December 2016

On the Design and Analysis of Secure Inference Networks

Venkata Sriram Siddhardh Nadendla
Syracuse University

ABSTRACT

Parallel-topology inference networks consist of spatially-distributed sensing agents that collect and transmit observations to a central node called the fusion center (FC), so that a global inference is made regarding the phenomenon-of-interest (PoI). In this dissertation, we address two types of statistical inference, namely binary-hypothesis testing and scalar-parameter estimation in parallel-topology inference networks. We address three different types of security threats in parallel-topology inference networks, namely Eavesdropping (Data-Confidentiality), Byzantine (Data-Integrity) or Jamming (Data-Availability) attacks. In an attempt to alleviate information leakage to the eavesdropper, we present optimal/near-optimal binary quantizers under two different frameworks, namely differential secrecy where the difference in performances between the FC and Eve is maximized, and constrained secrecy where FC's performance is maximized in the presence of tolerable secrecy constraints. We also propose near-optimal transmit-diversity mechanisms at the sensing agents in detection networks in the presence of tolerable secrecy constraints. In the context of distributed inference networks with M-ary quantized sensing data, we propose a novel Byzantine attack model and find optimal attack strategies that minimize KL Divergence at the FC in the presence of both ideal and non-ideal channels. Furthermore, we also propose a novel deviation-based reputation scheme to detect Byzantine nodes in a distributed inference network. Finally, we investigate optimal jamming attacks in detection networks where the jammer distributes its power across the sensing and the communication channels. We also model the interaction between the jammer and a centralized detection network as a complete-information zero-sum game. We find closed-form expressions for pure-strategy Nash equilibria and show that both the players converge to these equilibria in a repeated game. Finally, we show that the jammer finds no incentive to employ pure-strategy equilibria, and causes greater impact on the network performance by employing mixed strategies.

ON THE DESIGN AND ANALYSIS OF SECURE INFERENCE NETWORKS

By

Venkata Sriram Siddhardh Nadendla
B.E., Sri Chandrasekharendra Saraswathi Viswa Mahavidyalaya, 2007
M.S., Louisiana State University, 2009

DISSERTATION

Submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Electrical and Computer Engineering

Syracuse University
December 2016

ACKNOWLEDGMENTS

I owe my deepest gratitude to my adviser Prof. Pramod K. Varshney for his omnipresent support, and for teaching me to appreciate the elegance of signal processing and security beyond the walls of mathematical rigor. I thank all the committee members: Prof. Biao Chen, Prof. Makan Fardad, Prof. Mustafa Cenk Gursoy, Prof. Yingbin Liang and Prof. Lixin Shen for their valuable time and insightful remarks. I am very grateful to Prof. Yunghsiang S. Han and Prof. Vinod Sharma for their effective guidance and productive collaboration. I am greatly indebted to the current and former members of Sensor Fusion Lab who have been a great source of ingenious ideas throughout my stay. In particular, I have thoroughly enjoyed brainstorming sessions with Swastik Brahma, Arun Subramanian, Aditya Vempaty, Bhavya Kailkhura, Sijia Liu and Raghed El Bardan.

I express my sincere gratitude to Andrew L. Drozd for providing internships at ANDRO Computational Solutions and exposing me to various security perspectives in networked systems. I also thank all the other colleagues in the company, and in particular, Irina Kasperovich, Dr. Ashwin Amanna, Dr. Svetlana Foulke, Richard Lawrence, Jithin Jagannath and Travis Murphy, for valuable discussions and feedback.

It is my greatest honor to thank all my teachers, friends and family for igniting the spirit in me, and in particular, Monisha Lakshmipathi for her unconditional support. Most importantly, I owe a lot to my wife, Aishwarya, for her warm companionship and enduring effort in keeping me focused all the time. Above all, this

thesis wouldn't have been possible without the continuous inspiration from my parents, Venkata Chalapathi Rao Nadendla and Indira Nadendla, and my brother, Venkata Srikrishna Karthik Nadendla.

To my family:

past, present and future.

TABLE OF CONTENTS

LIST OF TABLES

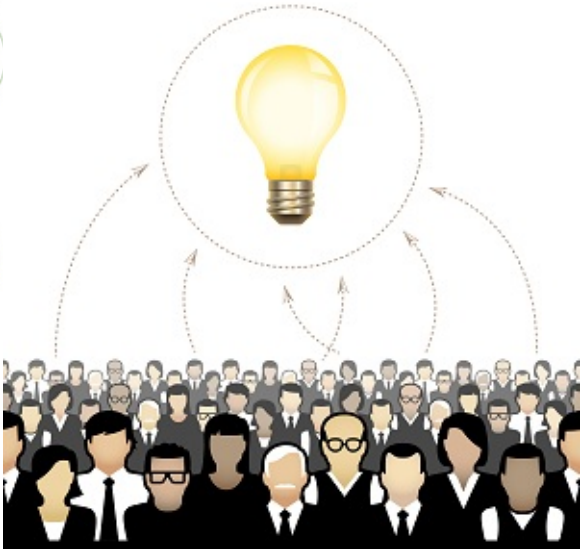
LIST OF FIGURES

CHAPTER 1
INTRODUCTION

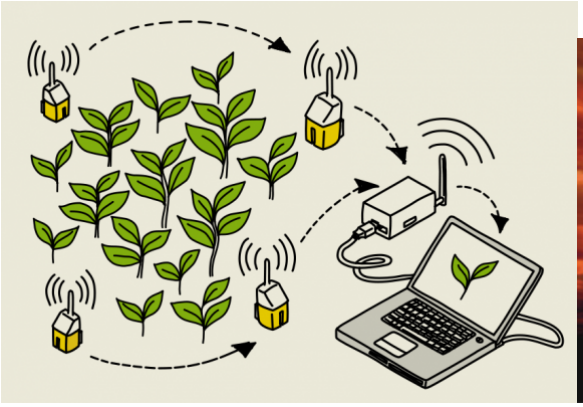
1.1 Inference Networks

Statistical inference has played a cardinal role in the growth of modern technology, and is quintessential in almost every application when there is uncertainty within the collected data. This demand for statistical inference has been bolstered by significant advancements in the design of sensors and their networks over the past decade. In the context of classical inference, a single powerful sensing agent is designed /chosen to collect data and make inferences about a phenomenon-of-interest (PoI). Such a sensor requires expensive technologies to facilitate high-performance inference-tasks. Furthermore, other practical difficulties such as PoI-shadowing and short battery life can severely degrade the inference performance. Therefore, inference networks have been proposed where several low-cost sensing agents are installed to collect data in a spatially distributed manner. Although these distributed sensing agents are inexpensive and have limited computational, bandwidth and energy resources, a significant number of such spatially-distributed low-cost agents collaborate and share resources and processing effort to achieve a prescribed performance. In order to make inference networks practically viable, several researchers had pursued extensive work in the design and analysis of these networks under different scenarios and

(a) E-Health Networks



(b) Crowd-Sensing



(c) Agricultural Sensor Networks

(d) Radar Networks

Figure 1.1: Inference Network Applications

applications (See [1, 9, 11, 12, 14, 15, 47, 60, 61, 64, 65, 69] and references therein). Today, inference networks span over a broad range of applications such as distributed radar surveillance in the military domain, traffic-control networks, agricultural sensor networks and disaster-monitoring in the commercial cyber-physical domain and, various other applications in stock markets, crowdsensing, smart-homes and wearable body sensors to facilitate e-health, as pointed out in Figure 1.1.

As shown in Figure 1.2, there are fundamentally two types of inference networks de-

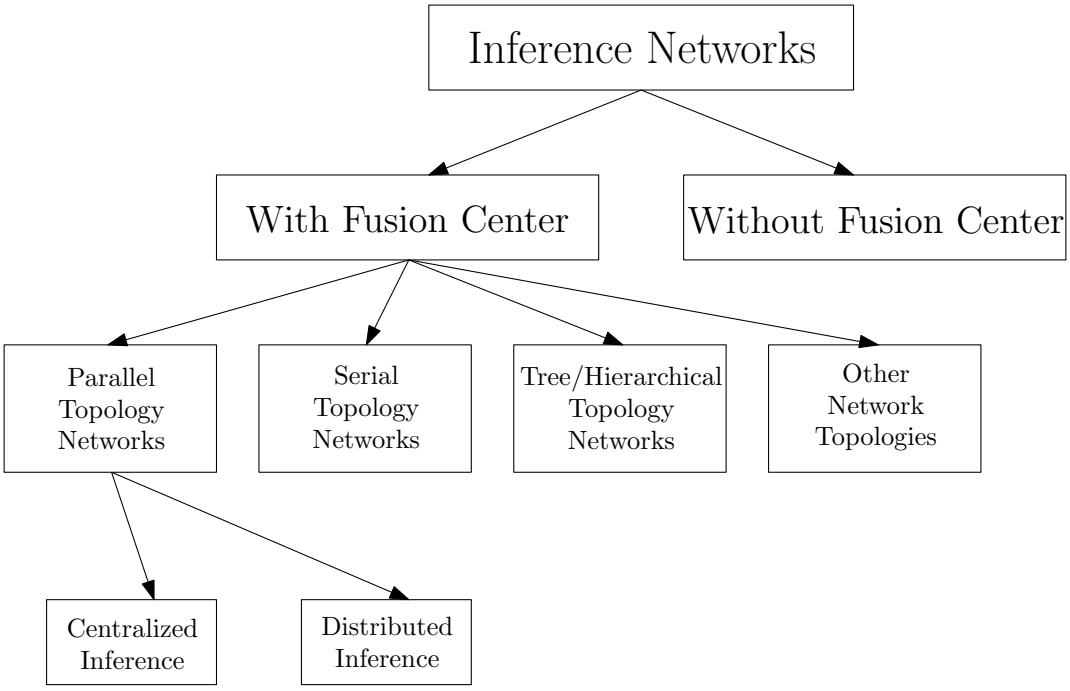


Figure 1.2: Topological Configurations of Inference Networks

pending on the presence/absence of a central node called the fusion center (FC), where a global inference is made regarding the PoI [69]. Furthermore, inference networks with the FC can be configured in different topologies based on their application needs. In *parallel-topology* inference networks, sensing agents collect and transmit either processed/unprocessed observations to the FC. In general, parallel-topology inference networks can be classified into two types, namely *centralized* and *distributed* inference networks, based on how the information is processed in the network. In a *centralized* inference network, the sensing agents transmit unprocessed observations to the FC using an amplify-and-forward strategy. The term *centralized* indicates that the data is only processed centrally at the FC. On the other hand, in a *distributed* inference network, sensing agents compress observations into a finite alphabet set and transmit compressed symbols to the FC. Due to the transmission of a finite alphabet set, the bandwidth requirement for the *distributed* inference network is significantly reduced at the expense of a minimal deterioration in inference performance.

Note that the parallel-topology inference network is an umbrella term used to label networks where sensors can communicate to the FC either over dedicated channels, multiple-access channels, or communication webs (internet) [65]. In contrast, *serial-topology* inference networks comprise of spatially distributed sensing agents that collaborate together in a linear hop-by-hop manner, in relaying their observations to the FC. Such networks are found in special applications such as vehicular networks where the sensors are aligned along the road. Other topological configurations that have been studied extensively, range from tree networks [58, 59] to collaborative inference networks [23, 32, 33] where sensors collaborate with each other in order to alleviate transmission costs. In some practical applications where the infrastructure cost is very high, inference networks are designed without an FC. In such cases, each sensing agent shares processed/unprocessed observations with the neighboring nodes in an ad-hoc manner, and makes a global inference based on the information shared over several iterations.

In this dissertation, we focus on parallel-topology inference networks that are designed to address two fundamental statistical inference problems, namely *binary hypothesis-testing* and *scalar parameter estimation*. In the case of *binary hypothesis-testing*, the goal is to detect the presence or absence of a given phenomenon-of-interest (PoI). On the other hand, in the case of *scalar parameter estimation*, the goal is to estimate a scalar parameter regarding the PoI. We analyze the vulnerabilities of a parallel-topology inference network, and design them under three different types of security threats, namely eavesdropping, Byzantine and jamming attacks. These attacks are discussed in greater detail in Section 1.2.

In the following subsection, we present our basic system model for the inference network, and introduce some notation to label the signals and decisions made at both the local sensing agents as well as the FC. Depending on the need, we may further introduce more notation in the future chapters.

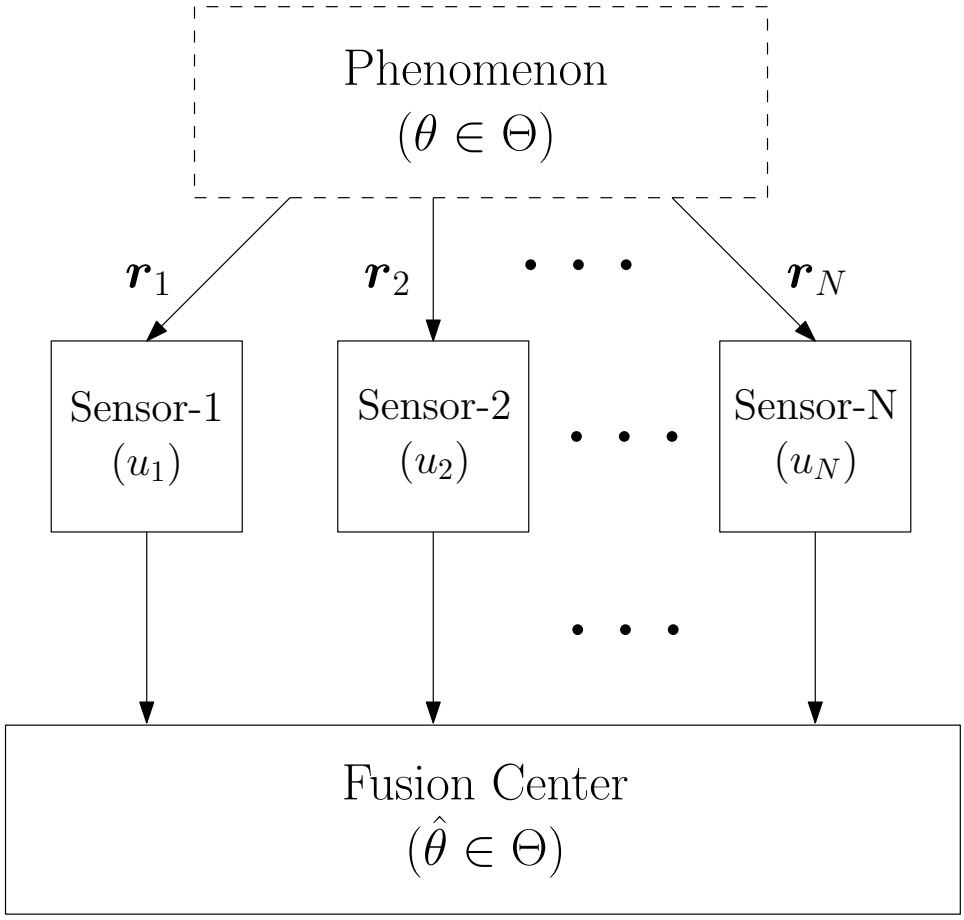


Figure 1.3: Parallel-Topology Inference Network Model

1.1.1 Notations and System Model

Consider a parallel-topology inference network with N sensing agents which sense a given PoI, as shown in Figure 1.3. Let $\theta \in \Theta$ denote the parameter representing the PoI's state, where Θ is the set of all possible states of the PoI. For example, in the case of *binary hypothesis-testing*, $\Theta \triangleq \{0, 1\}$. Similarly, in the case of *scalar parameter estimation*, $\Theta \triangleq \mathbb{R}$. Since there are two types of channels in any given inference network that a given sensor encounters, we resolve this confusion by labeling these two channels as follows. We refer to the channel between the PoI and the sensor as a *sensing channel*, and the channel between the sensor and the FC as a *communication channel*.

In this dissertation, we use regular and bold symbols to denote scalar and vector quantities respectively. For example, we denote the received signal (observation) at the i^{th} sensing agent as $\boldsymbol{r}_i$. In the case where the i^{th} sensing agent collects a scalar observation, we denote it as r_i. For the sake of generality, we will use the bold notation to denote various quantities, unless otherwise stated.

We assume the following signal received at the i^{th} sensing agent over its sensing channel:

$$\boldsymbol{r}_i = \boldsymbol{f}_i(\theta) + \boldsymbol{n}_i, \tag{1.1}$$

where $\boldsymbol{f}_i(\cdot)$ is a known, deterministic and invertible function for all $i = 1, \cdots, N$, and $\boldsymbol{n}_i$ is a zero-mean additive-white Gaussian noise with covariance matrix Σ_s. Having acquired the symbol $\boldsymbol{r}_i$, the i^{th} sensing agent processes (amplifies or compresses) it into a symbol u_i before relaying it to a central node called the fusion center (FC).

We consider two different types of communication channels in this dissertation. In the first type, we assume that the communication channel is discrete and memoryless, in which case, we denote the received symbol from the i^{th} sensing agent as v_i. In the second type of communication channels, we assume that the FC receives a real-valued signal which is denoted as r_{fc_j} at the j^{th} receiving antenna (or, channel use). Based on the received messages, FC makes a global inference $\hat{\theta}$ regarding the state of the PoI.

1.2 Security Threats in Inference Networks

Due to its wide range of applications and technological impact, vulnerabilities in the design of inference networks pose a very significant problem that ought to be addressed with great concern. Moreover, new security threats are discovered every day to bring down various networks, particularly in the context of cyber-physical systems. Although many traditional approaches have been proposed to address security, they also demand large amounts of resources such as computational power and latency. In addition, most of these approaches

Figure 1.4: Security Framework

rely on mathematical conjectures that assume the absence of any computationally tractable algorithm to crack them.

On the other hand, most inference networks consist of sensing agents with limited computational and bandwidth resources. Therefore, the design of secure inference networks demands novel techniques that protect the network using provably simple designs and computationally tractable algorithms. Initial attempts in this direction were made in the last decade by [24, 48, 49], which were mainly inspired by traditional security designs. As the security threats have evolved to be more powerful and directed specifically towards inference networks, there is an immediate need for system-level approaches to either prevent or mitigate these security threats from deteriorating the inference performance of the network. In the remaining section, we present a basic framework for security in order to broadly classify the attacker's endeavor, which has been discussed extensively in the traditional security literature.

Security threats have been traditionally classified into three types, based on the system aspects threatened by attackers: Confidentiality, Integrity and Availability (in short, CIA). In this dissertation, we follow the same taxonomy to discuss security threat models in

inference networks. Note that the above *CIA framework* for security can be configured to address different services in networking such as node identities, node location and data. In this dissertation, we investigate the *decision-theoretic* aspects of the *CIA* framework, where the attacker is assumed to interfere with various aspects of data flowing in the network. In the remaining section, we describe briefly how these security threats classified under the *decision-theoretic CIA framework* impact the inference performance.

- **Data Confidentiality:** Data confidentiality broadly addresses the problem of data exfiltration from the inference network to an unauthorized, third-party entity. There are several approaches that the attacker can employ, in order to extract any given information from the inference network. All of these approaches can be broadly classified into two types. In the first type, the attacker gains access to the data repositories of a given entity within the inference network, and extracts private information. This attack model is traditionally labeled as a threat on the *privacy* of data, and various solutions have been proposed to stall any unauthorized data access in the inference network. In the second approach, the attacker extracts useful information by eavesdropping the transmitted information via wiretapping the communication channels between the sensing agents and the FC. This attack model is labeled as a threat to the *secrecy* of data in an inference network, and the readers can refer to [22] for a survey on state-of-the-art solutions to this security threat.

- **Data Integrity:** Data integrity in inference networks refers to the authenticity of data in terms of its accuracy in value. Such discrepancies in data can be achieved by the attacker by either creating false identities through network infiltration, or by compromising and enslaving an existing sensing agent within the inference network. In the context of data accuracy, the attacker injects falsified sensing data into the network either to randomize the global inference, or to manipulate it in a specific manner. For more details on how data-falsification attacks with the intent of randomizing the global inference (labeled as *Byzantine* attacks) are mitigated, the readers may refer to

some of the solution approaches in [68]. On the other hand, one may find instances of data falsification attacks with the intent of manipulating the global inference in a specific manner in the case of spectrum sensing in cognitive-radio (CR) networks where CR agents compete for vacant spectrum.

- **Data Availability:** Data availability in inference networks points out to security attacks where data is made unavailable at any entity in the inference network. In the case of attacks which introduce temporal inconsistencies in data acquisition or transmission, there is a latency in the inference mechanism, thus resulting in untimely decisions. Another attack that falls into this category is the case where the attacker scrambles the identities of the sensing agents, in which case, the inference performance at the FC deteriorates significantly. Among all the attacks within the context of data availability, jamming attacks are extensively studied by several researchers. In this attack model, the jammer introduces disruptive interference in either the sensing channel, the communication channels, or both, in order to deteriorate the inference performance of the network.

In the real world, more complex threats can be found where the attackers adopt hybrid models to maximally disrupt the operation of the inference network. For example, a smart jammer may first eavesdrop the channel and use this information to optimize its jamming strategy in order to cause maximal impact on the performance of the inference network. This dissertation attempts to understand the design limitations in the context of three simple attack models, and provides a basic foundation for the future design and analysis of inference networks in the presence of complex, real-world attackers. In the remaining part of this chapter, we summarize the contributions of this dissertation.

1.3 Major Contributions

As pointed out in the earlier section, inference networks suffer from a wide variety of security threats. Since these security threats are driven by various motives, there is a demand for a unique design for securing inference networks from each of these security threats. Therefore, we design secure inference networks in the presence of three attack models, namely *eavesdropping*, *Byzantine* and *jamming* attacks, each of which is derived from one of the three issues within the *CIA* framework. We throw light on fundamental design-limits and simple mitigation techniques for inference networks under each of these attack models. In the remainder of this section, we summarize the contributions of this dissertation in the aforementioned attack scenarios.

Eavesdropping Attack

In this dissertation, we address *data confidentiality* in inference networks within the context of *secrecy* threats, by considering the problem of an eavesdropping attack in detection (binary hypothesis testing) networks. In the past, a few attempts have been made to address the problem of eavesdropping threats by designing stochastic ciphers at the sensing agents [3, 20, 41]. In contrast, Marano *et al.* had designed optimal decision rules for a censoring sensor network in the presence of eavesdroppers in [36], where they had assumed that Eve can only determine whether an individual sensor transmits its decision or not.

