \approx 0$, then $\ln(1+x) \approx x - \frac{1}{2}x^2$.

A.0.1 $(\alpha, \beta) \approx (0,0)$ and $\bar{\rho} \neq 0$

It can be shown that if $\bar{\rho} \neq 0$,

$$\begin{aligned}\frac{\bar{\alpha}-\bar{\beta}}{\bar{\beta}} &\approx 0\\ \frac{\bar{\beta}-\bar{\alpha}}{1-\bar{\beta}} &\approx 0\\ \frac{\bar{\alpha}}{\bar{\beta}} &\approx 1\\ \frac{1-\bar{\alpha}}{1-\bar{\beta}} &\approx 1\end{aligned} \tag{A.1}$$

$$\begin{aligned} D\left(\bar{\alpha},\bar{\beta}\right) &\approx \bar{\alpha}\left(\frac{\bar{\alpha}-\bar{\beta}}{\bar{\beta}}-\frac{1}{2}\left(\frac{\bar{\alpha}-\bar{\beta}}{\bar{\beta}}\right)^2\right)+(1-\bar{\alpha})\left(\frac{\bar{\beta}-\bar{\alpha}}{1-\bar{\beta}}-\frac{1}{2}\left(\frac{\bar{\beta}-\bar{\alpha}}{1-\bar{\beta}}\right)^2\right)\\ &= \left(\bar{\beta}-\bar{\alpha}\right)\left(\frac{1-\bar{\alpha}}{1-\bar{\beta}}-\frac{\bar{\alpha}}{\bar{\beta}}\right)-\frac{(\bar{\beta}-\bar{\alpha})^2}{2}\left(\frac{1-\bar{\alpha}}{(1-\bar{\beta})^2}+\frac{\bar{\alpha}}{\bar{\beta}^2}\right).\end{aligned}$$

Because of Equation (A.1),

$$\begin{aligned} D\left(\bar{\alpha},\bar{\beta}\right) &\approx \frac{\left(\bar{\beta}-\bar{\alpha}\right)^2}{\left(1-\bar{\beta}\right)\bar{\beta}}-\frac{\left(\bar{\beta}-\bar{\alpha}\right)^2}{2}\left(\frac{1}{1-\bar{\beta}}+\frac{1}{\bar{\beta}}\right)\\ &= \frac{1}{2}\frac{\left(\bar{\beta}-\bar{\alpha}\right)^2}{\left(1-\bar{\beta}\right)\bar{\beta}}\\ &= \frac{1}{2}\frac{\left(\beta-\alpha\right)^2\left(1-2\bar{\rho}\right)^2}{\left(1-\bar{\rho}-\left(1-2\bar{\rho}\right)\beta\right)\left(\bar{\rho}+(1-2\bar{\rho})\beta\right)}.\end{aligned} \tag{A.2}$$

When $(\alpha,\beta)\approx(0,0)$ and $\frac{\beta}{\alpha}\to\infty$,

$$D\left(\bar{\alpha},\bar{\beta}\right)\approx\frac{1}{2}\frac{\beta^2(1-2\bar{\rho})^2}{(1-\bar{\rho})\bar{\rho}}.$$

Using the same procedure, it is straightforward to show that

$$D\left(\bar{\alpha},\bar{\beta}\right)\approx D\left(\bar{\beta},\bar{\alpha}\right)\approx\frac{1}{2}\frac{\beta^2(1-2\bar{\rho})^2}{(1-\bar{\rho})\bar{\rho}}. \tag{A.3}$$