In this dissertation, we consider a more realistic scenario where Eve can extract more information than just merely determining the presence or absence of transmission, and hence can make a reasonably good decision regarding the PoI, based on its receptions. In the case of distributed detection networks [J1, C1], we consider the problem of designing binary quantizers at the local sensing agents in the presence of a tolerable constraint on eavesdropper's inference performance. In Chapter 2, we design optimal binary quantizers in a detection network where all the sensing agents are conditionally independent

and identically distributed (i.i.d.) under the true hypothesis. Furthermore, due to the intractable nature of the design of a general detection network with non-identical sensing agents, we propose an efficient design for the local quantizers using a greedy algorithm. On the other hand, in the case of centralized detection networks [C2], we propose a near-optimal transmit diversity mechanism at the sensing agents in Chapter 3, in order to maximize the detection performance at the FC while constraining the detection performance at the eavesdropper.

Byzantine Attack

In our attempt to address the issue of *data integrity*, we consider the problem of Byzantine attacks in distributed inference networks. In the past, several efforts have been made to address the problem of Byzantine attacks in distributed inference networks with different topologies, especially when the sensors employ binary quantization to compress their observations. Byzantine attacks on centralized inference networks have been addressed in the context of smart grids in [26]. For a detailed survey on related works on inference networks in the presence of Byzantine attacks, the reader may refer to the survey by Vempaty *et al.*, in [68].

In this dissertation, we investigate the problem of distributed inference with M-ary quantized data at the sensors in the presence of Byzantine attacks in Chapter 4. We propose a general attack model where the Byzantine nodes modify the original quantized message into another symbol within the quantization alphabet-set using a probability distribution. In the presence of noiseless communication channels, we show that the optimal Byzantine attack deteriorates drastically in terms of the *blinding* fraction of Byzantine nodes, as the quantization alphabet size increases. We also deduce the optimal Byzantine attack in the presence of discrete memoryless channels between the sensors and the FC. In the case where the Byzantine attack cannot afford to launch an optimal attack, we find an attack from a restricted space of highly-symmetric attack strategies, that maximally degrades

the performance of the inference network in the presence of resource-constrained Byzantine attacks. Furthermore, a reputation-based scheme for identifying Byzantine nodes is also presented as the network's strategy to mitigate the impact of Byzantine threats on the inference performance. We also provide asymptotic analysis to find the optimal reputation-based scheme as a function of the fraction of compromised nodes in the network.

Jamming Attack

Within the context of data-availability, we consider the problem of jamming attacks on inference networks which intentionally disrupt both the sensing and the communication channels simultaneously by introducing interference. In the past, several attempts have been made to address and mitigate jamming threats in inference networks. For more details, the reader may refer to [39, 75] and references therein. In this dissertation, we consider a novel decision-theoretic approach within the context of detection networks where the jammer optimizes its attack strategy so as to minimize the detection performance of the network.

In Chapter 5, we consider the problem of finding the optimal jamming attack in a simple detection network where there is only one sensing agent, for the sake of illustration. Furthermore, we also assume that all the entities (PoI, sensing agent, FC and the jammer) lie on a straight line. The goal of the jammer is to distribute its power between the sensing and the communication channels in such a way that the error probability at the FC is maximized. Since the problem is non-convex and intractable, we investigate the optimal solution numerically and illustrate the results for different example scenarios. On the other hand, in Chapter 6, we model a zero-sum game between a centralized detection network and the jammer, and investigate Nash equilibria. In particular, we find closed-form expressions for a family of pure-strategy Nash equilibria, and also show that the jammer has no incentive to employ any of these pure-strategy equilibria in a complete-information game under strict power constraints.

1.4 Bibliographic Notes

Most of the work presented in this dissertation has been either already published, or in some stage of publication. Table 1.1 shows the relationship between the different chapters, and the following list of publications.

Relevant Publications

Journal Papers and Magazine Articles

J1. **V. S. S. Nadendla** and P. K. Varshney, "Design of Binary Quantizers for Distributed Detection Under Secrecy Constraints," *IEEE Transactions on Signal Processing*, vol. 64, no. 10, pp. 2636-2648, May 15, 2016.

J2. **V. S. S. Nadendla**, Y. S. Han and P. K. Varshney, "Distributed Inference With M-Ary Quantized Data in the Presence of Byzantine Attacks," *IEEE Transactions on Signal Processing*, vol. 62, no. 10, pp. 2681-2695, May 15, 2014.

J3. **V. S. S. Nadendla**, V. Sharma, P. K. Varshney, "On Strategic Multi-Antenna Jamming in Centralized Detection Networks," Submitted to *IEEE Signal Processing Letters*, Availabile on ArXiv:1608.04705, Aug 2016.

Refereed Conferences

C1. **V. S. S. Nadendla**, H. Chen and P. K. Varshney, "Secure distributed detection in the presence of eavesdroppers," *Proceedings of the 44th Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, 2010.

C2. **V. S. S. Nadendla**, S. Liu and P. K. Varshney, "Design of transmit-diversity schemes in detection networks under secrecy constraints," *Proceedings of the 53rd Annual Allerton Conference on Communication, Control, and Computing*, Monticello, IL, 2015.

C3. **V. S. S. Nadendla**, H. Chen, and P. K. Varshney, "On jamming models against collaborative spectrum sensing in a simple cognitive radio network," *Proceedings of the 44th Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, Nov 2010.

C4. **V. S. S. Nadendla**, H. Chen, and P. Varshney, "Minimax games for cooperative spectrum sensing in a centralized cognitive radio network in the presence of interferers," *Proceedings of the 30th Military Communications Conference*, Baltimore, MD, Nov 2011.

Other Related Publications

Journal Papers and Magazine Articles

J4. B. Kailkhura, **V. S. S. Nadendla** and P. K. Varshney, "Distributed inference in the presence of eavesdroppers: a survey," *IEEE Communications Magazine*, vol. 53, no. 6, pp. 40-46, June 2015.

Refereed Conferences

C5. M. Gagrani, P. Sharma, S. Iyengar, **V. S. S. Nadendla**, A. Vempaty, H. Chen, P. K. Varshney, "On noise-enhanced distributed inference in the presence of Byzantines," *Proceedings of the 49th Annual Allerton Conference on Communication, Control, and Computing*, pp. 1222-1229, Monticello, IL, 2011.

C6. A. Vempaty, **V. S. S. Nadendla** and P. K. Varshney, "Further results on noise-enhanced distributed inference in the presence of Byzantines," *Proceedings of the 16th International Symposium on Wireless Personal Multimedia Communications*, Atlantic City, NJ, 2013.

Chapter	Security Threat	Contribution	Related Publications
2	*Data Confidentiality* (Eavesdropping Attack)	Design of Sensor Quantizers	[J1], [C1]
3	*Data Confidentiality* (Eavesdropping Attack)	Transmit Diversity Schemes	[C2]
4	*Data Integrity* (Byzantine Attack)	Optimal Attack for M-ary Quantized Data, Reputation-based identification	[J2]
5	*Data Availability* (Jamming Attack)	Optimal Jamming Attack: Power Allocation and Placement	[C3]
6	*Data Availability* (Jamming Attack)	Strategic Games	[J3],[C4]

Table 1.1: Chapter-wise Contributions and Related Publications

CHAPTER 2
SECRECY IN DISTRIBUTED DETECTION: DESIGN OF BINARY QUANTIZERS

Secrecy in the context of distributed detection networks is an important problem, especially when the network is a sub-system within a larger cyber-physical system. Following are some examples where confidentiality plays a very important role in the context of distributed inference. First, consider the example of a distributed radar network where the radars observe the presence or absence of an enemy aircraft. Any information about the radar decisions at the enemy aircraft can help it to adapt its strategy so as to remain invisible to the radar and in clandestine pursuit of its mission. Another example is the case of a cognitive-radio (CR) network where an eavesdropper may be able to use a given vacant PU channel, without paying any participation costs to the network moderator. Thus, selfishness and maliciousness can be two motives of any eavesdropper to compromise the confidentiality of any inference network. In this chapter, we address confidentiality in distributed detection networks and focus on the design of the network such that the eavesdropper may not acquire any information beyond tolerable limits.

In this chapter, we consider a distributed detection network in the presence of noisy channels between the sensors and the FC, as well as those between the sensors and the

Eve, whose transition probabilities are known to the network designer. We address the notion of secrecy in two different frameworks, namely *differential* secrecy and *constrained secrecy*. In Section 2.4, we address the framework of differential secrecy by designing optimal binary quantizers at the sensors that maximize the difference between the KL Divergences at FC and Eve. We show that the structure of the optimal sensor quantizers are either likelihood ratio test (LRT) based, or uninformative depending on the quality of the Eve's channels. On the other hand, in Section 2.5, we address the framework of constrained secrecy where we design optimal binary sensor quantizers that maximize KL Divergence at the FC while constraining the Eve's KL Divergence to a prescribed tolerance level. We consider two scenarios, one where the channels between the sensors and the FC (likewise, channels between sensors and the Eve) are modeled as identical binary symmetric channels (BSCs), and the second where the channels are modeled as non-identical BSCs. In Section 2.5.1, we consider the identical channel scenario, where we show that the structure of the optimal quantizer at the local sensors is a *likelihood ratio test* (LRT). We present an illustrative example where we assume that the sensors make noisy observations of a known deterministic signal. We present an algorithm to find the optimal threshold so as to maximize the KL Divergence at the FC while ensuring that the Eve's KL Divergence remains within tolerable limits. In Section 2.5.2, we consider the scenario where channels are non-identical, where we decompose the problem into N subproblems to be solved sequentially using dynamic programming. Consequently, we decouple the Eve's constraint into N individual constraints, thus allowing us to solve each of these decoupled problems as in the identical sensor case.

2.1 Literature Survey

In the past, a few attempts have been made to address the problem of eavesdropping threats by designing ciphers in the broader context of sensor networks. For example, Aysal *et al.*

in [3] investigated the problem of secure distributed estimation by incorporating a stochastic cipher in the existing sensor networks to improve secrecy. They showed a significant deterioration in Eve's performance (in terms of bias and mean squared error) at the cost of a marginal increase in the estimation variance at the FC. A similar attempt has been made in the context of distributed detection in sensor networks by Nadendla in [41], where the author presented an optimal network (sensor quantizers, flipping probabilities in the stochastic cipher and the fusion rule) that minimizes the error probability at the FC in the presence of a constraint on Eve's error probability. In [20], Jeon *et al.* proposed a cooperative transmission scheme for a sensor network where the sensors are partitioned into non-flipping, flipping and dormant sets, based on the thresholds dictated by the FC. The non-flipping set of sensors quantize the sensed data and transmit them to the FC, while the flipping sensors transmit flipped decisions in order to confuse the Eve. The sensors within the dormant set sleep, in order to conserve energy and we have an energy-efficient sensor network with longer lifetime.

In all of the above attempts, security in distributed detection systems was incorporated as an afterthought in that separate security blocks were added after the original system had been designed without considering the possible security threats. Marano *et al.* in [36], on the other hand, investigated the problem of designing optimal decision rules for a censoring sensor network in the presence of eavesdroppers. Although their framework of censoring sensor networks is more general, they assume that the Eve can only determine whether an individual sensor transmits its decision or not. In reality, Eve can extract more information than just merely determining the presence or absence of transmission, and hence can make a reasonably good decision regarding the PoI, based on its receptions. Therefore, in this chapter, we assume that the Eve is also interested in making similar inferences regarding the PoI, just as in the case of FC.

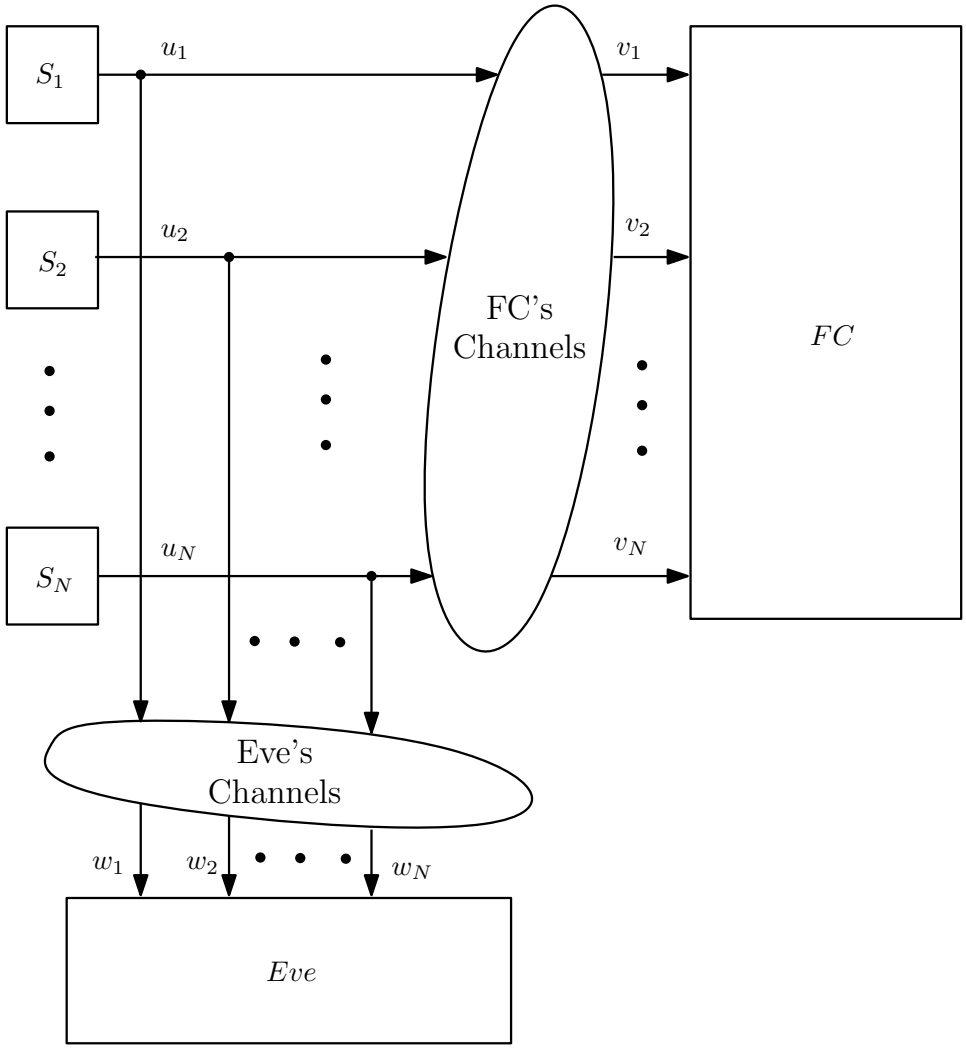


Figure 2.1: Distributed Inference Network in the Presence of an Eavesdropper

2.2 System Model

Consider a binary-hypothesis testing problem for distributed detection with N sensors under the Neyman Pearson framework, as shown in Figure 2.1. Let $\boldsymbol{r}_i = \{r_{i,t} : t = 1, \cdots, T\}$ denote a sequence of i.i.d. observations (in time) acquired by the i^{th} sensor over T time periods. Furthermore, we also assume that these observations $\boldsymbol{r}_i$ are independent across sensors, i.e., for $i = 1, \cdots, N$, but do not necessarily have identical distributions at different sensors. Let H_0 and H_1 denote the null and the alternate hypotheses respectively. We denote the conditional probability density functions of $r_{i,t}$ under hypotheses H_0 and H_1 as

$p_{i,0}(r) = p(r_{i,t} = r|H_0)$ and $p_{i,1}(r) = p(r_{i,t} = r|H_1)$ respectively. In this chapter, for all $i = 1, \cdots, N$, we assume that the i^{th} sensor employs binary quantization to compress its observation $r_{i,t}$ into $u_{i,t}$, as defined below, using a decision rule $\gamma_i(\cdot)$.

$$u_{i,t} = \gamma_i(r_{i,t}) = \begin{cases} 1, & \text{where } \Lambda(r_{i,t}) \geq \lambda_i \\ 0, & \text{otherwise.} \end{cases} \tag{2.1}$$

where $\Lambda(r_{i,t})$ is a test-statistic and λ_i is a suitable threshold to be designed.

Let $x_i = P(u_{i,t} = 1|H_0)$ and $y_i = P(u_{i,t} = 1|H_1)$ denote the false-alarm and detection probabilities at the i^{th} sensor respectively. The pair (x_i, y_i) is traditionally referred to, as the operating point of the i^{th} sensor, which can lie anywhere on the compact[1] unit-square $\mathscr{U} = [0, 1]^2$, which we call the *ROC space*. For any fixed test-statistic $\Lambda(\cdot)$, when the threshold λ_i is varied, the operating point of the i^{th} sensor follows a curve $y_i = g_\Lambda(x_i)$. This curve $y_i = g_\Lambda(x_i)$ is traditionally known as the *ROC curve*. In the rest of the chapter, we use the operating point (x_i, y_i) to represent the quantizer rule γ_i employed at the i^{th} sensor. Two quantizers γ_1 and γ_2 are considered identical (equivalent), if their operating points (x_1, y_1) and (x_2, y_2) are the same.

Let Γ_i denote the set of all feasible[2] operating points (x_i, y_i) at the i^{th} sensor. Then, the region Γ_i in the ROC space is upper-bounded by the set of operating points corresponding to the likelihood ratio tests (LRTs). We call this boundary as the *LRT curve*, and denote it as $y_i = g_{LRT_i}(x_i)$. Furthermore, we restrict our analysis only to those operating points that lie above the line $y_i = x_i$ in the ROC plane. This is because any point below the line $y_i = x_i$ contributes negatively to the overall performance in terms of error probability at the FC. In summary, the region Γ_i in the ROC space is upper-bounded by the LRT curve $y_i = g_{LRT_i}(x_i)$, and lower-bounded by the line $y_i = x_i$.

[1] In this context, compactness of the unit-square corresponds to the inclusion of the boundary points (0,0), (0,1), (1,0) and (1,1) within the set itself.

[2] The feasibility of an operating point is primarily dictated by the quality of the sensing observations. Note that the size of Γ_i diminishes as the sensor observations get corrupted due to multipath fading and/or thermal noise.

Note that the operating points $(0,0)$ and $(1,1)$ are the extreme points on the lower boundary $y_i = x_i$. These operating points $(0,0)$ and $(1,1)$ are achieved by always deciding $u_{i,t} = 0$ and $u_{i,t} = 1$ respectively. Since any operating point that lie on the line $y_i = x_i$ can be achieved via randomizing between the two operating points $(0,0)$ and $(1,1)$ (or, equivalently the corresponding quantizer rules), we refer to this boundary as the *uninformative* boundary (and, the corresponding quantizer rules and operating points as *uninformative* rules and points respectively), since these rules do not depend on sensor observations.

Given the operating point (x_i, y_i), the Kullback-Leibler (KL) Divergence of the i^{th} sensor is defined as follows.

$$D_i = x_i \log \frac{x_i}{y_i} + (1 - x_i) \log \frac{1 - x_i}{1 - y_i} \tag{2.2}$$

Let $\Upsilon = \{0,1\}^N$ denote the N-dimensional space of compressed symbols $\boldsymbol{u}_t = \{u_{1,t}, \cdots, u_{N,t}\}$ at all the sensors at a given time t. In this chapter, we assume that the i^{th} sensor transmits its compressed symbols $u_{i,t}$ to the FC through a binary-symmetric channel (BSC) with transition probability ρ_{fc_i}. In our model, we also assume that an eavesdropper wiretaps each of these sensor transmissions through a BSC with transition probability ρ_{e_i}.

If $\boldsymbol{v}_i = \{v_{1,t}, \cdots, v_{N,t}\}$ and $\boldsymbol{w}_i = \{w_{1,t}, \cdots, w_{N,t}\}$ denote the received symbols at the FC and Eve respectively, the operating point (x_i, y_i) at the i^{th} sensor gets transformed into (x_{fc_i}, y_{fc_i}) and (x_{e_i}, y_{e_i}) at the FC and Eve respectively, which are given as follows.

$$x_{fc_i} = P(v_{i,t} = 1|H_0) = \rho_{fc_i} + (1 - 2\rho_{fc_i})x_i \tag{2.3a}$$

$$y_{fc_i} = P(v_{i,t} = 1|H_1) = \rho_{fc_i} + (1 - 2\rho_{fc_i})y_i \tag{2.3b}$$

$$x_{e_i} = P(w_{i,t} = 1|H_0) = \rho_{e_i} + (1 - 2\rho_{e_i})x_i \tag{2.3c}$$

$$y_{e_i} = P(w_{i,t} = 1|H_1) = \rho_{e_i} + (1 - 2\rho_{e_i})y_i \tag{2.3d}$$

Let the contributions of the i^{th} sensor to the overall KL Divergence at the FC and Eve be denoted as D_{FC_i} and D_{E_i} respectively. Then, D_{FC_i} and D_{E_i} are defined as follows.

$$\begin{aligned} D_{FC_i} &= x_{fc_i} \log\left(\frac{x_{fc_i}}{y_{fc_i}}\right) + (1 - x_{fc_i}) \log\left(\frac{1 - x_{fc_i}}{1 - y_{fc_i}}\right) \\ D_{E_i} &= x_{e_i} \log\left(\frac{x_{e_i}}{y_{e_i}}\right) + (1 - x_{e_i}) \log\left(\frac{1 - x_{e_i}}{1 - y_{e_i}}\right). \end{aligned} \tag{2.4}$$

Let $\mathscr{A}_T^{FC}, \mathscr{A}_T^{E} \in \Upsilon^T$ denote the acceptance regions of the hypothesis H_1 at FC and Eve respectively, over a time-window $t = 1, \cdots, T$. Then, the global probabilities of false alarm and miss at the FC and Eve are given by

$$\begin{aligned} p_T^{FC} = Pr(\boldsymbol{v_i} \in \mathscr{A}_T^{FC} | H_0), \quad q_T^{FC} = Pr(\boldsymbol{v_i} \in \overline{\mathscr{A}}_T^{FC} | H_1). \\ p_T^{E} = Pr(\boldsymbol{w_i} \in \mathscr{A}_T^{E} | H_0), \qquad q_T^{E} = Pr(\boldsymbol{w_i} \in \overline{\mathscr{A}}_T^{E} | H_1). \end{aligned} \tag{2.5}$$

where $\overline{\mathscr{A}}_T^{FC}$ and $\overline{\mathscr{A}}_T^{E}$ are the rejection regions of the hypothesis H_1 at the FC and Eve respectively, and, $\boldsymbol{v_i} = \{v_{i,1}, \cdots, v_{i,T}\}$ and $\boldsymbol{w_i} = \{w_{i,1}, \cdots, w_{i,T}\}$ are the received symbols at the FC and Eve respectively, transmitted by the i^{th} sensor over a time window of length T. Next, we present Stein's Lemma that addresses the asymptotic properties of the global probability of miss q_T^{FC}.

Lemma 2.1 (*Stein's Lemma* [16])**.** *For any* $0 < \delta, \varphi < \frac{1}{2}$*, let* $q_{T,\delta}^{FC} = \min\limits_{p_T^{FC} < \delta} q_T^{FC}$ *and* $q_{T,\varphi}^{E} = \min\limits_{p_T^{E} < \varphi} q_T^{E}$*. Then, we have*

$$\begin{aligned} \lim_{\delta \to 0} \lim_{T \to \infty} -\frac{1}{T} \log q_{T,\delta}^{FC} &= \mathscr{D}_{FC} \\ \lim_{\varphi \to 0} \lim_{T \to \infty} -\frac{1}{T} \log q_{T,\varphi}^{E} &= \mathscr{D}_{E} \end{aligned} \tag{2.6}$$

where $\mathscr{D}_{FC}$ *and* $\mathscr{D}_{E}$ *are the KL divergences at the FC and Eve respectively, which are*

defined as follows.

$$\mathscr{D}_{FC} = \sum_{i=1}^{N} D_{FC_i} \quad \textit{and} \quad \mathscr{D}_{E} = \sum_{i=1}^{N} D_{E_i}. \tag{2.7}$$

Thus, KL Divergence is the error exponent for the global probability of miss when the global probability of false alarm is constrained (and diminishing to zero with time). Therefore, as a surrogate to the global probability of miss, we choose KL Divergence as the performance metric in this chapter. Note that $\mathscr{D}_{FC}$ and $\mathscr{D}_{E}$ are both convex functions of $\boldsymbol{x} = \{x_1, \cdots, x_N\}$ and $\boldsymbol{y} = \{y_1, \cdots, y_N\}$ in the hyper-cube $[0,1]^N$, which is made up of the ROC spaces of all the sensors in the detection network.

2.3 Linear Transformations in ROC Space

In this section, we focus our attention on the transformation of the operating point of a single sensor due to the presence of a binary symmetric channel (BSC) between a given sensor and both the FC, as well as between the same sensor and Eve. Let the operating point of a given quantizer be $A = (x, y)$. As mentioned earlier, the sensor's quantizer characteristics (x, y) are represented using its operating point in the sensor's ROC. Also, consider two BSCs with transition probabilities ρ_1 and ρ_2, each of which transforms the operating point $A = (x, y)$ into $B_1 = (x_1, y_1)$ and $B_2 = (x_2, y_2)$. Let $C = \left(\frac{1}{2}, \frac{1}{2}\right)$. In the following lemma, we present a useful relationship between A, B_1, B_2 and C.

Lemma 2.2. *Let $0 \leq \rho_1 \leq \rho_2 \leq \frac{1}{2}$. Then, B_1 and B_2 always lie on the line segment joining A and C. In addition, the following inequality holds true.*

$$\frac{x}{y} \leq \frac{x_1}{y_1} \leq \frac{x_2}{y_2} \leq 1 \leq \frac{1-x_2}{1-y_2} \leq \frac{1-x_1}{1-y_1} \leq \frac{1-x}{1-y} \tag{2.8}$$

Proof. Consider a BSC with transition probability ρ, which transforms the operating point

$A = (x, y)$ into $B = (\hat{x}, \hat{y})$. Then, the equation of the line joining A and B is given by

$$\frac{b-y}{a-x} = \frac{b-\hat{y}}{a-\hat{x}} \tag{2.9}$$

where (a, b) is some arbitrary point on the line.

Substituting $\hat{x} = \rho + (1-2\rho)x$ and $\hat{y} = \rho + (1-2\rho)y$, we have

$$\frac{b-y}{a-x} = \frac{b-\rho-(1-2\rho)y}{a-\rho-(1-2\rho)x}. \tag{2.10}$$

Rearranging the terms in Equation (2.10), we have

$$(b-y)[a-\rho-(1-2\rho)x] = (a-x)[b-\rho-(1-2\rho)y]. \tag{2.11}$$

Simplifying Equation (2.11), we have

$$(a-b) + (y-x) = 2(ay-bx). \tag{2.12}$$

Note that the line $a = b$ represents the set of operating points for which the KL Divergence becomes zero. Therefore, let us investigate the point where Equation (2.9) intersects the line $a = b$. Substituting $b = a$, we have

$$(2a-1)(y-x) = 0.$$

In other words, the line in Equation (2.9) intersects line $a = b = \frac{1}{2}$ for any transition probability ρ. In other words, the points A, B_1, B_2 and C are collinear.

In fact, as $\rho \to \frac{1}{2}$, $B \to C$. In other words, for a given sensor's operating point A, the transformed operating point B slides along the line segment joining A and C. This sliding behavior can be investigated by analyzing the distance between B and C, in terms of increasing ρ, as shown in Figure 2.2. We denote the Euclidian distance between B and

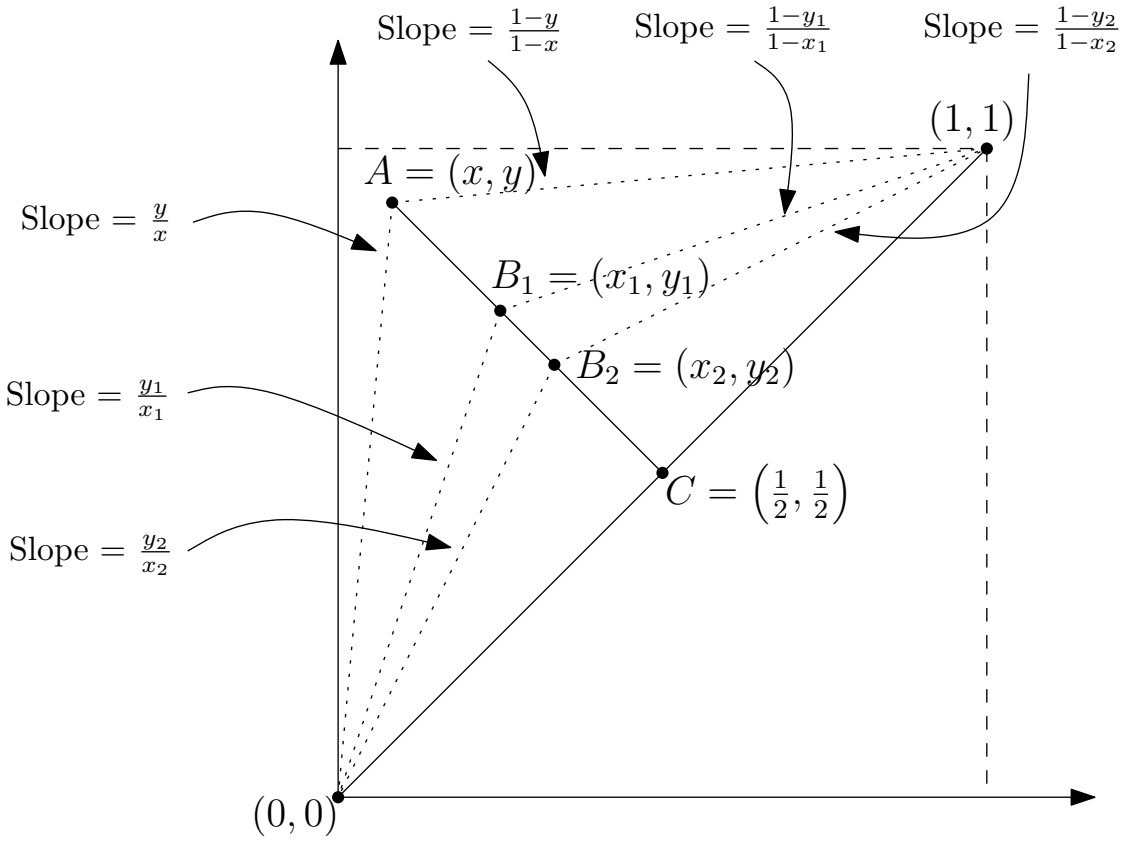


Figure 2.2: Transformations in the ROC

C as $\phi_{BC} = \sqrt{\left(\hat{x} - \frac{1}{2}\right)^2 + \left(\hat{y} - \frac{1}{2}\right)^2}$. Differentiating ϕ_{BC} with respect to ρ, we have

$$\begin{aligned}
\frac{d\phi_{BC}}{d\rho} &= \frac{1}{\phi_{BC}}\left[\left(\hat{x} - \frac{1}{2}\right)(1-2x) + \left(\hat{y} - \frac{1}{2}\right)(1-2y)\right] \\
&= \frac{-1 + \rho + (1-2\rho)[x(1-x) + y(1-y)]}{\phi_{BC}} \\
&= -\left(\frac{\rho + (1-2\rho)[1 - x(1-x) - y(1-y)]}{\phi_{BC}}\right) \\
&\leq 0,
\end{aligned} \tag{2.13}$$

since the function $x(1-x)+y(1-y)$ is concave and attains a maximum value of $\frac{1}{2}$ at $\left(\frac{1}{2}, \frac{1}{2}\right)$. In other words, B slides towards C as ρ increases. Consequently, as shown in Figure 2.2,

B_1 is farther away from C than B_2 on the line joining A and C, since $0 \leq \rho_1 \leq \rho_2 \leq 1$.

Note that the slope of the line joining $(0,0)$ and B_1 is $\frac{y_1}{x_1}$, and similarly, $\frac{y_2}{x_2}$ in the case of B_2. Since B_2 is closer to B_1 to C, as shown in Figure 2.2, $\frac{y_1}{x_1} \geq \frac{y_2}{x_2}$ and the slope tends to 1 as the transition probability approaches $\frac{1}{2}$. A similar argument holds for the slope of the lines that join B_1 and B_2 with $(1,1)$. Therefore, the inequality given in Equation (2.8) holds. □

In order to understand the impact of this transformation on the performance of the network, let us now analyze the KL Divergence at some arbitrary operating point $B = (\hat{x}, \hat{y})$ due to a BSC with transition probability ρ operating on the sensor operating point A. In the following lemma, we show that the KL Divergence decreases with increasing ρ.

Lemma 2.3. *Given the sensor operating point $A = (x, y)$, let $B = (\hat{x}, \hat{y})$ denote the transformed operating point due to a BSC with transition probability ρ. Let D_B denote the KL Divergence at B. Then, for $0 \leq \rho \leq \frac{1}{2}$, D_B is a monotonically decreasing function of ρ whenever $y \geq x$.*

Proof. The KL Divergence at the transformed operating point B is defined as follows.

$$D_B = \hat{x} \log \frac{\hat{x}}{\hat{y}} + (1 - \hat{x}) \log \frac{1 - \hat{x}}{1 - \hat{y}}. \tag{2.14}$$

Differentiating D_B with respect to ρ, we have

$$\begin{aligned} \frac{dD_B}{d\rho} &= (1-2y)\left[\frac{1-\hat{x}}{1-\hat{y}} - \frac{\hat{x}}{\hat{y}}\right] - (1-2x)\left[\log\left(\frac{1-\hat{x}}{1-\hat{y}}\right) - \log\left(\frac{\hat{x}}{\hat{y}}\right)\right] \\ &= \left(\frac{1-\hat{x}}{1-\hat{y}} - \frac{\hat{x}}{\hat{y}}\right)\left[(1-2y) - (1-2x)\left\{\frac{\log\left(\frac{1-\hat{x}}{1-\hat{y}}\right) - \log\left(\frac{\hat{x}}{\hat{y}}\right)}{\frac{1-\hat{x}}{1-\hat{y}} - \frac{\hat{x}}{\hat{y}}}\right\}\right] \end{aligned} \tag{2.15}$$

From Lemma 2.2, we have

$$\frac{\hat{x}}{\hat{y}} \leq \frac{1-\hat{x}}{1-\hat{y}}. \tag{2.16}$$

In other words, $\frac{1-\hat{x}}{1-\hat{y}} - \frac{\hat{x}}{\hat{y}} \geq 0$. Therefore, the sign of $\frac{dD_B}{d\rho}$ does not depend on $\frac{1-\hat{x}}{1-\hat{y}} - \frac{\hat{x}}{\hat{y}}$.

Also, using the properties of the $\log(\cdot)$ function, we have

$$\frac{1-\hat{y}}{1-\hat{x}} \leq \frac{\log\left(\frac{1-\hat{x}}{1-\hat{y}}\right) - \log\left(\frac{\hat{x}}{\hat{y}}\right)}{\frac{1-\hat{x}}{1-\hat{y}} - \frac{\hat{x}}{\hat{y}}} \leq \frac{\hat{y}}{\hat{x}}. \tag{2.17}$$

Substituting Equation (2.17) in Equation (2.15), we have

$$\begin{aligned} \left(\frac{1-\hat{x}}{1-\hat{y}} - \frac{\hat{x}}{\hat{y}}\right)^{-1} \frac{dD_B}{d\rho} &\leq (1-2y) - (1-2x)\left\{\frac{1-\hat{y}}{1-\hat{x}}\right\} \\ \left(\frac{1-\hat{x}}{1-\hat{y}} - \frac{\hat{x}}{\hat{y}}\right)^{-1} \frac{dD_B}{d\rho} &\leq \frac{-(y-x)}{1-\hat{x}} \end{aligned} \tag{2.18}$$

Since $\frac{dD_B}{d\rho} \leq 0$, D_B is a monotonically decreasing function of ρ, for all $\rho \in [0, \frac{1}{2}]$.

□

Having analyzed the impact of BSCs on the ROC, let us now shift our focus on finding those quantizers that maximize the KL Divergence at the sensor or the FC. Given any operating point $A = (x, y)$ at the sensor, we investigate the behavior of D_A with respect to y, for a fixed value of x.

Lemma 2.4. *The optimal quantizer always lies on the boundary of the set of all feasible quantizer designs.*

Proof. For a fixed value of x, we differentiate D_A with respect to y as follows.

$$\left.\frac{dD_A}{dy}\right|_{\text{fixed } x} = \frac{1-x}{1-y} - \frac{x}{y} \tag{2.19}$$

From Lemma 2.2, we have $\left.\frac{dD_A}{dy}\right|_{\text{fixed } x} \geq 0$. In other words, D_A is a monotonically increasing function of y, for a fixed value of x. Hence, we are always interested in quantizer rules whose operating points lie on the boundary of the set of all feasible quantizers. □

In summary, the sensor operating point chosen on the LRT boundary slides towards the point $(\frac{1}{2}, \frac{1}{2})$ as the channel deteriorates (increasing ρ), which, in turn, degrades the KLD of any decision rule γ to zero. Therefore, we address the problem of finding the operating point on the boundary which maximizes $\mathscr{D}_{FC}$, where the boundary is dictated by the Eve's constraint $\mathscr{D}_E = \alpha$ and the boundary of $\mathbf{\Gamma} = \{\Gamma_1, \cdots, \Gamma_N\}$.

2.4 Differential Secrecy

There exists a trade-off in the selection of binary sensor quantizers, as the loss in FC's performance is directly reflected in the performance loss at Eve. Therefore, in this section, we model this design trade-off by considering the difference in the KLDs at the FC and Eve as our performance metric. More specifically, we design sensor quantizers in a distributed detection network that maximize the difference in the KLDs at the FC and Eve, as stated in the following problem.

Problem 2.1. *Determine the set of optimal operating points* $(\boldsymbol{x}, \boldsymbol{y}) \in \mathbf{\Gamma}$ *at all the sensors that*

$$\begin{aligned} &\underset{(\boldsymbol{x},\boldsymbol{y})}{\textit{maximize}} \quad \mathscr{D}_{FC} - \mathscr{D}_E \\ &\textit{subject to} \quad (x_i, y_i) \in \Gamma_i, \textit{ for all } i = 1, 2, \ldots, N. \end{aligned}$$

Since the received symbols $\boldsymbol{v}_i$ and $\boldsymbol{w}_i$ at both the FC and Eve respectively, due to the i^{th} sensing agent are conditionally independent from those of other sensing agents, the difference $\mathscr{D}_{FC} - \mathscr{D}_E = \sum_{i=1}^{N} (D_{FC} - D_E)$ is linearly separable. Therefore, Problem 2.1 can be decomposed into N independent problems, which is stated below. For the sake of notational convenience, we ignore the sensor indices in the remaining section.

Problem 2.2. *Determine the optimal sensor operating point* $(x, y) \in \Gamma$ *that*

$$\underset{(x,y)}{\textit{maximize}} \quad D_{FC} - D_E$$

$$\textit{subject to} \quad (x, y) \in \Gamma.$$

We solve this problem in two different cases. In the first case, we assume that the Eve's channel is worse than FC's channel. In other words, we assume that $\rho_e \geq \rho_{fc}$. In the second case, we assume that the Eve has a better channel than FC, i.e., $\rho_e < \rho_{fc}$.

CASE 1: Eve has worse channels than FC

In this case, we assume that $\rho_e \geq \rho_{fc}$. Under this assumption, we show that the optimal binary quantizers always lie on the LRT boundary of the achievable region, as shown in Figure 2.3.

Lemma 2.5. *For a fixed* x*,* $L(y) = D_{FC} - D_E$ *is a monotonically increasing function of* y *in the achievable region* Γ*, if the Eve has a worse channel than FC.*

Proof. For a fixed x, we investigate the rate of change of $L(y) = D_{FC} - D_E$ by taking its derivative with respect to y, as follows.

$$\begin{aligned} \frac{dL(y)}{dy} &= \frac{dD_{FC}}{dy} - \frac{dD_E}{dy} \\ &= \frac{dy_{fc}}{dy} \cdot \frac{dD_{FC}}{dy_{fc}} - \frac{dy_e}{dy} \cdot \frac{dD_E}{dy_e} \\ &= (1 - 2\rho_{fc}) \left[\frac{1 - x_{fc}}{1 - y_{fc}} - \frac{x_{fc}}{y_{fc}} \right] - (1 - 2\rho_e) \left[\frac{1 - x_e}{1 - y_e} - \frac{x_e}{y_e} \right]. \end{aligned} \tag{2.20}$$

Given that $\rho_e \geq \rho_{fc}$, we have $1 - 2\rho_{fc} \geq 1 - 2\rho_e$. Furthermore, given an operating

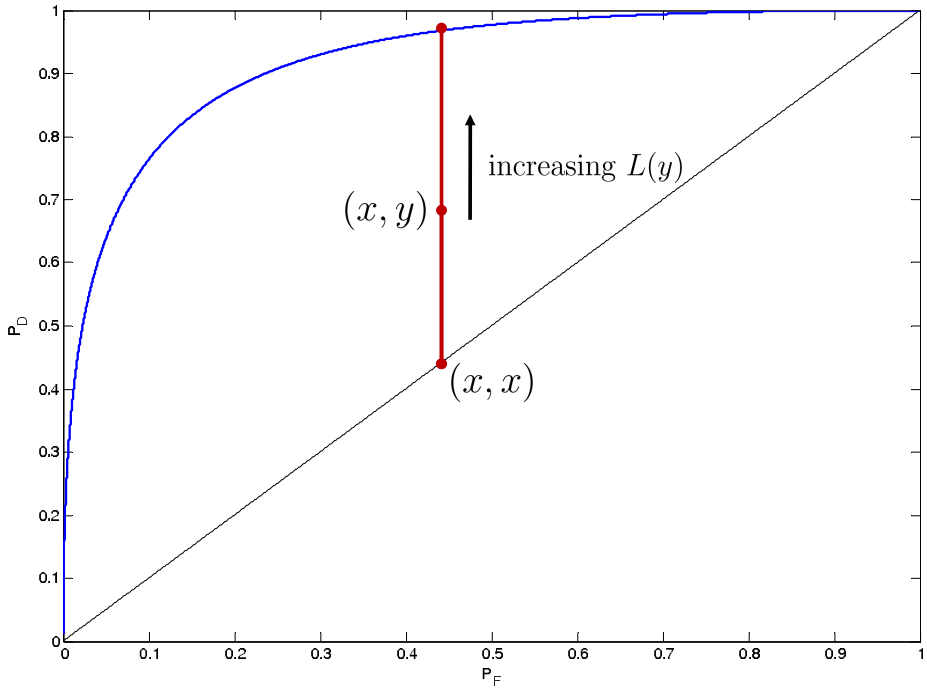


Figure 2.3: $L(y)$ increases with increasing y for a fixed value of x, when Eve has a worse channel than FC.

point (x, y), as stated in Lemma 2.2, we have

$$\frac{1 - x_{fc}}{1 - y_{fc}} - \frac{x_{fc}}{y_{fc}} \geq \frac{1 - x_e}{1 - y_e} - \frac{x_e}{y_e}. \tag{2.21}$$

As a result, we have

$$\frac{dL}{dy} \geq 0. \tag{2.22}$$

□

Since $L(y)$ is an increasing function of y for a given x, the optimal sensor operating point always lies on the LRT boundary when the Eve has a worse channel than the FC.

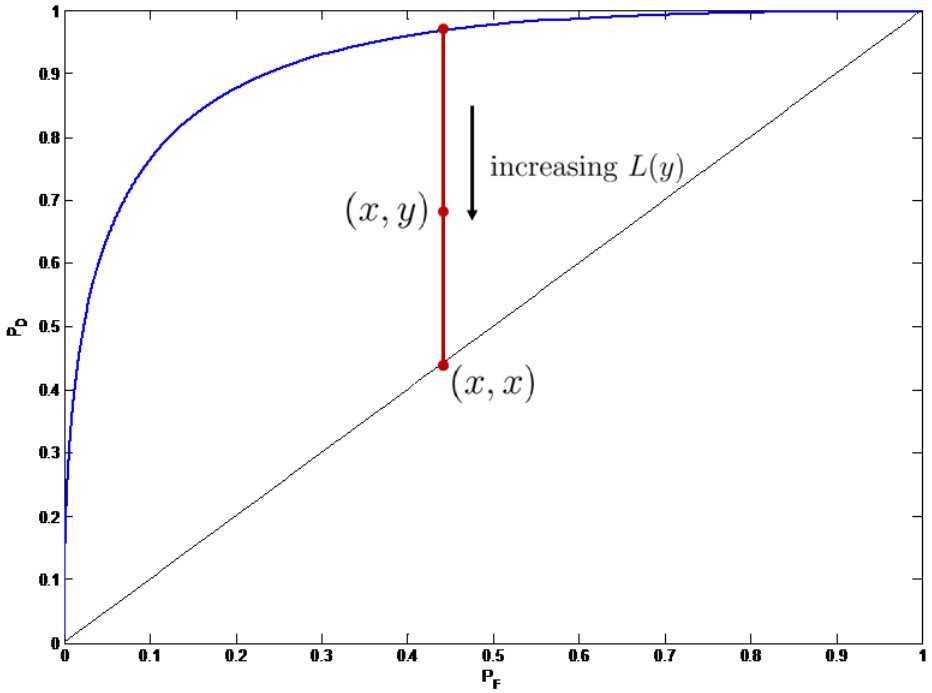


Figure 2.4: $L(y)$ increases with decreasing y for a fixed value of x, when Eve has a better channel than FC.