A.0.2 $(\alpha, \beta) \approx (0,0)$ and $\bar{\rho} = 0$

If $\bar{\rho} = 0$, one needs to use a different approximation as follows,

$$\begin{aligned} D\left(\bar{\alpha}, \bar{\beta}\right) &= \bar{\alpha} \ln \frac{\bar{\alpha}}{\bar{\beta}} + (1-\bar{\alpha}) \ln \frac{1-\bar{\alpha}}{1-\bar{\beta}} \\ &= \bar{\alpha} \ln \frac{\bar{\alpha}}{\bar{\beta}} + (1-\bar{\alpha}) \ln \left(1 + \frac{\bar{\beta}-\bar{\alpha}}{1-\bar{\beta}}\right) \\ &\approx \bar{\alpha} \ln \frac{\bar{\alpha}}{\bar{\beta}} + (1-\bar{\alpha}) \frac{\bar{\beta}-\bar{\alpha}}{1-\bar{\beta}} \\ &\approx \bar{\alpha} \ln \frac{\bar{\alpha}}{\bar{\beta}} + \bar{\beta} - \bar{\alpha} \\ &\approx \bar{\alpha} \frac{\bar{\beta}}{\bar{\alpha}} \\ &= \bar{\beta} \\ &= \beta \;\text{ as } \beta \gg \alpha \quad \text{when } \alpha \to 0. \end{aligned} \tag{A.4}$$

and

$$\begin{aligned} D(\bar{\beta}, \bar{\alpha}) &= \bar{\beta} \ln \frac{\bar{\beta}}{\bar{\alpha}} + \left(1-\bar{\beta}\right) \ln \left(1 + \frac{\bar{\alpha}-\bar{\beta}}{1-\bar{\alpha}}\right) \\ &\approx \bar{\beta} \ln \frac{\bar{\beta}}{\bar{\alpha}} + \left(1-\bar{\beta}\right) \left(\frac{\bar{\alpha}-\bar{\beta}}{1-\bar{\alpha}}\right) \\ &\approx \bar{\beta} \ln \frac{\bar{\beta}}{\bar{\alpha}} + \bar{\alpha} - \bar{\beta} \\ &= \bar{\beta} \left(\ln \frac{\bar{\beta}}{\bar{\alpha}} + \frac{\bar{\alpha}}{\bar{\beta}} - 1\right) \\ &\approx \bar{\beta} \left(\ln \frac{\bar{\beta}}{\bar{\alpha}} - 1\right). \end{aligned}$$

Since, $\bar{\rho} = 0$, $\bar{\beta} = \beta$, $\bar{\alpha} = \alpha$,

$$D(\bar{\beta}, \bar{\alpha}) \approx \beta \left(\ln \frac{\beta}{\alpha} - 1\right). \tag{A.5}$$

The proof for $(\alpha, \beta) \approx (1, 1)$ relative to $D(\bar{\alpha}, \bar{\beta})$ and $D(\bar{\beta}, \bar{\alpha})$ can be shown in a similar fashion.

APPENDIX B

PROOF OF MAXIMUM ACHIEVABLE PERFORMANCE TRADE-OFF UNDER NEYMAN- PEARSON FRAMEWORK

If $\ln\left(p_1(x)/p_0(x)\right)$ is unbounded below, then as $\alpha \to 1$, $\eta \to 0$, applying L'Hopital's rule to the ratio $R(\alpha)$,

$$\begin{aligned}
R(1) &= \lim_{\alpha \to 1} \frac{\frac{d}{d\alpha} D\left(\alpha_F, \beta_F\right)}{\frac{d}{d\alpha} D\left(\alpha_E, \beta_E\right)} \\
&= \lim_{\alpha \to 1} \frac{\frac{d\alpha_F}{d\alpha}\left(\frac{d}{d\alpha_F} D\left(\alpha_F, \beta_F\right)\right)}{\frac{d\alpha_E}{d\alpha}\left(\frac{d}{d\alpha_E} D\left(\alpha_E, \beta_E\right)\right)} \\
&= \lim_{\substack{\alpha \to 1 \\ \beta \to 1 \\ \eta \to 0}} \frac{\left(1 - 2\rho_F\right)\left(\eta \frac{\beta_F - \alpha_F}{(1-\beta_F)\beta_F} + \log \frac{\alpha_F (1-\beta_F)}{\beta_F (1-\alpha_F)}\right)}{\left(1 - 2\rho_E\right)\left(\eta \frac{\beta_E - \alpha_E}{(1-\beta_E)\beta_E} + \log \frac{\alpha_E (1-\beta_E)}{\beta_E (1-\alpha_E)}\right)} =
\end{aligned}$$

$$\lim_{\substack{\alpha \to 1 \\ \beta \to 1 \\ \eta \to 0}} \frac{\left(1 - 2\rho_F\right)^2 (\beta - \alpha)\left(\frac{\eta}{(1-\beta_F)\beta_F} + \frac{\log\left(1 + \frac{\alpha_F - \beta_F}{\beta_F(1-\alpha_F)}\right)}{\beta_F - \alpha_F}\right)}{\left(1 - 2\rho_E\right)^2 (\beta - \alpha)\left(\frac{\eta}{(1-\beta_E)\beta_E} + \frac{\log\left(1 + \frac{\alpha_E - \beta_E}{\beta_E(1-\alpha_E)}\right)}{\beta_E - \alpha_E}\right)}.$$

Since

$$\lim_{x \to 0} \frac{\log(1+x)}{x} = 1,$$

$$\lim_{\substack{\alpha \to 1 \\ \beta \to 1}} \frac{\log\left(1 + \frac{\alpha_F - \beta_F}{\beta_F(1-\alpha_F)}\right)}{\beta_F - \alpha_F} = \lim_{\substack{\alpha \to 1 \\ \beta \to 1}} \frac{-1}{\beta_F\left(1 - \alpha_F\right)},$$

Therefore,

$$\begin{aligned}
R\left(1\right) &= \lim_{\substack{\alpha \to 1 \\ \beta \to 1 \\ \eta \to 0}} \frac{\left(1 - 2\rho_F\right)^2\left(\frac{\eta}{(1-\beta_F)\beta_F} - \frac{1}{\beta_F(1-\alpha_F)}\right)}{\left(1 - 2\rho_E\right)^2\left(\frac{\eta}{(1-\beta_E)\beta_E} - \frac{1}{\beta_E(1-\alpha_E)}\right)} \\
&= \frac{\left(1 - 2\rho_F\right)^2\left(1 - \rho_E\right)\rho_E}{\left(1 - 2\rho_E\right)^2\left(1 - \rho_F\right)\rho_F}.
\end{aligned}$$

Similarly, if $\ln\left(p_1(x)/p_0(x)\right)$ is unbounded above, then as $\alpha \to 0$, $\eta \to \infty$,

$$R(0) = \frac{\left(1 - 2\rho_F\right)^2 \left(1 - \rho_E\right) \rho_E}{\left(1 - 2\rho_E\right)^2 \left(1 - \rho_F\right) \rho_F} = R(1). \tag{B.1}$$

APPENDIX C

SECRECY AND DETECTION TRADE-OFF UNDER BAYESIAN FRAMEWORK

Plug β_E, α_E from Equation (3.4) in Equation (3.18), we have

$$\left(\frac{(1-2\rho_E)\beta+\rho_E}{(1-2\rho_E)\alpha+\rho_E}\right)^N = \left(\frac{\beta+\left(\frac{\rho_E}{1-2\rho_E}\right)}{\alpha+\left(\frac{\rho_E}{1-2\rho_E}\right)}\right)^N \leq \frac{\pi_0}{\pi_1}.$$

Let $\tau = \frac{\rho_E}{1-2\rho_E}$, then

$$\begin{aligned}&\left(\frac{\beta+\tau}{\alpha+\tau}\right)^N \leq \frac{\pi_0}{\pi_1},\\ \Rightarrow&\beta+\tau \leq \left(\frac{\pi_0}{\pi_1}\right)^{\frac{1}{N}}(\alpha+\tau),\\ \Rightarrow&\beta \leq \left(\frac{\pi_0}{\pi_1}\right)^{\frac{1}{N}}(\alpha+\tau)-\tau,\end{aligned} \tag{C.1}$$

or

$$\alpha \geq \left(\frac{\pi_0}{\pi_1}\right)^{\frac{1}{N}}(\beta+\tau)-\tau. \tag{C.2}$$