CASE 2: Eve has better channels than FC

Similar to Case 1, we now investigate the optimal sensor quantizers when Eve has a better channel than the FC, i.e., when $\rho_e < \rho_{fc}$. We show that the optimal binary quantizers always lie on the uninformative boundary of the achievable region (the line $y = x$), as shown in Figure 2.4.

Lemma 2.6. *For a fixed x, $L(y) = D_{FC} - D_E$ is a monotonically decreasing function of y in the achievable region Γ, if the Eve has a better channel than FC.*

Proof. As shown in the proof of Lemma 2.5, we calculate the derivative of $L(y) = D_{FC} -$

D_E with respect to y for a fixed x, as follows.

$$\frac{dL(y)}{dy} = (1-2\rho_{fc})\left[\frac{1-x_{fc}}{1-y_{fc}} - \frac{x_{fc}}{y_{fc}}\right] - (1-2\rho_e)\left[\frac{1-x_e}{1-y_e} - \frac{x_e}{y_e}\right]. \tag{2.23}$$

Given that $\rho_e < \rho_{fc}$, we have $1-2\rho_{fc} < 1-2\rho_e$. Furthermore, given an operating point (x, y), as stated in Lemma 2.2, we have

$$\frac{1-x_{fc}}{1-y_{fc}} - \frac{x_{fc}}{y_{fc}} < \frac{1-x_e}{1-y_e} - \frac{x_e}{y_e}. \tag{2.24}$$

As a result, we have

$$\frac{dL}{dy} < 0. \tag{2.25}$$

□

Combining Lemmas 2.5 and 2.6, Problem 2.2 reduces to a quantizer selection problem, as summarized below in the following theorem.

Theorem 2.1. *The structure of an optimal sensor quantizer is LRT-based, if Eve has a worse channel than the FC. Else, the optimal sensor quantizer is an uninformative rule.*

Note that the objective considered in this section, namely the difference in KLDs at the FC and Eve, does not constrain the Eve's performance. Consequently, Eve may acquire an intolerable amount of information from the sensors, and therefore, the solution (quantizer design) provided in this section may not be attractive to the network designer in many practical scenarios. Therefore, in the following section, we present another framework for secrecy where we impose a tolerable constraint on the Eve's performance in the design of optimal binary quantizers at the sensors.

2.5 Constrained Secrecy

In this section, we design a distributed detection network where $\mathscr{D}_{FC}$ is maximized while constraining $\mathscr{D}_E$ to a prescribed tolerance limit, denoted as α. We present the formal problem statement and discuss the various scenarios that are addressed in this chapter, as follows.

Problem 2.3. *Find*

$$\arg\max_{\gamma} \quad \mathscr{D}_{FC} \quad \textit{s.t.}$$

$$\textit{1.} \quad \mathscr{D}_E \leq \alpha$$

$$\textit{2.} \quad (x_i, y_i) \in \Gamma_i, \textit{ for all } i = 1, \cdots, N.$$

Note that Constraint 1 in the above problem statement becomes degenerate for large values of α. More specifically, Problem 2.3 is meaningful only when $0 \leq \alpha < \alpha^*$ so that it has a non-degenerate Constraint 1 in Problem 2.3. This critical value α^* is equal to Eve's KL Divergence $\mathscr{D}_E^*$, which Eve attains when FC attains the maximum KL Divergence $\mathscr{D}_{FC}^*$. This maximum KL Divergence $\mathscr{D}_{FC}^*$ can be found by solving Problem 2.3 in the absence of Constraint 1.

Let $\mathscr{R} \triangleq \cap_{i=1}^{N} \Gamma_i \cap \{ (\boldsymbol{x}, \boldsymbol{y}) \mid \mathscr{D}_E \leq \alpha \}$ denote the search space in Problem 2.3. Note that $\{(\boldsymbol{x}, \boldsymbol{y}) \mid \mathscr{D}_E \leq \alpha\}$ is a convex level-set of $\mathscr{D}_E$ [54], because $\mathscr{D}_E$ is a convex function of $(\boldsymbol{x}, \boldsymbol{y})$. Similarly, since LRTs are optimal in the absence of Eve (For a detailed proof, please refer to Proposition 4.1 in [62]), Γ_i is also a convex set in the ROC space. Also, $\mathscr{R}$ is an intersection of two convex sets, and therefore, $\mathscr{R}$ is a convex set.

Since $\mathscr{D}_{FC}$ is a convex function of $(\boldsymbol{x}, \boldsymbol{y})$, Problem 2.3 is a convex maximization problem, and therefore, the optimal solution is one of the extreme points of $\mathscr{R}$ [54]. Note that $\{(\boldsymbol{x}, \boldsymbol{y}) \mid \mathscr{D}_E \leq \alpha\}$ is not necessarily a subset of Γ_i, and therefore, the optimal set of binary

quantizers need not necessarily be LRTs. Furthermore, the search space $\mathscr{R}$ in Problem 2.3 is not a simple polytope. $\mathscr{R}$ is an intersection of two convex sets with smooth boundaries and therefore, its boundary does not necessarily have a smooth differential at every point. Consequently, optimal search algorithms proposed to solve traditional convex maximization problems with polytope search spaces cannot be applied to find the optimal solution of Problem 2.3, as our problem demands a more detailed analysis of the boundary of the search space.

Therefore, in Section 2.5.1, we first restrict our attention to a simpler scenario[3] where all the sensors' observations are identically distributed and, where all the channels between the sensors and the FC (likewise, channels between sensors and the Eve) are identical. This assumption results in the received symbols at the FC (likewise, received symbols at the Eve) being conditionally i.i.d., thus decomposing the problem into a distributed framework of N identical sub-problems. In Section 2.5.2, we consider a more general scenario[4] where the sensor observations are conditionally independent and non-identically distributed, and the channels between the sensors and the FC (likewise, channels between sensors and the Eve) are also non-identical. In both these scenarios, we investigate the design of secure binary quantizers when $\alpha < \mathscr{D}_E^*$.

2.5.1 Identical Sensors and Channels

In this section, we address the problem of designing optimal quantizers when all the sensors and the channels between the sensors and the FC (likewise, channels between sensors and the Eve) are identical.

[3] In this chapter, we call this scenario as "identical sensors and channels".
[4] Similarly, we call this scenario as "non-identical sensors and channels".

For all $i = 1, \cdots, N$, we have

$$\begin{aligned} p_{i,0}(x) &= p_0(x), \quad & p_{i,1}(x) &= p_1(x) \\ x_i &= x, & y_i &= y \\ \rho_{fc_i} &= \rho_{fc}, & \rho_{e_i} &= \rho_e \end{aligned} \tag{2.26}$$

Since all the sensors and their corresponding channels are identical, we remove the sensor-indices for notational simplicity. Therefore, we have $x_{fc_i} = x_{fc}$, $y_{fc_i} = y_{fc}$, $x_{e_i} = x_e$ and $y_{e_i} = y_e$ for all $i = 1, \cdots, N$. Because of this, $D_i = D$, $D_{FC_i} = D_{FC}$ and $D_{E_i} = D_E$ for all $i = 1, \cdots, N$, and consequently, the KLD at the FC and Eve reduces to $\mathscr{D}_{FC} = ND_{FC}$ and $\mathscr{D}_E = ND_E$. In other words, Problem 2.3 reduces to the design of the quantizer at one of the identical sensors as follows.

Problem 2.4. *Find*

$$\begin{aligned} \arg\max_{\gamma} \quad & D_{FC} \quad \textit{s.t.} \\ & \textit{1.} \quad D_E \leq \tilde{\alpha} \\ & \textit{2.} \quad (x, y) \in \Gamma. \end{aligned}$$

where $\tilde{\alpha} = \dfrac{\alpha}{N}$.

Note that, although Problem 2.4 is still a convex maximization problem, due to its reduced dimensionality, the problem becomes tractable. In the remaining section, we find the optimal quantizer in two stages. First, we find the structure of the optimal binary quantizers by gaining insights into the behavior of D_{FC} on the boundary of the Eve's constraint $\{(x, y) \mid D_E \leq \tilde{\alpha}\}$. Then, we present an algorithm to find the optimal threshold for this quantizer.

We start our investigation of the behavior of D_{FC} on the boundary of the Eve's constraint $\{(x, y) \mid D_E \leq \tilde{\alpha}\}$ by determining the necessary conditions for guaranteeing $D_E = \tilde{\alpha}$ in the following lemma.

Lemma 2.7. *If the transition probability of the Eve's BSCs satisfies* $\rho_e < \frac{1}{2}$*, the two necessary conditions for any sensor operating point* (x, y) *to guarantee* $D_E = \tilde{\alpha}$ *in the ROC space are stated as follows.*

$$\frac{dy}{dx} = \frac{\log\left(\frac{1-x_e}{1-y_e}\right) - \log\left(\frac{x_e}{y_e}\right)}{\frac{1-x_e}{1-y_e} - \frac{x_e}{y_e}} \tag{2.27}$$

and

$$\begin{aligned}\left(\frac{1-x_e}{1-y_e} - \frac{x_e}{y_e}\right)\frac{d^2y}{dx^2} &= (1-2\rho_e)\left[-\left(\frac{1-x_e}{(1-y_e)^2} + \frac{x_e}{y_e^2}\right)\left(\frac{dy}{dx}\right)^2\right. \\ &\qquad \left. +2\left(\frac{1}{y_e} + \frac{1}{1-y_e}\right)\frac{dy}{dx} - \left(\frac{1}{x_e} + \frac{1}{1-x_e}\right)\right].\end{aligned} \tag{2.28}$$

Proof. Since D_E is a constant (equal to the fixed design-parameter $\tilde{\alpha}$), its first two derivatives are equal to zero. We employ these to prove the lemma.

First, we differentiate D_E with respect to x and equate it to zero, as follows.

$$\begin{aligned}\frac{dD_E}{dx} &= \frac{d}{dx}\left[x_e \log\frac{x_e}{y_e} + (1-x_e)\log\left(\frac{1-x_e}{1-y_e}\right)\right] \\ &= (1-2\rho_e)\left[\left(\frac{1-x_e}{1-y_e} - \frac{x_e}{y_e}\right)\frac{dy}{dx} - \left\{\log\left(\frac{1-x_e}{1-y_e}\right) - \log\left(\frac{x_e}{y_e}\right)\right\}\right] \qquad (2.29) \\ &= 0.\end{aligned}$$

Rearranging the terms in Equation (2.29), we can obtain Equation (2.27).

Next, we differentiate Equation (2.29) again with respect to x as follows, in order to

find a closed-form expression for $\frac{d^2y}{dx^2}$.

$$\begin{aligned}
\frac{d^2D_E}{dx^2} &= (1-2\rho_e)\frac{d}{dx}\left[\left(\frac{1-x_e}{1-y_e}-\frac{x_e}{y_e}\right)\frac{dy}{dx}-\left\{\log\left(\frac{1-x_e}{1-y_e}\right)-\log\left(\frac{x_e}{y_e}\right)\right\}\right] \\
&= (1-2\rho_e)\left[\left(\frac{1-x_e}{1-y_e}-\frac{x_e}{y_e}\right)\frac{d^2y}{dx^2}+(1-2\rho_e)\left(\frac{1-x_e}{(1-y_e)^2}+\frac{x_e}{y_e^2}\right)\left(\frac{dy}{dx}\right)^2\right. \\
&\qquad \left.-2(1-2\rho_e)\left(\frac{1}{y_e}+\frac{1}{1-y_e}\right)\frac{dy}{dx}+(1-2\rho_e)\left(\frac{1}{x_e}+\frac{1}{1-x_e}\right)\right]. \\
&= 0.
\end{aligned} \tag{2.30}$$

Rearranging the terms in Equation (2.30), we can obtain Equation (2.28). □

Note that Equation (2.29) in Lemma 2.7 provides the slope of the Eve's constraint boundary $D_E = \tilde{\alpha}$. Since the slope of y with respect to x along the boundary $D_E = \tilde{\alpha}$ has a structure similar to the slope of a line joining two points on a logarithmic curve as seen in Equation (2.27), we present lower and upper bounds for the slope of this boundary curve $D_E = \tilde{\alpha}$ in the ROC plane in the following lemma.

Lemma 2.8. *The slope of the Eve's constraint boundary in the ROC plane, as defined by the set of points* $\{\ (x,y) \mid D_E = \tilde{\alpha}\ \}$*, is bounded on both sides as follows.*

$$\frac{x_e}{y_e} \leq \frac{dy}{dx} \leq \frac{1-x_e}{1-y_e}. \tag{2.31}$$

Proof. Given two points $a \geq b$, due to the concavity of the $\log(\cdot)$ function, the slope of the line joining $(a, \log a)$ and $(b, \log b)$ always lies between the slopes of the $\log(\cdot)$ at points a and b respectively Hence, this results in Equation (2.31). □

Note that the necessary conditions for any operating point (x, y) to lie on the Eve's constraint boundary $\{\ (\boldsymbol{x}, \boldsymbol{y}) \mid D_E = \tilde{\alpha}\ \}$, as stated in Lemma 2.7, and the bounds on the

slope of the same boundary curve, as given in Lemma 2.8, are essential to our analysis of the behavior of the sensor's KL divergence D, and the FC's KL Divergence, D_{FC}, in terms of the false alarm probability x along the Eve's constraint, which is defined by $D_E = \tilde{\alpha}$.

First, we investigate the behavior of the KL Divergence at the sensor, which is denoted as $D(x, y)$, along the Eve's constraint $D_E(x, y) = \tilde{\alpha}$. Note that this analysis can be equivalently interpreted as the case where we investigate the behavior of D_{FC} when the channels between the sensors and the FC are ideal. In the following proposition, we prove that $D(x, y)$ is a convex function of x along the curve $D_E(x, y) = \tilde{\alpha}$.

Proposition 2.1. *Given that the Eve's channel is a BSC with transition probability $\rho_e < \frac{1}{2}$, D is strictly a convex function of x, for all operating points that lie in the set $\{(x, y) \mid D_E = \tilde{\alpha}\}$.*

Proof. To show that D is a convex function of x in the presence of a constraint on Eve, we investigate the second-order differential of D with respect to x.

The closed-form expression for the first-order differential of D with respect to x

$$\begin{aligned} \frac{dD}{dx} &= \frac{d}{dx}\left[x \log \frac{x}{y} + (1-x) \log\left(\frac{1-x}{1-y}\right)\right] \\ &= \left(\frac{1-x}{1-y} - \frac{x}{y}\right)\frac{dy}{dx} - \left[\log\left(\frac{1-x}{1-y}\right) - \log\left(\frac{x}{y}\right)\right]. \end{aligned} \tag{2.32}$$

The second-order differential of D can therefore be obtained by differentiating Equation (2.32) with respect to x as follows.

$$\begin{aligned} \frac{d^2D}{dx^2} &= \left(\frac{1-x}{1-y} - \frac{x}{y}\right)\frac{d^2y}{dx^2} + \left(\frac{1-x}{(1-y)^2} + \frac{x}{y^2}\right)\left(\frac{dy}{dx}\right)^2 \\ &\quad -2\left(\frac{1}{y} + \frac{1}{1-y}\right)\frac{dy}{dx} + \left(\frac{1}{x} + \frac{1}{1-x}\right). \end{aligned} \tag{2.33}$$

Note that the first term in Equation (2.33) can be rewritten as follows.

$$\left(\frac{1-x}{1-y}-\frac{x}{y}\right)\frac{d^2y}{dx^2} = \frac{\left(\frac{1-x}{1-y}-\frac{x}{y}\right)}{\left(\frac{1-\hat{x}}{1-\hat{y}}-\frac{\hat{x}}{\hat{y}}\right)}\left(\frac{1-\hat{x}}{1-\hat{y}}-\frac{\hat{x}}{\hat{y}}\right)\frac{d^2y}{dx^2} \tag{2.34}$$

$$= \frac{\hat{y}(1-\hat{y})}{y(1-y)}\cdot\frac{1}{(1-2\rho)}\cdot\left(\frac{1-\hat{x}}{1-\hat{y}}-\frac{\hat{x}}{\hat{y}}\right)\frac{d^2y}{dx^2}$$

Note that Equation (2.34) allows us to use the necessary condition for the operating point (x, y) to lie on the Eve's constraint curve $D_E = \tilde{\alpha}$, as given in Equation (2.28). Therefore, we substitute Equation (2.28) from the Lemma 2.7 in Equation (2.34), and use this in Equation (2.33) to have the following.

$$\frac{d^2D}{dx^2} = T_1\left(\frac{dy}{dx}\right)^2 - 2T_2\frac{dy}{dx} + T_3 \tag{2.35}$$

where

$$T_1 = \left(\frac{1-x}{(1-y)^2}+\frac{x}{y^2}\right) - \frac{\hat{y}(1-\hat{y})}{y(1-y)}\left(\frac{1-\hat{x}}{(1-\hat{y})^2}+\frac{\hat{x}}{\hat{y}^2}\right) \tag{2.36a}$$

$$T_2 = \left(\frac{1}{y}+\frac{1}{(1-y)}\right) - \frac{\hat{y}(1-\hat{y})}{y(1-y)}\left(\frac{1}{\hat{y}}+\frac{1}{(1-\hat{y})}\right) \tag{2.36b}$$

$$T_3 = \left(\frac{1}{x}+\frac{1}{(1-x)}\right) - \frac{\hat{y}(1-\hat{y})}{y(1-y)}\left(\frac{1}{\hat{x}}+\frac{1}{(1-\hat{x})}\right). \tag{2.36c}$$

It is easy to show that $T_2 = 0$.

So, let us first consider T_1. Expanding Equation (2.36a), we have

$$\begin{aligned} T_1 &= \frac{(x\hat{y} - \hat{x}y) - (x\hat{y}^2 - \hat{x}y^2) + y\hat{y}\left\{(y - \hat{y}) - 2(x - \hat{x}) + 2(x\hat{y} - \hat{x}y)\right\}}{y^2(1-y)^2\hat{y}(1-\hat{y})} \\ &= \frac{-\rho(y-x) - \{\rho^2 x - \rho y^2 + 2\rho(1-2\rho)xy - 2\rho(1-2\rho)xy^2\} + y\hat{y}(\rho - 2\rho x)}{y^2(1-y)^2\hat{y}(1-\hat{y})} \\ &= \frac{\rho(1-\rho)(y-x)(2y-1)}{y^2(1-y)^2\hat{y}(1-\hat{y})} \end{aligned} \tag{2.37}$$

Similarly, expanding Equation 2.36c for T_3, we have

$$\begin{aligned} T_3 &= \frac{1}{y(1-y)}\left[\frac{y(1-y)}{x(1-x)} - \frac{\hat{y}(1-\hat{y})}{\hat{x}(1-\hat{x})}\right] \\ &= \frac{\rho(1-\rho)}{y(1-y)} \cdot \frac{(y-x)(1-x-y)}{x(1-x)\hat{x}(1-\hat{x})} \end{aligned} \tag{2.38}$$

Substituting Equations (2.37) and (2.38) in Equation (2.35), we simplify Equation (2.35) into the following.

$$\frac{d^2D}{dx^2} = \frac{\rho(1-\rho)(y-x)}{y(1-y)} \cdot T_4 \tag{2.39}$$

where

$$T_4 = \frac{2y-1}{y\hat{y}(1-y)(1-\hat{y})}\left(\frac{dy}{dx}\right)^2 + \frac{1-x-y}{x\hat{x}(1-x)(1-\hat{x})}. \tag{2.40}$$

Note that, if $T_4 \geq 0$, D is a convex function of x along the Eve's constraint curve $D_E = \tilde{\alpha}$. Since we are only interested in the region where $y \geq x$ and $\rho < \frac{1}{2}$ for all practical purposes, we restrict our analysis of the sign of T_4 in this region.

In order to analyze the sign of T_4, we divide the achievable region in the receiver-operating characteristics into three regions, as shown in Figure 2.5.