In this case, we can choose the randomization between $A = (\alpha_A, \beta_A)$ and $B = (\alpha_B, \beta_B)$ in Figure C.1, where the region of operation is the region inside the two red curves. However, it is not ideal for Bayesian framework in that it does not consider the prior information. We only need to consider the limiting property of A and B. Utilizing the fact that $A \to \tilde{A}$ and $B \to \tilde{B}$ as $N \to \infty$ where $\tilde{A} = (0, \beta_{\tilde{A}})$ and $\tilde{B} = (\alpha_{\tilde{B}}, 1)$ and $P_e(A, B) > P_e(\tilde{A}, \tilde{B})$, we consider $\tilde{A}$ and $\tilde{B}$ instead.

C.0.1 Case1: At Point $\tilde{A} = (0, \beta_F)$

In this case, the probability of false alarm of the FC is zero, since $\alpha_F = 0$, from Equation (3.3), we know that $\alpha = 0$, and after inserting in Equation (C.1)

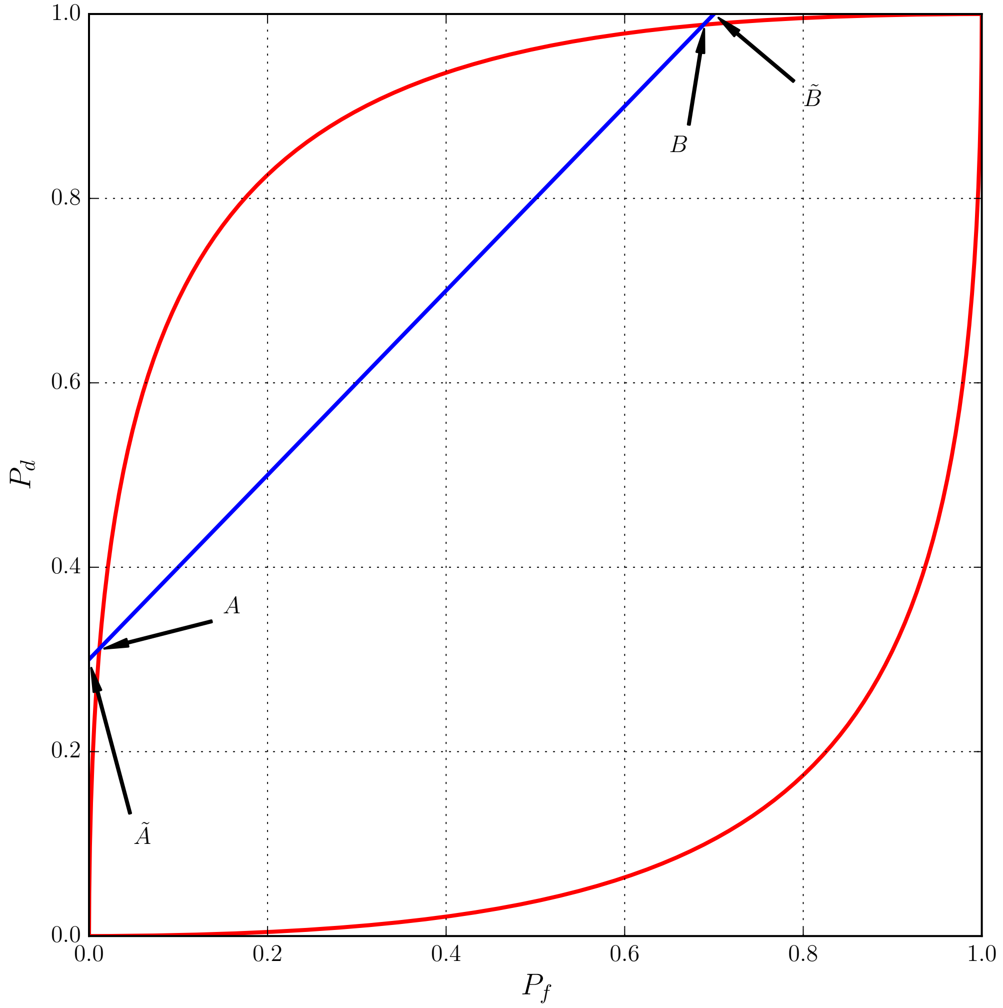

Figure C.1: Operation Region is the inside area of two ROC curves.