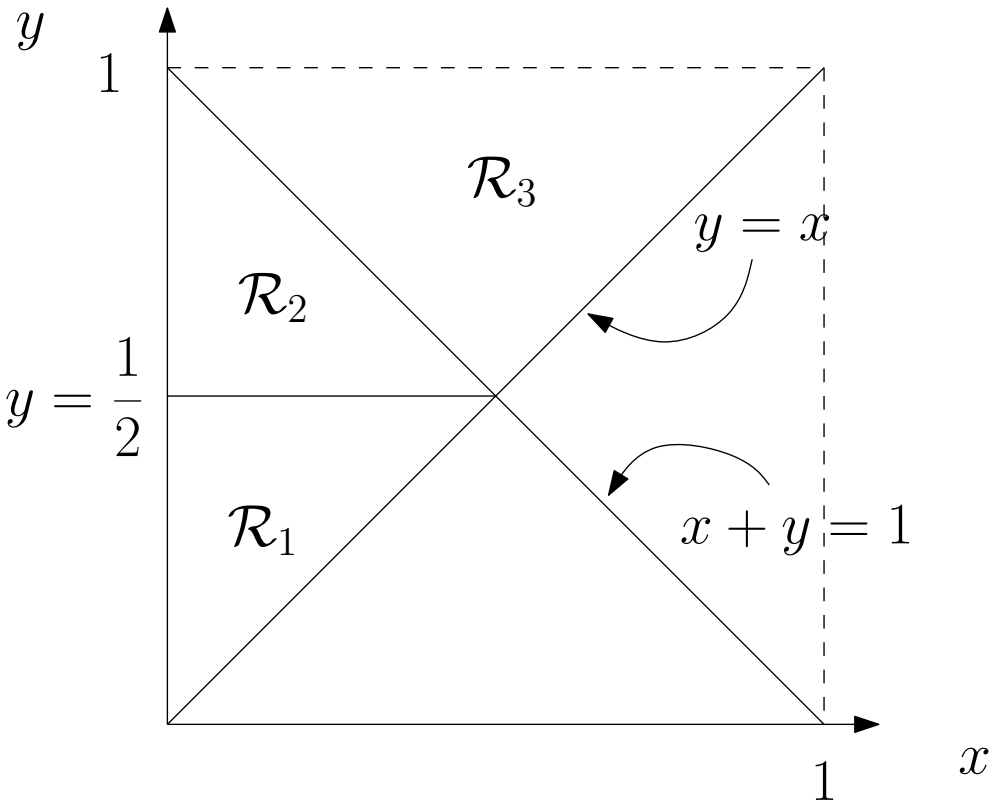


Figure 2.5: Partition of ROC into three regions

$$\begin{aligned}
\mathscr{R}_1 &: \left(y \leq \frac{1}{2}\right) \& \left(x + y \leq 1\right) \\
\mathscr{R}_2 &: \left(y \geq \frac{1}{2}\right) \& \left(x + y \leq 1\right) \\
\mathscr{R}_3 &: \left(y \geq \frac{1}{2}\right) \& \left(x + y \geq 1\right).
\end{aligned} \tag{2.41}$$

Obviously, in region $\mathscr{R}_2$, $2y - 1 \geq 0$ and $1 - x - y \geq 0$. Therefore, $\frac{d^2 D}{dx^2} \geq 0$. Henceforth, we analyze the sign of T_4 in the remaining regions $\mathscr{R}_1$ and $\mathscr{R}_3$.

Region $\mathscr{R}_1$ In this region, $2y - 1 \leq 0$. Therefore, we use the upper bound on $\frac{dy}{dx}$, presented in Equation (2.31), to find the sign of T_4 as follows.

Substituting Equation (2.31) in Equation (2.40), we have

$$\begin{aligned} T_4 &\geq \frac{1-x-y}{x\hat{x}(1-x)(1-\hat{x})} - \frac{1-2y}{y\hat{y}(1-y)(1-\hat{y})}\frac{y\hat{y}}{x\hat{x}} \\ &= \frac{1}{x\hat{x}}\left[\frac{1-x-y}{(1-x)(1-\hat{x})} - \frac{1-2y}{(1-y)(1-\hat{y})}\right] \\ &= \frac{(1-x-y)(1-y)(1-\hat{y}) - (1-y)(1-x)(1-\hat{x}) + y(1-x)(1-\hat{x})}{x\hat{x}(1-x)(1-\hat{x})(1-y)(1-\hat{y})} \end{aligned} \tag{2.42}$$

Equation (2.42) can be rearranged as follows.

$$T_4 \geq \frac{(y-x)\left[y(1-\rho) + (1-2\rho)\left\{(2y-1)(1-x) - y^2\right\}\right]}{x\hat{x}(1-x)(1-\hat{x})(1-y)(1-\hat{y})} \tag{2.43}$$

Since $1 - x - y \geq 0$ in region $\mathscr{R}_1$, we have $1 - x \geq y$. Therefore, substituting this inequality in Equation (2.43), we have

$$\begin{aligned} T_4 &\geq \frac{(y-x)\left[y(1-\rho) + (1-2\rho)\left\{(2y-1)y - y^2\right\}\right]}{x\hat{x}(1-x)(1-\hat{x})(1-y)(1-\hat{y})} \\ &= \frac{(y-x)\left[y(1-\rho) + (1-2\rho)y(y-1)\right]}{x\hat{x}(1-x)(1-\hat{x})(1-y)(1-\hat{y})} \\ &= \frac{(y-x)y\left[(1-\rho) + (1-2\rho)(y-1)\right]}{x\hat{x}(1-x)(1-\hat{x})(1-y)(1-\hat{y})} \\ &= \frac{(y-x)y\hat{y}}{x\hat{x}(1-x)(1-\hat{x})(1-y)(1-\hat{y})} \\ &\geq 0. \end{aligned} \tag{2.44}$$

Region $\mathscr{R}_3$ In this region, since $2y-1 \geq 0$, we use the lower bound on $\dfrac{dy}{dx}$, presented in Equation (2.31), in order to find the sign of T_4.

Substituting Equation (2.31) in Equation (2.40), we have

$$\begin{aligned}
T_4 &\geq \frac{2y-1}{y\hat{y}(1-y)(1-\hat{y})}\frac{(1-y)(1-\hat{y})}{(1-x)(1-\hat{x})} - \frac{x+y-1}{x\hat{x}(1-x)(1-\hat{x})} \\
&= \frac{1}{(1-x)(1-\hat{x})}\left[\frac{2y-1}{y\hat{y}} - \frac{x+y-1}{x\hat{x}}\right]. \\
&= \frac{(y\hat{y}-x\hat{x}) - y(y\hat{y}-x\hat{x}) - xy(\hat{y}-\hat{x})}{x\hat{x}y\hat{y}(1-x)(1-\hat{x})} \qquad (2.45) \\
&= \frac{(1-y)\left\{\rho(y-x)+(1-2\rho)(y^2-x^2)\right\} - xy(1-2\rho)(y-x)}{x\hat{x}y\hat{y}(1-x)(1-\hat{x})} \\
&= \frac{(y-x)\left\{\rho(1-y)+(1-2\rho)y(1-y) - (1-2\rho)x(1-2y)\right\}}{x\hat{x}y\hat{y}(1-x)(1-\hat{x})}
\end{aligned}$$

Since we are only interested in the region where $y \geq x$, Equation (2.46) can be lower-bounded as follows.

$$\begin{aligned}
T_4 &\geq \frac{1}{x\hat{x}y\hat{y}(1-x)(1-\hat{x})}\left[(y-x)\left\{\rho(1-y)\right.\right. \\
&\qquad \left.\left. +(1-2\rho)x(1-y) - (1-2\rho)x(1-2y)\right\}\right] \qquad (2.46) \\
&= \frac{(y-x)\left[\rho(1-y)+(1-2\rho)xy\right]}{x\hat{x}y\hat{y}(1-x)(1-\hat{x})} \\
&\geq 0.
\end{aligned}$$

Hence, for BSCs with $\rho < \frac{1}{2}$, D is a convex function of x along the constraint $D_E = \alpha$. $\square$

For any general BSC between the sensors and the FC, the sensor's operating point (x, y) transforms linearly into (x_{fc}, y_{fc}). Consequently, we have the following proposition, where we analyze the behavior of D_{FC} for any general BSC.

Proposition 2.2. *Let the BSCs corresponding to the FC and Eve have transition probabilities $0 < \rho_{fc}, \rho_e < \frac{1}{2}$. Then, D_{FC} is strictly a convex function of x, for all operating points that lie in the set $\{(x, y) \mid D_E = \tilde{\alpha}\}$.*

Proof. Note that (x_{fc}, y_{fc}) is a linear transformation of (x, y). This can be mathematically expressed as follows.

$$\begin{bmatrix} x_{fc} \\ y_{fc} \end{bmatrix} = \rho_{fc} \begin{bmatrix} 1 \\ 1 \end{bmatrix} + (1 - 2\rho_{fc}) \begin{bmatrix} x \\ y \end{bmatrix}. \tag{2.47}$$

In other words, a composition of D with an affine transformation, as given in Equation (2.47), results in D_{FC}. Consequently, since D is a convex function, D_{FC} is also a convex function [10]. □

Thus, for any BSC with transition probability ρ_{fc} corresponding to the FC, D_{FC} is a convex function of x. In other words, among the set of operating points that lie on the Eve's constraint boundary $D_E = \tilde{\alpha}$, the quantizers that maximize D_{FC} always lie on the intersection of the LRT curve $y = g_{LRT}(x)$ and the Eve's constraint boundary $D_E = \tilde{\alpha}$. As a consequence, the optimal quantizer is LRT-based, which we state in the following theorem.

Theorem 2.2. *The optimal quantizer that maximizes the FC's KL Divergence D_{FC} in the presence of a constraint on Eve's KL Divergence $D_E = \tilde{\alpha}$ is a likelihood ratio quantizer.*

Proof. Let $\mathscr{R}_i \triangleq \Gamma_i \cup \{(x, y) \mid D_E = \tilde{\alpha}\}$ denote the search space in Problem 2.4. We know, from Proposition 2.1, that D_{FC} is convex with respect to x along the Eve's constraint boundary on the ROC plane. Therefore, the solution of Problem 2.4 always lies

on the extreme points of the set of operating points on the Eve's constraint boundary $\{(x, y) \mid D_E = \tilde{\alpha}\}$. Note that the region of the Eve's constraint boundary that lies within $\mathscr{R}_i$ depends on the choice of $\tilde{\alpha}$.

Let D_E^* be the maximum KL Divergence at the Eve when the sensor employs the optimal solution to the unconstrained problem where Constraint 1 is not considered in Problem 2.4. In the regard, the following two cases arise:

- Case-1 [$\tilde{\alpha} \geq D_E^*$]: Note that, $\Gamma_i \subseteq \{(x, y) \mid D_E \leq \tilde{\alpha}\}$ in this case because the Eve's KL Divergence is always within the tolerable limit when the sensor employs any operating point $(x, y) \in \Gamma_i$. Therefore, the solution to Problem 2.4 is the optimal LRT in this case [62].

- Case-2 [$\tilde{\alpha} \leq D_E^*$]: This is equivalent to the case where $\Gamma_i \nsubseteq \{(x, y) \mid D_E \leq \tilde{\alpha}\}$. Note that we also have $\Gamma_i \not\supset \{(x, y) \mid D_E \leq \tilde{\alpha}\}$ since there always exist operating points $(x, y) \in \Gamma_i$ such that $D_E \leq \tilde{\alpha}$. Therefore, the boundaries of Γ_i and $\{(x, y) \mid D_E \leq \tilde{\alpha}\}$ both intersect each other. As discussed earlier in this proof, since the optimal solution is an extreme point of the Eve's constraint boundary $D_E = \tilde{\alpha}$, this is one of the intersection points that also lies on the boundary of Γ_i. In other words, the optimal sensor quantizer that solves Problem 2.4 is a LRT.

□

As discussed in the proof of Theorem 2.2, the problem of finding the optimal quantizer reduces to the problem of finding the intersection points of the boundaries of Γ_i and the Eve's constraint $\{(x, y) \mid D_E \leq \tilde{\alpha}\}$, and thereby, finding the corresponding threshold for the optimal LRT at the sensor.

Algorithm to find the Optimal Threshold

Let $f(x) \triangleq D_{FC}(x, y = g_{LRT}(x))$. For the sake of tractability, we consider the problem of finding optimal thresholds when $f(x)$ is a quasi-concave[5] function of x. As shown in Proposition 2.1, since the Eve's constraint translates into the convexity of D_{FC} with respect to x, there are at most two points of intersection for the curves $y = g_{LRT}(x)$ and $D_E = \tilde{\alpha}$, of which, one of them corresponds to the optimal quantizer. We present this formally in the following claim.

Claim 2.1. *Let $f(x) \triangleq D_{FC}(x, y = g_{LRT}(x))$. If $f(x)$ is a quasi-concave function of x, then there are at most two intersection points for the curves $y = g_{LRT}(x)$ and $D_E = \tilde{\alpha}$. The optimal quantizer corresponds to one of the two intersection points.*

Therefore, the problem reduces to finding these two intersection points and comparing them with respect to each other in terms of their respective D_{FC}. Moreover, we wish to find the threshold λ^* for the LRT that maximizes D_{FC} in the presence of Eve's constraint. Since, both x and y are tail-probabilities where the start of the tail is the threshold, x and y are both monotonically decreasing functions of the threshold λ. Therefore, we have the following claim.

Claim 2.2. *The two intersection points can be found by investigating the zeros of the function $h(\lambda) \triangleq D_E(x(\lambda), y(\lambda)) - \tilde{\alpha}$, where x and y are parameterized by the LRT threshold λ.*

Let $\tilde{\alpha}_{max}$ denote the value of KL Divergence at which D_E reaches its maximum value. In other words, the optimal quantizer in the absence of Eve (equivalent to $\tilde{\alpha} = \infty$), denoted as the operating point (x_∞, y_∞), is the same as the optimal quantizer for any $\tilde{\alpha} \geq \tilde{\alpha}_{max}$.

[5]Note that

$$\lim_{x \to 0} f(x) = 0, \quad \lim_{x \to 1} f(x) = 0 \tag{2.48}$$

Since, KLD is always non-negative, we always have $f(x) \geq 0$. Also, since any LRT curve $y = g_{LRT}(x)$ cuts through the level-sets of D_{FC} and is concave, $f(x)$ is a quasi-concave function of x.

Obviously, the function $h(\lambda)$ has two real zeros only when $\tilde{\alpha} < \tilde{\alpha}_{max}$. Note that only one of them provides the maximum KL Divergence at the FC.

In order to find both zeros of the function $h(\lambda) = 0$, we use the bisection method where we first find the point λ^* at which $h(\lambda)$ attains its maximum value. Then, consider two points, one on either side of λ^* (which are at a significant distance from λ^*) as initial points and use the bisection algorithm to find the roots of $h(\lambda) = 0$. We call these two zeros as λ_1 and λ_2. Then, we compute and compare D_{FC} at the operating points $(x(\lambda_1), y(\lambda_1))$ and $(x(\lambda_2), y(\lambda_2))$. We choose that threshold as the optimal choice, which results in the maximum D_{FC}.

For the sake of illustration, we present an example where the sensors observe the presence or absence of a known deterministic signal, which is corrupted by additive Gaussian noise.

Illustrative Example

We have so far shown that the optimal quantizer lies at the intersection of the curves $D_E = \tilde{\alpha}$ and the LRT boundary in the ROC. But, the structure of the LRT is specific to the observation model, and therefore, it is difficult to characterize the optimal sensor quantizer, in general. Therefore, we illustrate the design methodology for an example, where the sensors observe the presence or absence of a known deterministic signal. In other words, the observations at the i^{th} sensor are modeled as follows.

$$r_{i,t} = \begin{cases} n_{i,t} & \text{if } H_0 \\ \theta + n_{i,t} & \text{if } H_1 \end{cases} \tag{2.49}$$

where θ is the signal-of-interest and $n_{i,t} \sim \mathcal{N}(0, \sigma^2)$ is the additive Gaussian noise with zero mean and variance σ^2. Then, the probabilities of false alarm and detection are given by

$$x = Q\left(\frac{\lambda}{\sigma}\right),\ y = Q\left(\frac{\lambda - \theta}{\sigma}\right) \tag{2.50}$$

where $Q(\cdot)$ is the tail probability of the standard normal distribution $\mathcal{N}(0,1)$.

Substituting Equation (2.50) in Equation (2.2), we obtain the KL Divergence at the sensor, which is observed to be concave for this example. Therefore, as stated in Claim 2.1, the optimum quantizer is given by the intersection of the LRT boundary in the ROC with the Eve's constraint $D_E = \tilde{\alpha}$.

Note that Equation (2.50) is a parameterization of the LRT boundary, where both the ROC's coordinates are parameterized with the threshold of the LRT. Since we are interested in the intersection of the LRT's boundary in the ROC with the Eve's constraint $D_E = \tilde{\alpha}$, we substitute $x_e = \rho_e + (1-2\rho_e)Q\left(\frac{\lambda}{\sigma}\right)$ and $y_e = \rho_e + (1-2\rho_e)Q\left(\frac{\lambda-\theta}{\sigma}\right)$ in D_E to obtain $h(\lambda) = D_E(x(\lambda), y(\lambda)) - \tilde{\alpha}$.

As shown in Figure 2.6, $h(\lambda)$ is a quasi-concave function of λ, with the tails converging to $-\tilde{\alpha}$. In other words, there are at most two zero-crossings since the function $h(\lambda)$ is unimodal with the two tails converging to a value less than zero. Therefore, there are at most two solutions to the equation $h(\lambda) = 0$. The optimum sensor threshold can be found by investigating the two zeros of $h(\lambda)$, as suggested in Claim 2.2, and comparing them in terms of D_{FC}.

Discussion and Results

In this subsection, we first discuss the impact of the secrecy constraint on the performance of the sensor network. Obviously, when we consider $\tilde{\alpha} = 0$, the network achieves perfect secrecy. But, this also forces the network to be *blind* in that $D_{FC} \to 0$. On the other extreme, consider a scenario where $\tilde{\alpha} \to \infty$. This is equivalent to the case where there is no eavesdropper present in the network. In other words, the optimal quantizer is given by (x_∞, y_∞). For any finite $\tilde{\alpha} > 0$, we numerically investigate the tradeoff between secrecy and performance of a given distributed detection system.

Since $\tilde{\alpha}$ is the tolerable limit on the performance of Eve, the greater the information leakage we can tolerate, the better the performance of the distributed detection network.

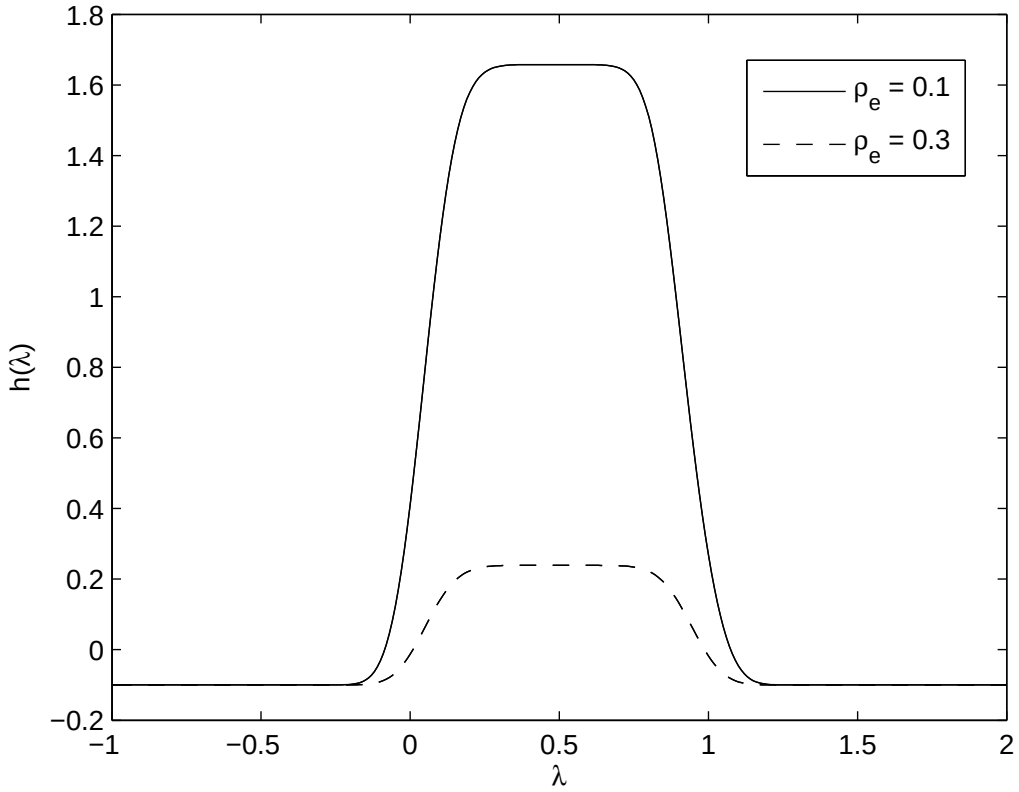


Figure 2.6: Plot of $h(\lambda)$ as a function of λ

This tradeoff is captured by Figure 2.7, where the maximum D_{FC} in the presence of a constrained Eve increases with increasing $\tilde{\alpha}$. Note that, beyond a certain value of $\tilde{\alpha}$, the maximum D_{FC} gets saturated to the optimal KLD at the FC in the absence of Eve. This saturation level for this example is 5.8 and it is dictated by the fundamental limits enforced by the imperfect observations and channel models within the network.