$$\beta_F = \beta = \left(\left(\frac{\pi_0}{\pi_1} \right)^{\frac{1}{N}} - 1 \right) \tau. \tag{C.3}$$

Therefore, the FC decision rule should be

$$\begin{cases} \text{decides } H_0 \text{ when receiving all zeros,} \\ \text{decides } H_1 \text{ otherwise.} \end{cases}$$

The probability of error of the FC is

$$\begin{aligned} P_e &= \alpha_F^N \pi_0 + (1-\beta_F)^N \pi_1, \\ &= (1-\beta_F)^N \pi_1. \end{aligned}$$

$P_m = 1 - P_d$, where $P_d = \beta_F$ from Equation (C.3), with

$$\begin{aligned} \lim_{N\to\infty} \ln P_m &= \lim_{N\to\infty} N \ln \left(1 - \left(\left(\frac{\pi_0}{\pi_1}\right)^{\frac{1}{N}} - 1\right)\tau\right) \\ &= \lim_{N\to\infty} \frac{\ln\left(1 - \left(\left(\frac{\pi_0}{\pi_1}\right)^{\frac{1}{N}} - 1\right)\tau\right)}{\frac{1}{N}}. \end{aligned}$$

Apply L'Hopital's rule, and let $x = \frac{1}{N}$ we have

$$\begin{aligned} \lim_{x\to 0} \ln P_m &= \lim_{x\to 0} \frac{\left(1 - \left(\left(\frac{\pi_0}{\pi_1}\right)^x - 1\right)\tau\right)'}{1 - \left(\left(\frac{\pi_0}{\pi_1}\right)^x - 1\right)\tau} \\ &= \lim_{x\to 0} \left(-\tau\left(\left(\frac{\pi_0}{\pi_1}\right)^x - 1\right)\right)' \\ &= \lim_{x\to 0} \left(-\tau\left(\frac{\pi_0}{\pi_1}\right)^x\right)' \\ &= \lim_{x\to 0} \left(-\tau \ln\frac{\pi_0}{\pi_1}\left(\frac{\pi_0}{\pi_1}\right)^x\right) \\ &= -\tau \ln\frac{\pi_0}{\pi_1}. \end{aligned}$$

Therefore,

$$P_m = \left(\frac{\pi_0}{\pi_1}\right)^{-\tau} = \left(\frac{\pi_0}{\pi_1}\right)^{-\frac{\rho_E}{1-2\rho_E}}.$$

The probability of error of the FC at operating point $\tilde{A}$ is

$$P_e = P_m \pi_1 = \pi_1 \left(\frac{\pi_0}{\pi_1}\right)^{-\frac{\rho_E}{1-2\rho_E}}.$$

This shows that as long as $\rho_E < 0.5$, the probability of error of the FC at $\tilde{A}$ is not zero.

C.0.2 Case2: At Point $\tilde{B} = (\alpha_F, 1)$

The detection probability of the FC is one, plug $\beta = 1$ in Equation (C.2), we have $\alpha_F = \alpha = (1+\tau)\left(\frac{\pi_1}{\pi_0}\right)^{\frac{1}{N}} - \tau$. Thus, the FC decision rule should be

$$\begin{cases} \text{decides } H_1 \text{ when receiving all ones,} \\ \text{decides } H_0 \text{ otherwise.} \end{cases}$$