Next, we demonstrate the impact of the Eve's constraint on the ROC, as well as the KL Divergence at the FC, in Figure 2.8, when the FC's channels are ideal ($\rho_{fc} = 0$). Note that this argument can be carried over to any general BSC at the FC, as the operating point (x_{fc}, y_{fc}) is a linear transformation of (x, y). In Figure 2.8, we assume $\rho_e = 0.1$ and consider two different values of $\tilde{\alpha}$. In Figure 2.8a, we plot the constraint curve $D_E = \tilde{\alpha}$ along with the sensor's ROC. Note that the constraint curve intersects the LRT curve at two distinct points, as stated earlier. One of these two intersection points (the intersection

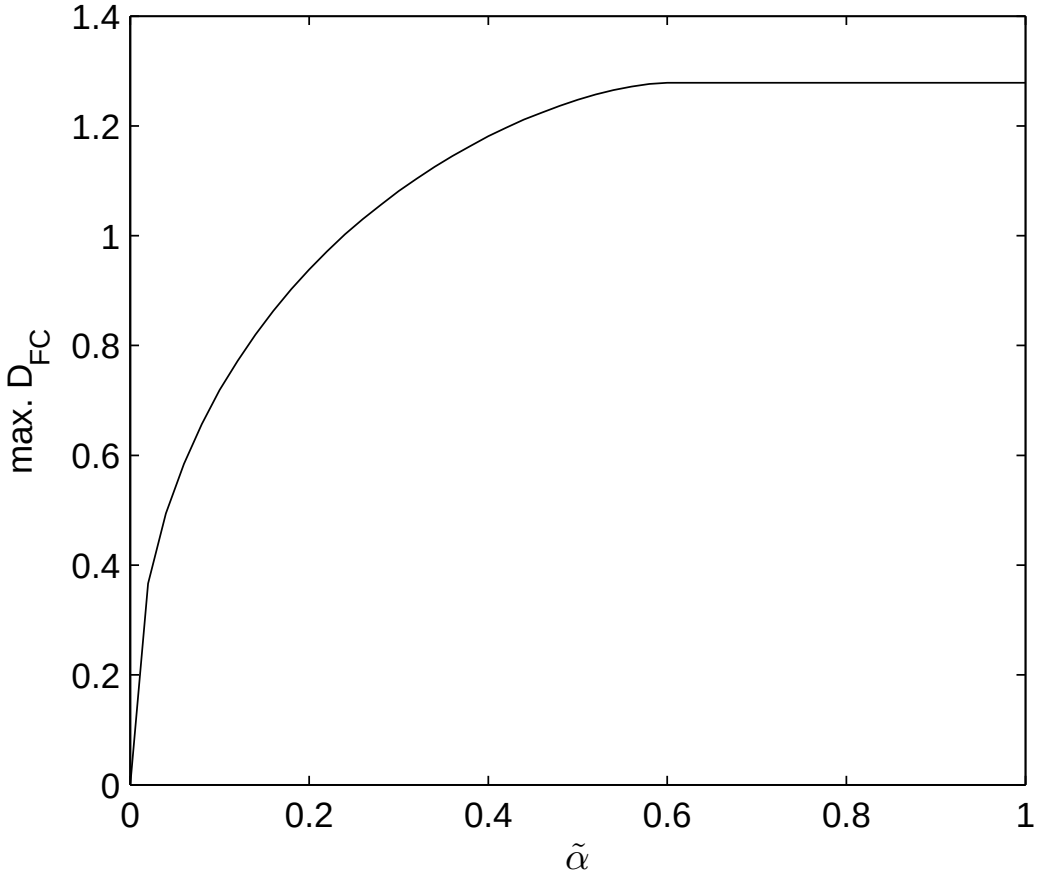
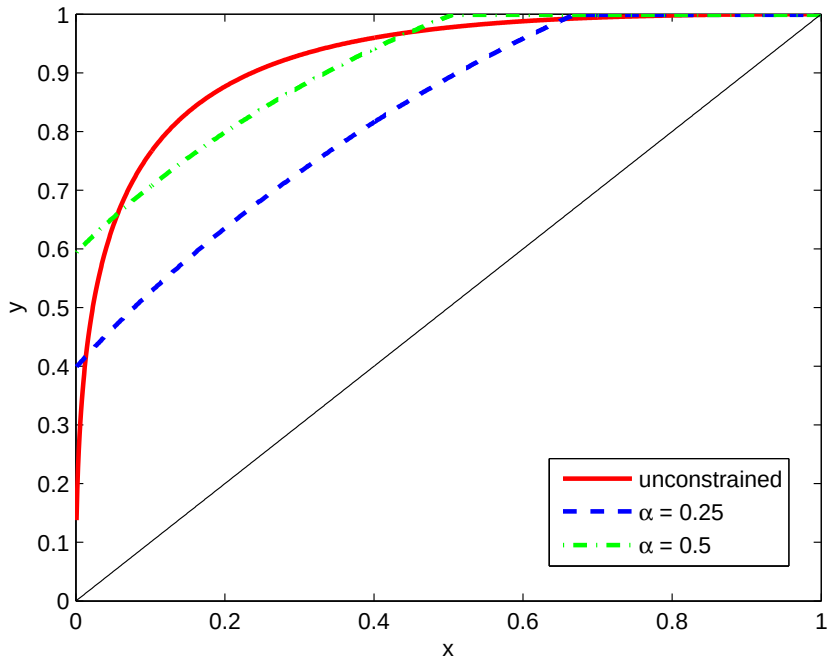
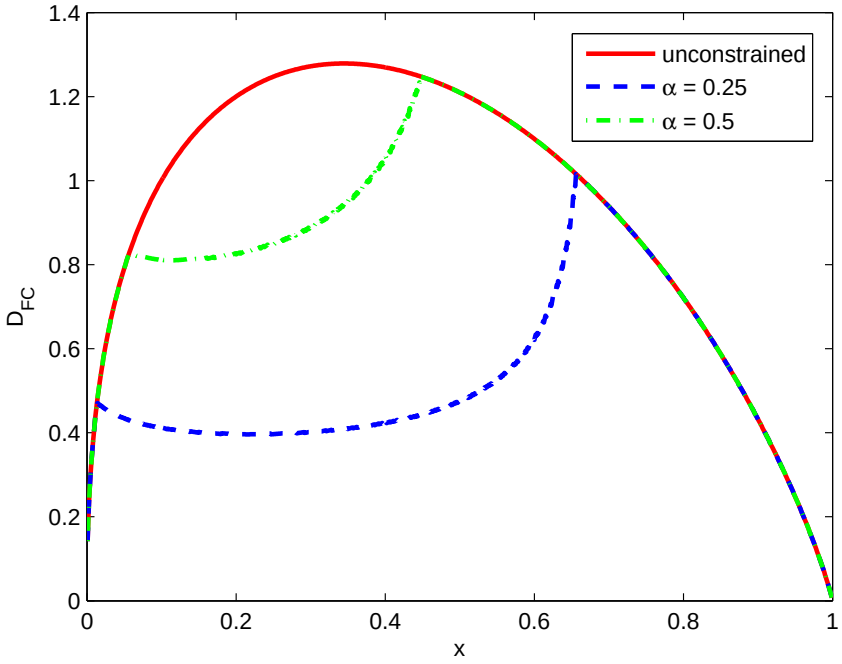


Figure 2.7: Tradeoff between maximum D_{FC} and $\tilde{\alpha}$.

point to the right, in this example) is optimal, as shown in Figure 2.8b. Note that the skewness in the ellipses in Figure 2.8b is due to the asymmetry in the KL divergence. Also, as $\tilde{\alpha}$ decreases, D_{FC} becomes deeper and flat-bottomed as a function of x over the Eve's constraint curve $D_E = \tilde{\alpha}$. Another important observation to be made is the fact that the optimal solution in the presence and absence of Eve (red curves) always is on the boundary of the LRT curve, although the thresholds vary depending on the scenario. Since the sufficient test-statistic is the same irrespective of the presence or absence of Eve, the network designer may implement the system in terms of a threshold that can be varied.



(a) Sensor's ROC in the presence of Eve



(b) D_{FC} as a function of x

Figure 2.8: Sensor performance in the presence of a constraint, $D_E \leq \tilde{\alpha}$, where $\rho_e = 0.1$.

In practice, there exist many conditional probability distributions $p_0(r)$ and $p_1(r)$ for which the computation of likelihood-ratios is intractable. Also, there may be situations where these distributions are not even known to the network designer. In both these cases, the network designer may choose to employ a tractable test that is not LRT.

Let Λ be the test-statistic employed in the sensor quantizer γ, as defined in Equation (2.1). Note that, by allowing randomization (linear stochastic combination of operating points) between quantizers, Carathèodary's theorem [54] and Lemma 2.2 in Section 2.3 together makes every operating point (x, y) inside the set $\Psi \triangleq \mathrm{conv}\left(\{(x, y) \mid y \leq g_\Lambda(x)\}\right)$ feasible, where $\mathrm{conv}(\mathscr{S})$ represents the convex-hull of a given set $\mathscr{S}$.

Since Ψ_i is convex, all of our arguments presented in Section 2.5.1 also hold for the case of any general non-LRT quantizer. We summarize this in the following claim:

Claim 2.3. *Given any ROC curve $y = g_\Lambda(x)$ based on a test-statistic Λ, the optimal quantizer that maximizes the FC's KL Divergence D_{FC} in the presence of a constraint on Eve's KL Divergence $D_E = \tilde{\alpha}$ within the set $\tilde{\Psi}_i \triangleq \mathrm{conv}\{(x, y) \mid y \leq g_\Lambda(x)\}$ always lies on the boundary of $\tilde{\Psi}_i$.*

As discussed earlier in this subsection, this optimal operating point can be implemented by randomizing over a finite set of quantizers, all defined using the same test statistic Λ.

2.5.2 Non-Identical Sensors and Channels

In Section 2.5.1, we investigated the case of identical sensors and channels which was similar to the case of designing the quantizer at a single sensor. In this section, we investigate Problem 2.3 when the network has non-identical sensors and/or has non-identical channels. Since Problem 2.3 is NP-Hard in general, we propose an efficient methodology for quantizer design that satisfies the Eve's constraint $\mathscr{D}_E \leq \alpha$.

Note that the objective function $\mathscr{D}_{FC}$ is linearly separable since the sensor observations

are conditionally independent. Therefore, we define

$$\Phi_n = \Phi_{n-1} + D_{FC_n},\ \forall\, n = 2, \cdots, N. \tag{2.51}$$

where $\Phi_1 = D_{FC_1}$. If, at any given intermediate stage, if Φ_{n-1} is a constant, then the problem of maximizing Ψ_n reduces to the problem of maximizing D_{FC_n}.

This above property of KL Divergence at the FC motivates us to employ dynamic programming [8] to decompose Problem 2.3 into N sub-problems by breaking down the Eve's constraint parameter α into $\boldsymbol{\alpha} = \{\alpha_1, \cdots, \alpha_N\}$ using a *greedy* algorithm. Here, for the sake of ensuring the feasibility of our solution, we assume the following.

$$\sum_{i=1}^{N} \alpha_i \leq \alpha.$$

Therefore, for a given $\boldsymbol{\alpha}$, Problem 2.3 becomes:

Problem 2.5. *For every $i = 1, \cdots, N$, find*

$$\begin{aligned} \arg\max_{\boldsymbol{\gamma}} \quad & D_{FC_i} \quad \textit{s.t.} \\ & \textit{1.} \quad D_{E_i} \leq \alpha_i \\ & \textit{2.} \quad (x_i, y_i) \in \Gamma_i, \textit{ for all } i = 1, \cdots, N. \end{aligned}$$

Note that the performance of this proposed design-methodology completely depends on the choice of $\boldsymbol{\alpha} = \{\alpha_1, \cdots, \alpha_N\}$. To be more precise, the exact solution to Problem 2.3 can be equivalently expressed in terms of an optimal decomposition of α into $\boldsymbol{\alpha} = \{\alpha_1, \cdots, \alpha_N\}$. Since the problem of finding optimal $\boldsymbol{\alpha}$ is intractable, we present a suboptimal (greedy) algorithm to find an efficient decomposition of α as follows.

Let $D^*_{FC_i}$ denote the maximum KL Divergence achievable at the FC, due to the i^{th} sensor. In such a setting, Eve attains a KL Divergence $D^*_{E_i}$ due to the i^{th} sensor. We define the

quality[6] of the FC's and the Eve's channels corresponding to the i^{th} sensor as $k_i = \frac{D^*_{FC_i}}{D^*_{E_i}}$. The quality k_i represents the tradeoff between the detection performance and secrecy. Let the sensors be ordered in terms of the increasing quality as $k_{i_1} \geq \cdots \geq k_{i_N}$. In other words, we obtain the best tradeoff in terms of the sensor quality by considering sensors in the order of decreasing quality in our sequential allocation mechanism. Therefore, we propose a greedy decomposition of Problem 2.5 into N sequential problems based on the sensors' quality, where $\boldsymbol{\alpha} = \{\alpha_1, \cdots, \alpha_N\}$ is chosen such that $\mathscr{D}_{FC}$ is maximized in the presence of Eve's constraint $\mathscr{D}_E \leq \alpha$. Note that this decoupling of α into $\boldsymbol{\alpha}$ allows us to solve each of the individual problems in Problem 2.5 using the same method as presented in Section 2.5.1.

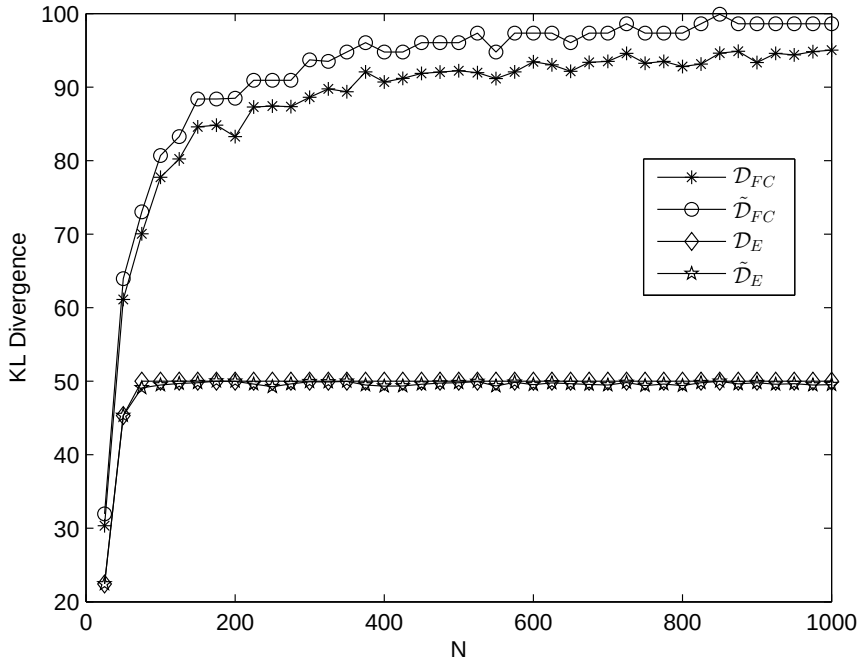
Having ordered the nodes in terms of decreasing k_i^*, we know that node i achieves better tradeoff than node j, if $i > j$. This allows us to select nodes with lower indices to achieve the best tradeoffs between detection performance and secrecy until the resource (constraint on Eve, α) is completely utilized. Therefore, the decomposition of $\mathscr{D}_{FC}$, as shown in Equation (2.51), allows us to sequentially select the individual sensors in an increasing order of indices. Therefore, for index $i = 1$, we allocate $\alpha_1 = D^*_{E_1}$ if $\alpha \geq D^*_{E_1}$. Otherwise, $\alpha_1 = \alpha$. Having allocated the Eve's constraint to Sensor 1, we move to Sensor 2. Now, the remaining tolerable leakage information at the Eve is given by $[\alpha - D^*_{E_1}]_+$, where $[x]_+ = x$ if $x \geq 0$, or, 0 otherwise. Therefore, we solve the problem at Sensor 2 with a new constraint $[\alpha - D^*_{E_1}]_+$.

As the process of selecting the nodes progresses, we reach a point where N^* sensors are already selected and the remaining resource left, given by $\alpha - \sum_{i=1}^{N^*} D^*_{E_i}$, is less than $D_{E_{N^*+1}}$. Therefore, we let $\alpha_{N^*+1} = \alpha - \sum_{i=1}^{N^*} D^*_{E_i}$ and let the remaining sensors sleep in order to satisfy the secrecy constraint.

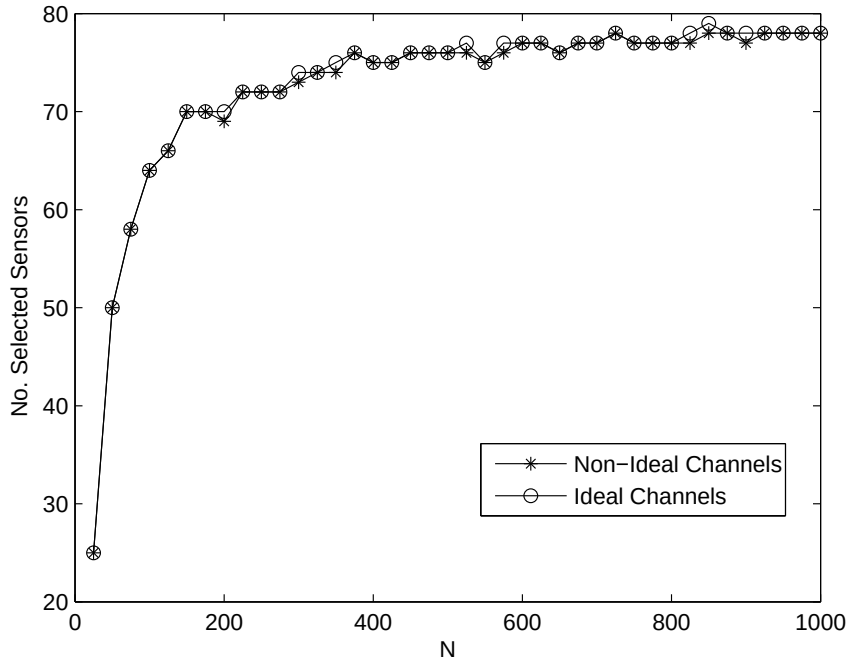
[6]Note that this definition for k_i is one possible heuristic. Another potential heuristic is to define k_i as the difference $D^*_{FC_i} - D^*_{E_i}$, for which we will investigate the network performance in our future work.

Numerical Results

In order to illustrate the performance of the proposed algorithm, we consider a simple example where, for each $i = 1, \cdots, N$, the i^{th} sensor's observation follows $\mathscr{N}(0, \sigma^2)$ under hypothesis H_0 and $\mathscr{N}(\mu_i, \sigma^2)$ under hypothesis H_1. Note that this example demonstrates a scenario where the signal source is spaced at different distances from different sensors in the network, and the sensor observations are modelled using a path-loss attenuation channel model. In such a case, the detection probability at the i^{th} sensor can be defined as $y_i = Q\left(Q^{-1}(x) - \eta_i\right)$ in terms of the false alarm probability x_i, where $\eta_i = \frac{\mu_i}{\sigma}$ is the corresponding SNR. Assuming that the FC has a perfect channel ($\rho_{fc_i} = 0$), while the Eve has a binary symmetric channel with transition probability $\rho_{e_i} = \rho_i$ at the i^{th} sensor, we have $x_{fc_i} = x_i$, $y_{fc_i} = y_i$, $x_{e_i} = \rho_i + (1 - 2\rho_i)x_i$ and $y_{e_i} = \rho_i + (1 - 2\rho_i)y_i$. Then, the KL divergences at the FC and Eve are computed as shown in Equation (2.7).



(a) KLD vs. N



(b) Number of Active Sensors vs. N

Figure 2.9: Performance of the Proposed Greedy Algorithm in a Distributed Inference Network when $\alpha = 50$.

For the sake of illustration, we consider a specific example in order to demonstrate the performance of the proposed greedy algorithm. We assume that all the sensors have identical sensing channels by letting $\eta_i = 1$, for all $i = 1, \cdots, N$. The transition probabilities of the BSCs between the sensors and the FC are sampled randomly from a uniform distribution $\mathscr{U}(0, 0.01)$. Similarly, we let the Eve's channels' transition probabilities be sampled randomly from a uniform distribution $\mathscr{U}(0, 0.1)$. We present a single run of our simulation results in Figure 2.9, where we present both the KL Divergence at the FC and Eve, along with the number of sensors selected in the network, as a function of N when $\alpha = 50$. Note that, for $\alpha = 50$, the difference between the KL divergences between the FC and Eve is about 40 units. We also provide an upper bound on this difference using a benchmark comparison where we present the case where the FC has ideal channels. In the case where FC has ideal channels, the KL Divergences at the FC and Eve are denoted as $\tilde{\mathscr{D}}_{FC}$ and $\tilde{\mathscr{D}}_{E}$ respectively. Although the FC's KL divergence is always lower-bounded by Eve's KL divergence, the difference in the KL Divergences at the FC and Eve depend on the quality of the channels at both FC and Eve.

Also, note that, in Figure 2.9a, as the number of sensors increases, both $\mathscr{D}_{FC}$ and $\mathscr{D}_{E}$ monotonically increase until N reaches a critical point where $\mathscr{D}_{E} = \alpha$. Beyond this critical point, the algorithm starts to select only those sensors that are prioritized according to the decreasing order of k_i. Furthermore, in Figure 2.9b, the number of selected sensors increases with increasing number of sensors in the network at the similar rate as that of $\mathscr{D}_{FC}$. Lastly, note that the performance of the distributed inference network in terms of KL Divergence saturates as N increases as per intuition.

2.6 Summary

In summary, we have considered two secrecy frameworks, namely differential and constrained secrecy, in a distributed detection network when all the communication channels

are binary symmetric channels. In the case of differential secrecy, we have shown that the structure of optimal quantizer at any sensor is either LRT-based, or uninformative. In the case of constrained secrecy, we have proved that the optimal quantizer is always LRT-based in the presence of identical sensors and channels. We have presented an algorithm to find optimal LRT thresholds, and presented numerical results to illustrate the performance of our network design. In the case of non-identical sensors and channels, we have proposed an efficient design for sensor quantizers by decomposing the original problem in N sub-problems using a dynamic programming approach. Numerical results have been presented to illustrate the efficiency of our proposed design under different scenarios.

CHAPTER 3
SECRECY IN CENTRALIZED DETECTION: TRANSMIT-DIVERSITY

Security in detection networks is a well-studied research topic in the past literature, in which several mitigation techniques such as stochastic encryption, optimal design of local detectors and so on, have been proposed to mitigate information leakage to eavesdroppers. While secrecy has been addressed in distributed detection networks in the past, as discussed in Chapter 2, the problem still remains open in the context of centralized detection networks. In this chapter, we consider the problem of designing a secure centralized detection network in the presence of tolerable secrecy constraints in this chapter.

We propose a transmit-diversity mechanism in a centralized detection network where the sensors construct transmission signals by combining artificial noise with the amplified observation so as to maximize the KL Divergence at the FC in the presence of a tolerable constraint on Eve's KL Divergence. While the amplify-forward mechanism is designed to increase the detection performance at the FC, the artificial noise is chosen in such a way that the Eve's performance is severely affected. In this chapter, we derive efficient sensor transmission policies in the proposed framework for detection networks by solving a non-convex optimization problem approximately using a two-stage algorithm. In the first stage,

our proposed algorithm decomposes the original problem into N sub-problems, where N is the number of sensing units in the network. In the second stage, each sub-problem is solved by relaxing it into a semidefinite program (SDP). In our simulation results, we show that the FC's KL Divergence increases as the number of sensor antennas increases. Furthermore, with enough number of sensor antennas, we show that the FC can always attain a greater KL Divergence than that of Eve by employing our proposed approach.

3.1 Literature Survey

Transmit-diversity mechanisms in multiple-input multiple-output (MIMO) systems have gained a lot of attention of many researchers over the last decade in the context of physical-layer security of communication systems. For a detailed account on this literature, the reader may refer to [40] for an in-depth survey on MIMO communication, and [28, 30] for an in-depth survey on MIMO detection. In this vast literature on MIMO systems, the most relevant framework to this work is the design of secure relay networks[1] with MIMO/beamforming capabilities, which are surveyed in great detail in [40], [70] and [19]. In these works, communication metrics such as Shannon's equivocation rate were chosen as design-objectives.

In contrast to securing relay networks and other communication networks using traditional MIMO/beamforming methods, we design an optimal transmission mechanism over multiple antennas at the sensors that employ amplify-and-forward transmission of observations, while simultaneously injecting artificial noise at the sensors in order to reduce information-leakage to the Eve. This work is inspired by the work carried out by Goel and Negi in [18, 45], where they proposed a MIMO-based scheme for point-to-point communication links to mitigate information-transfer to the Eve by adding artificial noise in the nulls of the beam.