The probability of error of the FC is

$$\begin{aligned} P_e &= \alpha_F^N \pi_0 + (1-\beta_F)^N \pi_1, \\ &= \alpha_F^N \pi_0. \end{aligned}$$

With

$$\begin{aligned}\lim_{N\to\infty} \ln P_f &= \lim_{N\to\infty} N \ln\left((1+\tau)\left(\frac{\pi_0}{\pi_1}\right)^{\frac{1}{N}} - \tau\right),\\ &= \lim_{N\to\infty} \frac{\ln\left((1+\tau)\left(\frac{\pi_0}{\pi_1}\right)^{\frac{1}{N}} - \tau\right)}{\frac{1}{N}}.\end{aligned}$$

Apply L'Hopital's rule with $x = \frac{1}{N}$ we have

$$\begin{aligned}\lim_{x\to 0} \ln P_f &= \lim_{x\to 0} \frac{\left((1+\tau)\left(\frac{\pi_1}{\pi_0}\right)^{x}\right)'}{(1+\tau)\left(\frac{\pi_1}{\pi_0}\right)^{x} - \tau},\\ &= \lim_{x\to 0} \left((1+\tau)\left(\frac{\pi_1}{\pi_0}\right)^{x}\right)',\\ &= \lim_{x\to 0} (1+\tau)\ln\left(\frac{\pi_1}{\pi_0}\right)\left(\frac{\pi_1}{\pi_0}\right)^{x},\\ &= (1+\tau)\ln\frac{\pi_1}{\pi_0}.\end{aligned}$$

Therefore,

$$P_f = \left(\frac{\pi_1}{\pi_0}\right)^{1+\tau} = \left(\frac{\pi_1}{\pi_0}\right)\left(\frac{\pi_1}{\pi_0}\right)^{\frac{\rho_E}{1-2\rho_E}}.$$

Similarly, the probability of error of the FC given operation point $\tilde{B}$ is

$$P_e = P_f\pi_0 = \pi_1\left(\frac{\pi_0}{\pi_1}\right)^{-\frac{\rho_E}{1-2\rho_E}},$$

which is equivalent to the probability of error at $\tilde{A}$.

APPENDIX D

DECISION RULE THRESHOLD UNDER BAYESIAN FRAMEWORK

In Equation (3.20), we have $D(T,\bar{\alpha})$ which is an increasing function of $T \in (\alpha, 1)$ and $D(T,\bar{\beta})$ is a decreasing function of $T \in (0,\beta)$.Therefore, P_e is minimized when $D(T,\bar{\alpha}) = D(T,\bar{\beta})$. This leads to

$$\begin{aligned}
&D(T,\bar{\alpha}) - D(T,\bar{\beta}) = 0 \implies \\
&T \ln \frac{T}{\bar{\alpha}} + (1-T)\ln\frac{1-T}{1-\bar{\alpha}} - T\ln\frac{T}{\bar{\beta}} - (1-T)\ln\frac{1-T}{1-\bar{\beta}} = 0 \\
&\implies T = \frac{\ln\frac{1-\bar{\alpha}}{1-\bar{\beta}}}{\ln\frac{\bar{\beta}}{\bar{\alpha}} + \ln\frac{1-\bar{\alpha}}{1-\bar{\beta}}}, \ \ \forall(\alpha,\beta).
\end{aligned}$$

Alternatively, notice that

$$\begin{aligned}
T - \bar{\alpha} &= \frac{\bar{\alpha}\ln\frac{\bar{\alpha}}{\bar{\beta}} + (1-\bar{\alpha})\ln\frac{1-\bar{\alpha}}{1-\bar{\beta}}}{\ln\frac{\bar{\beta}}{\bar{\alpha}} + \ln\frac{1-\bar{\alpha}}{1-\bar{\beta}}} \\
&= \frac{D(\bar{\alpha},\bar{\beta})}{\ln\frac{\bar{\beta}}{\bar{\alpha}} + \ln\frac{1-\bar{\alpha}}{1-\bar{\beta}}},
\end{aligned}$$

and

$$\bar{\beta} - T = \frac{D(\bar{\beta},\bar{\alpha})}{\ln\frac{\bar{\beta}}{\bar{\alpha}} + \ln\frac{1-\bar{\alpha}}{1-\bar{\beta}}}.$$