[1]Relay networks are similar to parallel-topology detection networks in terms of their functioning and architecture, which are traditionally designed to optimize Shannon's information rate between the end-users.

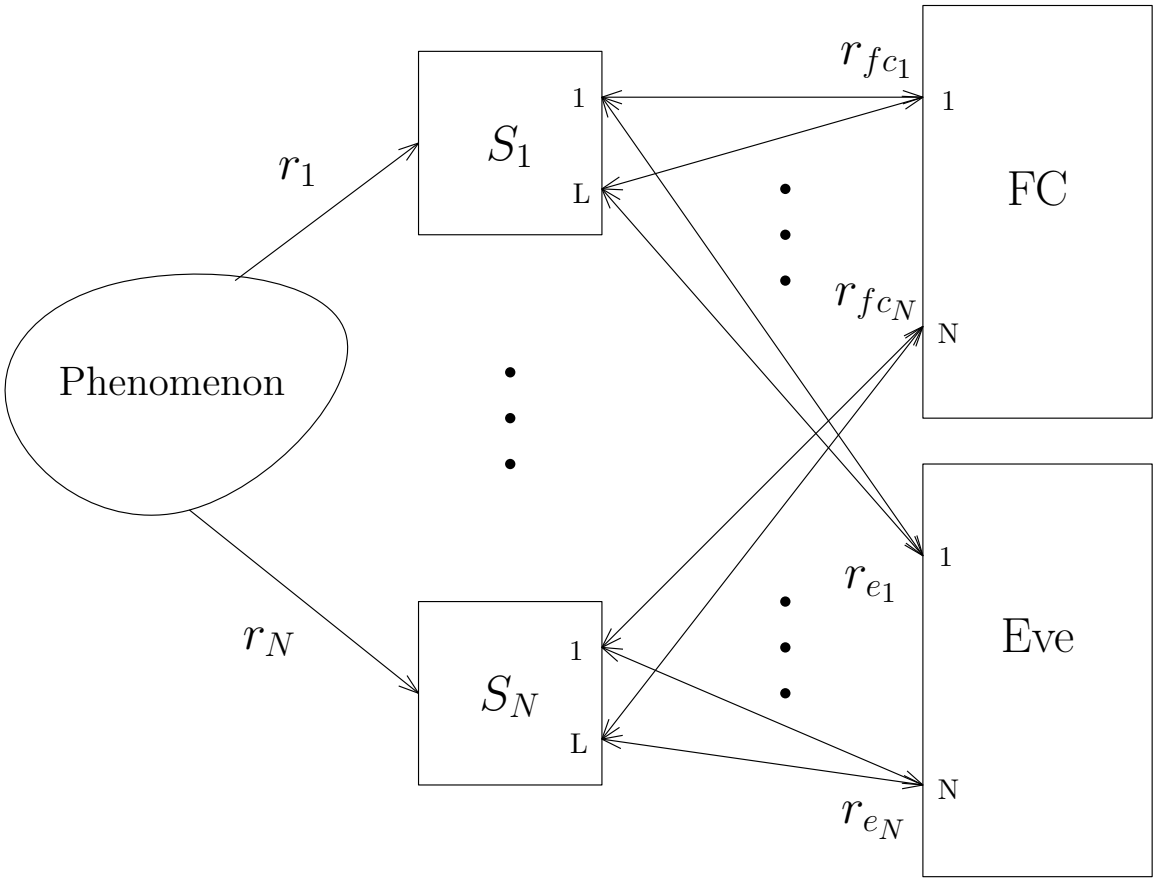


Figure 3.1: Centralized Inference Network in the Presence of an Eavesdropper

3.2 System Model and Problem Statement

Consider a detection network of N sensors, as shown in Figure 3.1, communicating with the FC through a parallel-topology of wireless links. Let the binary hypotheses H_1 and H_0 denote the presence and absence of a phenomenon-of-interest (PoI), with prior probabilities π_1 and π_0 respectively. Let r_i be the observation made by the i^{th} sensor, whose signal model is given as follows.

$$r_i = \begin{cases} n_i; & \text{under } H_0 \\ \theta + n_i; & \text{under } H_1 \end{cases} \tag{3.1}$$

where n_i is a zero-mean AWGN noise with variance σ_s^2, and θ is a known real-valued PoI signal, i.e., $\theta \in \mathscr{R}$.

Let each sensor be equipped with a transmitting array of L antennas, while the FC

receives each of the sensor's transmissions using a single antenna for each parallel channel. The i^{th} sensor processes the received signal r_i and generates $\mathbf{s}_i = [s_{i1}, \cdots, s_{iL}]$, and each signal s_{ik} is transmitted through the k^{th} antenna respectively, and is chosen as follows.

$$s_{ik} = r_i x_{ik} + w_{ik} \tag{3.2}$$

where x_{ik} is the weight of the i^{th} sensor's observation r_i at the k^{th} antenna, and $w_{ik} \sim \mathcal{N}(0, y_{ik}^2)$ is the artificial noise added independently to the k^{th} antenna at the i^{th} sensor. In other words, y_{ik} is the root-mean square (RMS) power of the artificial noise w_{ik}.

Equation (3.2) can be written in a vector form, as follows:

$$\mathbf{s}_i = r_i \mathbf{x}_i + \mathbf{w}_i \tag{3.3}$$

where $\mathbf{s}_i$, $\mathbf{x}_i$ and $\mathbf{w}_i$ are $L \times 1$ vectors for all $i = 1, \cdots, N$.

In this chapter, we assume that the i^{th} sensor transmits the signals $\mathbf{s}_i$ in one-shot using all its L transmitting antennas. In practice, a sensing unit has a limited total energy available for its transmission. This constraint on the total energy consumed to transmit $\mathbf{s}_i = \{s_{i1}, \cdots, s_{iL}\}$ at the i^{th} sensor is given as follows.

$$E\left(\mathbf{s}_i^T \mathbf{s}_i\right) \leq \mathscr{E}. \tag{3.4}$$

Expanding and simplifying Equation (3.4), we have

$$\left(\sigma_s^2 + \pi_1 \theta^2\right) \mathbf{x}_i^T \mathbf{x}_i + \mathbf{y}_i^T \mathbf{y}_i \leq \mathscr{E} \tag{3.5}$$

The sensors transmit their respective messages over N dedicated orthogonal channels, such that the FC and Eve receive N signals (one antenna per channel), which are denoted as $\{r_{fc_1}, \cdots, r_{fc_N}\}$ and $\{r_{e_1}, \cdots, r_{e_N}\}$ respectively. We model these received signals at

the FC and Eve respectively, as follows.

$$r_{fc_i} = \sum_{k=1}^{L} h_{ik} s_{ik} + n_{fc_i}, \tag{3.6a}$$

$$r_{e_i} = \sum_{k=1}^{L} g_{ik} s_{ik} + n_{e_i}, \tag{3.6b}$$

where h_{ik} and g_{ik} are the channel-gains between the k^{th} antenna at the i^{th} sensor and the i^{th} receiving antenna at the FC and Eve respectively, and, $n_{fc_i} \sim \mathcal{N}(0, \sigma_{fc}^2)$ and $n_{e_i} \sim \mathcal{N}(0, \sigma_e^2)$ are AWGN noises at the FC and Eve respectively.

For the sake of notational simplicity, let $\mathbf{h}_i = \{h_{i1}, \cdots, h_{iL}\}$ and $\mathbf{g}_i = \{g_{i1}, \cdots, g_{iL}\}$ denote the channel-gain vectors at the FC and Eve respectively, corresponding to the i^{th} sensor. Let $\mathbb{I}_N$ denote a $N \times N$ identity matrix. Then, Equations (3.6a) and (3.6b) can be rewritten as

$$\mathbf{r}_{fc} = \mathbb{R} \begin{bmatrix} \mathbf{h}_1^T \mathbf{x}_1 \\ \vdots \\ \mathbf{h}_N^T \mathbf{x}_N \end{bmatrix} + \begin{bmatrix} \mathbf{h}_1^T \mathbf{w}_1 \\ \vdots \\ \mathbf{h}_1^T \mathbf{w}_1 \end{bmatrix} + \mathbf{n}_{fc} \tag{3.7a}$$

$$\mathbf{r}_e = \mathbb{R} \begin{bmatrix} \mathbf{g}_1^T \mathbf{x}_1 \\ \vdots \\ \mathbf{g}_N^T \mathbf{x}_N \end{bmatrix} + \begin{bmatrix} \mathbf{g}_1^T \mathbf{w}_1 \\ \vdots \\ \mathbf{g}_N^T \mathbf{w}_N \end{bmatrix} + \mathbf{n}_e \tag{3.7b}$$

where $\mathbb{R}$ is a $N \times N$ diagonal matrix with r_i being the i^{th} diagonal entry, and, $\mathbf{n}_{fc} \sim \mathcal{N}(\mathbf{0}, \sigma_{fc}^2 \mathbb{I}_N)$ and $\mathbf{n}_e \sim \mathcal{N}(\mathbf{0}, \sigma_e^2 \mathbb{I}_N)$ are additive noise vectors at the FC and Eve respectively. Being linear combinations of conditionally normal random variables, both $\mathbf{r}_{fc}$ and $\mathbf{r}_e$ are also normally distributed when conditioned under any given hypothesis. More specifically, we have

$$\mathbf{r}_{fc}|H_0 \sim \mathcal{N}(\mathbf{0}, \Sigma_{fc}), \quad \mathbf{r}_{fc}|H_1 \sim \mathcal{N}(\boldsymbol{\mu}_{fc}, \Sigma_{fc}),$$

and

$$\mathbf{r}_e|H_0 \sim \mathscr{N}(\mathbf{0}, \Sigma_e), \quad \mathbf{r}_e|H_1 \sim \mathscr{N}(\boldsymbol{\mu}_e, \Sigma_e),$$

where $\boldsymbol{\mu}_{fc}$, $\boldsymbol{\mu}_e$, Σ_{fc} and Σ_e are computed as follows.

$$\boldsymbol{\mu}_{fc} = E(\mathbf{r}_{fc}|H_1) = \theta \begin{bmatrix} \mathbf{h}_1^T\mathbf{x}_1 \\ \vdots \\ \mathbf{h}_N^T\mathbf{x}_N \end{bmatrix}, \tag{3.8a}$$

$$\boldsymbol{\mu}_e = E(\mathbf{r}_e|H_1) = \theta \begin{bmatrix} \mathbf{g}_1^T\mathbf{x}_1 \\ \vdots \\ \mathbf{g}_N^T\mathbf{x}_N \end{bmatrix}, \tag{3.8b}$$

$$\begin{aligned} \Sigma_{fc} &= E[\mathbf{r}_{fc}\mathbf{r}_{fc}^T|H_0] = E[(\mathbf{r}_{fc} - \boldsymbol{\mu}_{fc})(\mathbf{r}_{fc} - \boldsymbol{\mu}_{fc})^T|H_1] \\ &= \sigma_s^2 \begin{bmatrix} \mathbf{x}_1^T\mathbb{H}_{11}\mathbf{x}_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \mathbf{x}_N^T\mathbb{H}_{N1}\mathbf{x}_N \end{bmatrix} \\ &\quad + \begin{bmatrix} \mathbf{y}_1^T\mathbb{H}_{12}\mathbf{y}_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \mathbf{y}_N^T\mathbb{H}_{N2}\mathbf{y}_N \end{bmatrix} + \sigma_{fc}^2\mathbb{I}, \end{aligned} \tag{3.8c}$$

and

$$\begin{aligned}\Sigma_e &= E[\mathbf{r}_{fc}\mathbf{r}_{fc}^T|H_0] = E[(\mathbf{r}_{fc}-\boldsymbol{\mu}_{fc})(\mathbf{r}_{fc}-\boldsymbol{\mu}_{fc})^T|H_1] \\ &= \sigma_s^2 \begin{bmatrix} \mathbf{x}_1^T\mathbb{G}_{11}\mathbf{x}_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \mathbf{x}_N^T\mathbb{G}_{N1}\mathbf{x}_N \end{bmatrix} \\ &\quad + \begin{bmatrix} \mathbf{y}_1^T\mathbb{G}_{12}\mathbf{y}_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \mathbf{y}_N^T\mathbb{G}_{N2}\mathbf{y}_N \end{bmatrix} + \sigma_e^2\mathbb{I}.\end{aligned} \tag{3.8d}$$

where, for all $i = 1, \cdots, N$,

$$\begin{aligned}\mathbb{H}_{i1} &= \mathbf{h}_i\mathbf{h}_i^T, & \mathbb{G}_{i1} &= \mathbf{h}_i\mathbf{h}_i^T \\ \mathbb{H}_{i2} &= [\text{diag}(\mathbf{h}_i)]^2, & \mathbb{G}_{i2} &= [\text{diag}(\mathbf{g}_i)]^2.\end{aligned} \tag{3.9}$$

In this chapter, we choose KL Divergence as the performance metric at both FC and Eve, denoted by $\mathscr{D}_{fc}$ and $\mathscr{D}_e$ at the FC and Eve respectively, for the sake of tractability. Note that KL Divergence is the error exponent for the global miss probability in the Neyman Pearson framework, where the global false alarm probability is constrained to a fixed tolerable amount. Therefore, $\mathscr{D}_{fc}$ and $\mathscr{D}_e$ act as a surrogate to the global error probability at both the FC and Eve respectively.

We compute both $\mathscr{D}_{fc}$ and $\mathscr{D}_e$ of the received signals $\mathbf{r}_{fc}$ and $\mathbf{r}_e$ respectively, as follows.

$$\begin{aligned}
\mathscr{D}_{fc} &= E_{H_0}\left[\log\frac{p(\mathbf{r}_{fc}|H_0)}{p(\mathbf{r}_{fc}|H_1)}\right] \\
&= \frac{1}{2}\boldsymbol{\mu}_{fc}^T\Sigma_{fc}^{-1}\boldsymbol{\mu}_{fc} \qquad (3.10a) \\
&= \frac{\theta^2}{2}\sum_{i=1}^{N}\left[\frac{\mathbf{x}_i^T\mathbb{H}_{i1}\mathbf{x}_i}{\sigma_s^2\mathbf{x}_i^T\mathbb{H}_{i1}\mathbf{x}_i+\mathbf{y}_i^T\mathbb{H}_{i2}\mathbf{y}_i+\sigma_{fc}^2}\right] \\
\mathscr{D}_{e} &= E_{H_0}\left[\log\frac{p(\mathbf{r}_{e}|H_0)}{p(\mathbf{r}_{e}|H_1)}\right] \\
&= \frac{1}{2}\boldsymbol{\mu}_{e}^T\Sigma_{e}^{-1}\boldsymbol{\mu}_{e} \qquad (3.10b) \\
&= \frac{\theta^2}{2}\sum_{i=1}^{N}\left[\frac{\mathbf{x}_i^T\mathbb{G}_{i1}\mathbf{x}_i}{\sigma_s^2\mathbf{x}_i^T\mathbb{G}_{i1}\mathbf{x}_i+\mathbf{y}_i^T\mathbb{G}_{i2}\mathbf{y}_i+\sigma_{fc}^2}\right]
\end{aligned}$$

Our goal is to design a secure detection network that maximizes the KL Divergence at the FC, while constraining the Eve's KL Divergence to a fixed value α, in the presence of an energy constraint at each sensor. This is formally stated as follows.

Problem 3.1.

$$\begin{aligned}
&\underset{\{\mathbf{x}_i,\mathbf{y}_i\}_{i=1,\cdots,N}}{\textit{maximize}} && \mathscr{D}_{fc} \\
&\textit{subject to} && \textit{1.}\ \ \mathscr{D}_e \le \alpha \\
& && \textit{2.}\ \ \left(\sigma_s^2+\pi_1\theta^2\right)\mathbf{x}_i^T\mathbf{x}_i+\mathbf{y}_i^T\mathbf{y}_i \le \mathscr{E}, \\
& && \quad \textit{for all}\ i=1,2,\ldots,N.
\end{aligned}$$

For the sake of notational simplicity, let us assume

$$\mathbf{z}_i = \begin{bmatrix} \mathbf{x}_i \\ \\ \mathbf{y}_i \end{bmatrix}.$$

Then, Equations (3.10a) and (3.10b) can be rewritten as follows.

$$\mathscr{D}_{fc} = \frac{\theta^2}{2\sigma_s^2} \sum_{i=1}^{N} \left[\frac{\mathbf{z}_i^T \mathbb{A}_i \mathbf{z}_i}{\mathbf{z}_i^T \mathbb{B}_i \mathbf{z}_i + c} \right] \tag{3.11a}$$

$$\mathscr{D}_{e} = \frac{\theta^2}{2\sigma_s^2} \sum_{i=1}^{N} \left[\frac{\mathbf{z}_i^T \mathbb{C}_i \mathbf{z}_i}{\mathbf{z}_i^T \mathbb{D}_i \mathbf{z}_i + e} \right] \tag{3.11b}$$

where

$$\mathbb{A}_i = \begin{bmatrix} \mathbb{H}_{i1} & \mathbb{0} \\ \\ \mathbb{0} & \mathbb{0} \end{bmatrix}, \qquad \mathbb{B}_i = \begin{bmatrix} \mathbb{H}_{i1} & \mathbb{0} \\ \\ \mathbb{0} & \dfrac{1}{\sigma_s^2}\mathbb{H}_{i2} \end{bmatrix},$$

$$\mathbb{C}_i = \begin{bmatrix} \mathbb{G}_{i1} & \mathbb{0} \\ \\ \mathbb{0} & \mathbb{0} \end{bmatrix}, \qquad \mathbb{D}_i = \begin{bmatrix} \mathbb{G}_{i1} & \mathbb{0} \\ \\ \mathbb{0} & \dfrac{1}{\sigma_s^2}\mathbb{G}_{i2} \end{bmatrix},$$

$$c = \frac{\sigma_{fc}^2}{\sigma_s^2} \qquad \text{and} \qquad e = \frac{\sigma_e^2}{\sigma_s^2}.$$

Note that ignoring the constant $\dfrac{\theta^2}{2\sigma_s^2}$ in the objective function does not affect the optimal solution of Problem 3.1. Let

$$\beta = \frac{2\sigma_s^2}{\theta^2}\alpha \quad \text{and} \quad \mathbb{E} = \begin{bmatrix} (\sigma_s^2 + \pi_1\theta^2)\,\mathbb{I}_L & \mathbb{0} \\ \\ \mathbb{0} & \mathbb{I}_L \end{bmatrix}.$$

Then, Problem 3.1 can be equivalently written as follows.

Problem 3.2.

$$\begin{aligned} \underset{\{\mathbf{z}_1,\cdots,\mathbf{z}_N\}}{\textit{maximize}} \quad & \sum_{i=1}^{N}\left[\frac{\mathbf{z}_i^T\mathbb{A}_i\mathbf{z}_i}{\mathbf{z}_i^T\mathbb{B}_i\mathbf{z}_i+c}\right] \\ \textit{subject to} \quad & \textit{1.} \ \sum_{i=1}^{N}\left[\frac{\mathbf{z}_i^T\mathbb{C}_i\mathbf{z}_i}{\mathbf{z}_i^T\mathbb{D}_i\mathbf{z}_i+e}\right] \leq \beta \\ & \textit{2.} \ \mathbf{z}_i^T\mathbb{E}\mathbf{z}_i \leq \mathscr{E}, \\ & \quad \textit{for all } i=1,2,\ldots,N. \end{aligned}$$

Note that Problem 3.1 is a hard, non-convex problem. In fact, the problem is non-convex even if $N = 1$. Therefore, we propose two approximate solutions to Problem 3.2 via employing semidefinite relaxation and convex-concave restriction.

3.3 Approximation via Semidefinite Relaxation

Without any loss of generality, we introduce N slack variables $\beta_1, \cdots, \beta_N$ by assuming

$$\frac{\mathbf{z}_i^T\mathbb{C}_i\mathbf{z}_i}{\mathbf{z}_i^T\mathbb{D}_i\mathbf{z}_i+e} \leq \beta_i, \ \forall\, i=1,\cdots,N. \tag{3.12}$$

In other words, β_i is an upper bound on the contribution of the i^{th} sensor to the Eve's KL Divergence. Substituting Equation (3.12) in Problem 3.2, we have the following equivalent problem statement.

Problem 3.3.

$$\underset{\{\mathbf{z}_1,\cdots,\mathbf{z}_N\},\{\beta_1,\cdots,\beta_N\}}{\textit{maximize}} \quad \sum_{i=1}^{N}\left[\frac{\mathbf{z}_i^T\mathbb{A}_i\mathbf{z}_i}{\mathbf{z}_i^T\mathbb{B}_i\mathbf{z}_i+c}\right]$$

$$\textit{subject to} \qquad 1.\ \frac{\mathbf{z}_i^T\mathbb{C}_i\mathbf{z}_i}{\mathbf{z}_i^T\mathbb{D}_i\mathbf{z}_i+e}\leq\beta_i,$$

$$\textit{for all } i=1,2,\ldots,N$$

$$2.\ \sum_{i=1}^{N}\beta_i\leq\beta,$$

$$3.\ \mathbf{z}_i^T\mathbb{E}\mathbf{z}_i\leq\mathscr{E},$$

$$\textit{for all } i=1,2,\ldots,N.$$

Note that, if the optimal decomposition $\boldsymbol{\beta}^*=\{\beta_1^*,\cdots,\beta_N^*\}$ are known beforehand, the above problem can be decomposed into N independent problems without any loss of optimality, as shown below in Problem 3.4, because of two reasons:

- The objective function is linearly separable.
- Constraint 2 is the only coupling condition in Problem 3.3, which is itself a linearly separable function.

Problem 3.4. *For a given β_i (which is chosen such that $\sum_{i=1}^{N}\beta_i\leq\beta$),*

$$\underset{\mathbf{z}_i}{\textit{maximize}} \quad \frac{\mathbf{z}_i^T\mathbb{A}_i\mathbf{z}_i}{\mathbf{z}_i^T\mathbb{B}_i\mathbf{z}_i+c}$$

$$\textit{subject to} \quad 1.\ \frac{\mathbf{z}_i^T\mathbb{C}_i\mathbf{z}_i}{\mathbf{z}_i^T\mathbb{D}_i\mathbf{z}_i+e}\leq\beta_i,$$

$$2.\ \mathbf{z}_i^T\mathbb{E}\mathbf{z}_i\leq\mathscr{E}.$$

Before we determine $\{\beta_1, \cdots, \beta_N\}$ and the corresponding approximate solution, let us consider the problem of optimal network design in the absence of an Eve's constraint. This is formally stated in Problem 3.5.