Thus

$$\frac{\bar{\beta} - T}{T - \bar{\alpha}} = \frac{D(\bar{\beta},\bar{\alpha})}{D(\bar{\alpha},\bar{\beta})},$$

or

$$T = \frac{D(\bar{\alpha},\bar{\beta})\bar{\beta} + D(\bar{\beta},\bar{\alpha})\bar{\alpha}}{D(\bar{\beta},\bar{\alpha}) + D(\bar{\alpha},\bar{\beta})},$$

According to the approximation in Table 3.1, if $\bar{\rho} > 0$, when $(\alpha,\beta) \to (0,0)$ or $(\alpha,\beta) \to (1,1)$, $D(\bar{\alpha},\bar{\beta}) = D(\bar{\beta},\bar{\alpha})$, therefore

$$\begin{aligned} T &\approx \frac{\bar{\alpha}+\bar{\beta}}{2} \\ &= \bar{\rho} + \frac{(1-\bar{\rho})\,(\alpha+\beta)}{2}. \end{aligned}$$

When $\bar{\rho}=0$, $(\alpha,\beta)\to(0,0)$, $\bar{\beta}\approx\beta$ and $\bar{\alpha}=\alpha$, then

$$\begin{aligned} T &= \frac{D(\bar{\alpha},\bar{\beta})\beta + D(\bar{\beta},\bar{\alpha})\alpha}{D(\bar{\alpha},\bar{\beta}) + D(\bar{\beta},\bar{\alpha})}, \\ &\approx \frac{\beta^2 + \beta(\ln\frac{\beta}{\alpha}-1)\alpha}{\beta+\beta(\ln\frac{\beta}{\alpha}-1)}, \\ &= \frac{\beta + (\ln\frac{\beta}{\alpha}-1)\alpha}{\ln\frac{\beta}{\alpha}} \\ &= \frac{\beta-\alpha}{\ln\frac{\beta}{\alpha}} + \alpha. \end{aligned}$$

Similarly, we can calculate the corresponding threshold when $(\alpha,\beta)\to(1,1)$,

$$\begin{aligned} T &= \frac{D(\bar{\alpha},\bar{\beta})\beta + D(\bar{\beta},\bar{\alpha})\alpha}{D(\bar{\alpha},\bar{\beta}) + D(\bar{\beta},\bar{\alpha})}, \\ &\approx \frac{(1-\alpha)\left(\ln\frac{1-\alpha}{1-\beta}-1\right)\beta + (1-\alpha)\,\alpha}{(1-\alpha)\left(\ln\frac{1-\alpha}{1-\beta}-1\right) + (1-\alpha)}, \\ &= \beta + \frac{\alpha-\beta}{\ln\frac{1-\alpha}{1-\beta}}. \end{aligned}$$

We summarize the decision rule threshold T in Table 3.2.

APPENDIX E

MINIMIZED FISHER INFORMATION RATIO

From Equation (4.5), we can see that the minimized FI ratio is achieved when the CDF $Q\left(\eta-\theta\right)=\frac{1}{2}$, therefore, plug that into equation (4.5), we have

$$\begin{aligned}R&=\frac{(1-2\rho_F)^2(\rho_E+\big(1-2\rho_E)\frac{1}{2}\big)}{(1-2\rho_E)^2(\rho_F+(1-2\rho_F)\frac{1}{2})}\frac{(1-\rho_E-(1-2\rho_E)\frac{1}{2})}{(1-\rho_F-(1-2\rho_F)\frac{1}{2})}\\&=\frac{\left(1-2\rho_F\right)^2}{\left(1-2\rho_E\right)^2}\end{aligned}$$

This minimized ratio can also be used to guarantee the performance of the FC due to channel disparity.