Problem 3.5.

$$\underset{\mathbf{z}_i}{\textit{maximize}} \quad \frac{\mathbf{z}_i^T \mathbb{A}_i \mathbf{z}_i}{\mathbf{z}_i^T \mathbb{B}_i \mathbf{z}_i + c}$$

$$\textit{subject to} \quad 1. \;\; \mathbf{z}_i^T \mathbb{E} \mathbf{z}_i \leq \mathscr{E}.$$

Let $\tilde{D}_{FC_i}$ and $\tilde{\mathbf{z}}_i$ denote the optimal solution to Problem 3.5. Then, we have

$$\tilde{D}_{FC_i} = \lambda_{max}\left(\mathbb{A}_i, \mathbb{B}_i + \frac{c\mathbb{E}}{\mathscr{E}}\right) \tag{3.13a}$$

$$\tilde{\mathbf{z}}_i = \boldsymbol{\eta}\sqrt{\frac{\mathscr{E}}{\boldsymbol{\eta}^T \mathbb{E} \boldsymbol{\eta}}} \tag{3.13b}$$

where $\lambda_{max}(\Lambda_1, \Lambda_2)$ is the generalized eigenvalue[2] of the two given matrices Λ_1 and Λ_2, and $\boldsymbol{\eta}$ is the generalized eigenvector corresponding to the aforementioned eigenvalue given in Equation (3.13a). For a detailed account on the computation of generalized eigenvalues and eigenvectors, the reader may refer to [7, 37].

We can compute the KL Divergence $\tilde{D}_{E_i}$ that the Eve attains, due to the i^{th} sensor, as follows.

$$\tilde{D}_{E_i} = \frac{\tilde{\mathbf{z}}_i{}^T \mathbb{C}_i \tilde{\mathbf{z}}_i}{\tilde{\mathbf{z}}_i{}^T \mathbb{D}_i \tilde{\mathbf{z}}_i + e} \tag{3.14}$$

Note that both $\tilde{D}_{FC_i}$ and $\tilde{D}_{E_i}$ are both upper-bounds on the optimal values of D_{FC_i} and D_{E_i} as they are the solutions of the unconstrained problem, which are found by solving Problem 3.1.

In the remaining section, we present our proposed methodology to find an efficient

[2]The generalized eigenvalue ψ and eigenvector $\boldsymbol{\mu}$ of the matrices Λ_1 and Λ_2 satisfy $\Lambda_1 \boldsymbol{\mu} = \psi \Lambda_2 \boldsymbol{\mu}$.

solution to Problem 3.3 by solving it in three stages. In the first stage, we find reasonable values for β_i based on the channel conditions at the FC and Eve using a greedy algorithm. In the second stage, we provide an optimal solution to each of the following decomposed problems using semi-definite relaxation (SDR). In the final stage, we use the above solution to construct an efficient and a feasible solution using randomization techniques.

Stage 1: Efficient Decomposition

In this subsection, we propose a sequential methodology in order to compute the value of $\{\beta_1, \cdots, \beta_N\}$ based on a specific ordering of the sensors. Intuitively, we expect that the choice of this sensor-ordering is based on the combined effect of the sensing observations and the channel models at both the FC and Eve.

Therefore, we first construct a vector $\mathbf{k} = \{k_1, \cdots, k_N\}$ which, when sorted in a decreasing manner, gives the sensor-ordering in terms of their allocation quality. For the sake of tractability, we assume that $\mathbf{k}$ is independent of the choice of $\{\mathbf{z}_1, \cdots, \mathbf{z}_N\}$. With this assumption, we deviate from the optimal solution of Problem 3.3. Note that the construction of the sensor-ordering is key to the efficiency of our solution approach.

In this chapter, we define the ordering metric k_i based on the objective function as follows.

$$k_i = \tilde{D}_{FC_i}. \tag{3.15}$$

Our intuition behind choosing this definition for k_i is that the sensors shall be ordered in terms of their respective contribution to the overall KL Divergence at the FC. Note that, since D_{FC_i} and D_{E_i} have a monotonic relationship, increasing k_i increases the value of β_i. Therefore, we denote the sorted vector of sensor-indices, as $\mathbf{i}_{sort} = \{i_1, \cdots, i_N\}$, which are ordered in a decreasing order of $\mathbf{k}$. In other words, $k_{i_j} \geq k_{i_{j+1}}$ for all $j = 1, \cdots, N-1$.

Let i_j denote the j^{th} entry in $\mathbf{i}_{sort}$. Then, β_{i_j} is given by

$$\beta_{i_j} = \min\left\{\Delta_{j-1}, \tilde{D}_{E_i}\right\}, \tag{3.16}$$

where $\Delta_{j-1} = \beta - \sum_{k=1}^{j-1} \beta_{i_k}$ is the residual value that needs to be allocated to the rest of the sensors with indices beyond the $(j-1)^{th}$ entry in the vector $\mathbf{i}_{sort}$, with its initial value defined as $\Delta_0 = \beta$.

Note that our choice of $\{\beta_1, \cdots, \beta_N\}$ satisfies Constraint 2 in Problem 3.3. Therefore, the proposed solution always lies within the feasible region of Problem 3.3.

Stage 2: Semidefinite Programming

In this subsection, we present a SDR-based methodology to approximately solve Problem 3.4. Note that since Problem 3.4 is completely defined locally at the i^{th} sensor, we remove the index i for notational convenience. In the case of β_i, we replace the notation with δ in order to avoid any confusion. Therefore, we restate Problem 3.4 by removing the index i, as follows.

Problem 3.6.

$$\begin{aligned} \underset{\mathbf{z}}{\textit{maximize}} \quad & \frac{\mathbf{z}^T \mathbb{A} \mathbf{z}}{\mathbf{z}^T \mathbb{B} \mathbf{z} + c} \\ \textit{subject to} \quad & \textit{1.} \ \frac{\mathbf{z}^T \mathbb{C} \mathbf{z}}{\mathbf{z}^T \mathbb{D} \mathbf{z} + e} \leq \delta, \\ & \textit{2.} \ \mathbf{z}^T \mathbb{E} \mathbf{z} \leq \mathscr{E}. \end{aligned} \tag{3.17}$$

Without any loss of generality, let

$$\frac{1}{\mathbf{z}^T \mathbb{B} \mathbf{z} + c} = u^2. \tag{3.18}$$

where u is a non-negative slack variable. Replacing the variable $\mathbf{z}$ with a new variable $\mathbf{t} = u\mathbf{z}$, Problem 3.6 can be equivalently expressed as follows.

Problem 3.7.

$$\begin{aligned} \underset{\mathbf{t},u}{\textit{maximize}} \quad & \mathbf{t}^T \mathbb{A} \mathbf{t} \\ \textit{subject to} \quad & \textit{1.} \ \mathbf{t}^T \left(\mathbb{C} - \delta \mathbb{D} \right) \mathbf{t} \leq e\delta u^2, \\ & \textit{2.} \ \mathbf{t}^T \mathbb{E} \mathbf{t} \leq \mathscr{E} u^2 \\ & \textit{3.} \ \mathbf{t}^T \mathbb{B} \mathbf{t} + c \cdot u^2 = 1 \end{aligned}$$

Note that, since the matrix $\mathbb{A}$ is positive semidefinite, the objective function is convex. Therefore, Constraint 3 can be relaxed into an inequality $\mathbf{t}^T \mathbb{B} \mathbf{t} + c \cdot u^2 \leq 1$ without any loss of generality, since Problem 3.7 is a convex-maximization problem. Furthermore, If $(\mathbf{t}^*, u^*)$ is the optimal solution to Problem 3.7, then the optimal solution is given by $\mathbf{z}^* = \mathbf{t}^* / u^*$.

To simplify further, we define the combined optimization variable $\mathbf{v} = \begin{bmatrix} \mathbf{t} \\ u \end{bmatrix}$ in order to have the following problem.

Problem 3.8.

$$\begin{aligned} \underset{\mathbf{v}}{\textit{maximize}} \quad & \mathbf{v}^T \mathbb{M}_0 \mathbf{v} \\ \textit{subject to} \quad & \textit{1.} \ \mathbf{v}^T \mathbb{M}_1 \mathbf{v} \leq 0, \\ & \textit{2.} \ \mathbf{v}^T \mathbb{M}_2 \mathbf{v} \leq 0 \\ & \textit{3.} \ \mathbf{v}^T \mathbb{M}_3 \mathbf{v} \leq 1 \end{aligned}$$

where

$$\mathbb{M}_0 = \begin{bmatrix} \mathbb{A} & \mathbf{0} \\ \mathbf{0}^T & 0 \end{bmatrix}, \qquad \mathbb{M}_1 = \begin{bmatrix} \mathbb{C} - \delta\mathbb{D} & \mathbf{0} \\ \mathbf{0}^T & -e\delta \end{bmatrix},$$

$$\mathbb{M}_2 = \begin{bmatrix} \mathbb{E} & \mathbf{0} \\ \mathbf{0}^T & -\mathscr{E} \end{bmatrix} \quad \text{and} \quad \mathbb{M}_3 = \begin{bmatrix} \mathbb{B} & \mathbf{0} \\ \mathbf{0}^T & c \end{bmatrix}.$$

Let $\mathbb{V} = \mathbf{v}\mathbf{v}^T$. Note that $\mathbb{V}$ is rank-1 and positive semidefinite. Therefore, Problem 3.8 can be equivalently written in a matrix-form as follows.

Problem 3.9.

$$\begin{aligned} \underset{\mathbb{V}}{\textit{maximize}} \quad & \mathrm{Tr}\left(\mathbb{V}\mathbb{M}_0\right) \\ \textit{subject to} \quad & \textit{1.} \quad \mathrm{Tr}\left(\mathbb{V}\mathbb{M}_1\right) \leq 0, \\ & \textit{2.} \quad \mathrm{Tr}\left(\mathbb{V}\mathbb{M}_2\right) \leq 0 \\ & \textit{3.} \quad \mathrm{Tr}\left(\mathbb{V}\mathbb{M}_3\right) \leq 1 \\ & \textit{4.} \quad \mathbb{V} \succeq 0, \\ & \textit{5.} \quad \mathrm{rank}(\mathbb{V}) = 1. \end{aligned}$$

Note that, in Problem 3.9, if Constraint 5 does not exist, we have a standard SDP which can be solved exactly in polynomial time [10]. Therefore, we relax the problem by removing the rank-constraint as follows.

Problem 3.10.

$$\begin{aligned} \underset{\mathbb{V}}{\textit{maximize}} \quad & \mathrm{Tr}\left(\mathbb{V}\mathbb{M}_0\right) \\ \textit{subject to} \quad & \textit{1.} \quad \mathrm{Tr}\left(\mathbb{V}\mathbb{M}_1\right) \leq 0, \\ & \textit{2.} \quad \mathrm{Tr}\left(\mathbb{V}\mathbb{M}_2\right) \leq 0 \\ & \textit{3.} \quad \mathrm{Tr}\left(\mathbb{V}\mathbb{M}_3\right) \leq 1 \\ & \textit{4.} \quad \mathbb{V} \succeq 0. \end{aligned}$$

Note that Problem 3.10 is a standard SDP and can be solved exactly in polynomial time using standard algorithms such as the interior-point algorithm [10]. This solution acts as an upper bound to the solution of Problem 3.9 since the search space gets expanded with the removal of Constraint 5. Furthermore, in our simulation experiments, we have observed that Problem 3.10 does not yield a rank-1 solution. Therefore, we investigate Problem 3.8 using approximations based on randomization.

Stage 3: Randomization

Let $\boldsymbol{\epsilon} \sim \mathscr{N}(\mathbf{0}, \mathbb{V})$ denote a random vector of size $(2L+1)$. In other words, since $\mathbb{V} = \mathbb{E}(\boldsymbol{\epsilon}\boldsymbol{\epsilon}^T)$, Problem 3.9 can be interpreted [34] as follows.

Problem 3.11.

$$\begin{aligned} \underset{\mathbb{V}}{\textit{maximize}} \quad & \mathrm{Tr}\left[\mathbb{E}(\boldsymbol{\epsilon}\boldsymbol{\epsilon}^T)\mathbb{M}_0\right] \\ \textit{subject to} \quad & \textit{1.} \quad \mathrm{Tr}\left[\mathbb{E}(\boldsymbol{\epsilon}\boldsymbol{\epsilon}^T)\mathbb{M}_1\right] \leq 0, \\ & \textit{2.} \quad \mathrm{Tr}\left[\mathbb{E}(\boldsymbol{\epsilon}\boldsymbol{\epsilon}^T)\mathbb{M}_2\right] \leq 0 \\ & \textit{3.} \quad \mathrm{Tr}\left[\mathbb{E}(\boldsymbol{\epsilon}\boldsymbol{\epsilon}^T)\mathbb{M}_3\right] \leq 1. \end{aligned}$$

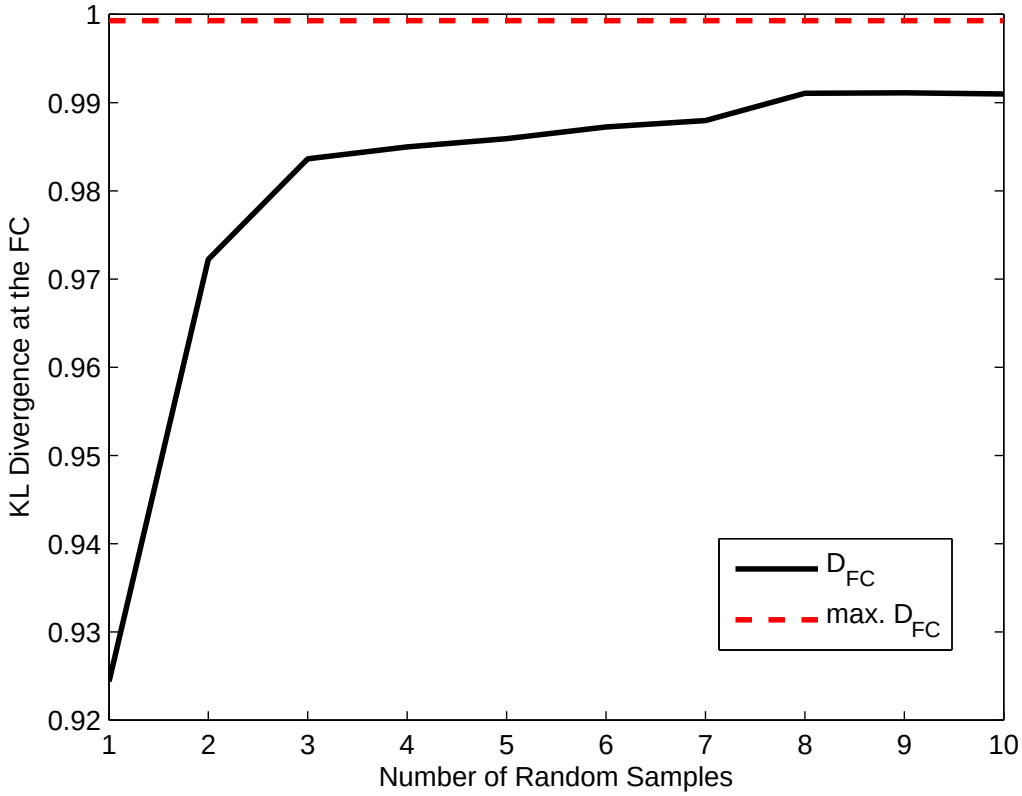
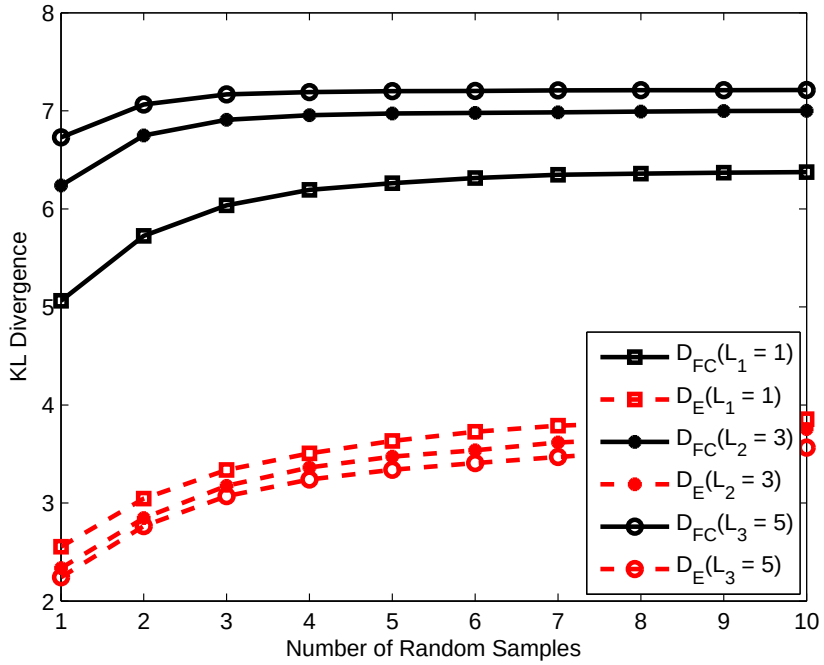
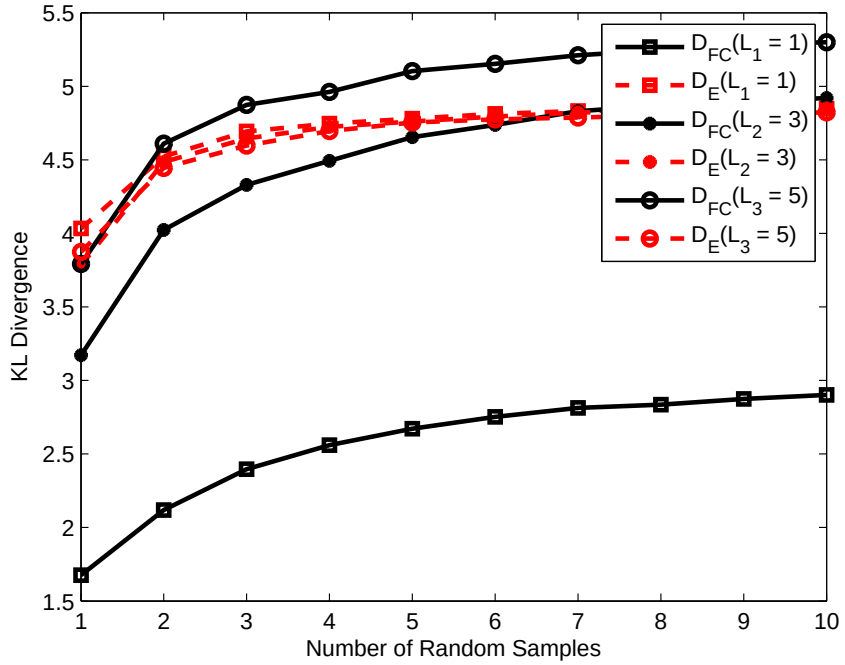


Figure 3.2: Improvement in KL Divergence with increasing number of samples M in the randomization procedure

This interpretation motivates us to construct $\mathbf{v}^*$ in the following manner.



(a) ($\sigma_{fc}^2 = 0.1, \sigma_e^2 = 0.5$)



(b) ($\sigma_{fc}^2 = 0.5, \sigma_e^2 = 0.1$)

Figure 3.3: KL Divergences at both FC and Eve for increasing number of random samples M, when $L = 1, 3, 5$, $N = 10$ and $\alpha = 5$.

Let $\boldsymbol{\epsilon}_1, \cdots, \boldsymbol{\epsilon}_M$, be M vectors that are sampled from the distribution $\mathscr{N}(\mathbf{0}, \mathbb{V}^*)$, where $\mathbb{V}^*$ is the solution to Problem 3.10. Note that each of these vectors $\boldsymbol{\epsilon}_1, \cdots, \boldsymbol{\epsilon}_M$ are potential candidates to approximate the solution of Problem 3.8. Therefore, we pick the best solution from the M available vectors as follows.

$$\begin{aligned} &\text{For } m = 1, \cdots, M, \\ &\qquad \text{Sample } \boldsymbol{\epsilon}_m \text{ from } \mathscr{N}(\mathbf{0}, \mathbb{V}^*). \\ &\qquad \text{Evaluate } f(m) = \boldsymbol{\epsilon}_m^T \mathbb{M}_0 \boldsymbol{\epsilon}_m \\ &\text{Find } m^* = \underset{m=1,\cdots,M}{\arg\max}\, f(m). \end{aligned} \tag{3.19}$$

Therefore, we propose $\mathbf{v}^* = \boldsymbol{\epsilon}_{m^*}$ as the solution to Problem 3.8. From $\mathbf{v}^*$, we evaluate the approximate solution $\mathbf{z}^*$ to Problem 3.6, as follows:

$$\mathbf{z}^* = \begin{bmatrix} \mathbf{x}^* \\ \mathbf{y}^* \end{bmatrix} = \left(\boldsymbol{\mu}_2^T \mathbf{v}\right)^{-1} \left(\boldsymbol{\mu}_1^T \mathbf{v}\right) \tag{3.20}$$

where

$$\boldsymbol{\mu}_1 = \begin{bmatrix} 1 \\ \vdots \\ 1 \\ 0 \end{bmatrix} \quad \text{and} \quad \boldsymbol{\mu}_2 = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}.$$

3.4 Simulation Results

Consider an example network with $N = 10$ sensors, each equipped with $L = 5$ transmitting antennas, a total energy budget $\mathscr{E} = 2$ and $\sigma_s^2 = 1$. We assume that the PoI is characterized by $\theta = 1$, with prior probabilities $\pi_0 = 1 - \pi_1 = 0.8$. Furthermore, we assume that each of the channel gains at both the FC and Eve are realizations of a standard Rayleigh distribution. In our simulations, we take 100 different realizations of this Rayleigh distribution in

order to evaluate the average performance of our system in all our simulation results.

Before we investigate the performance of our network design, we first focus our attention on the randomization procedure used in Stage 2 of our proposed algorithm. It is well known that the randomization procedure approaches the optimal value with increasing number of random samples, when there is only one non-convex constraint in the problem statement [34]. But, since $\mathbb{M}_1$ and $\mathbb{M}_2$ are not positive semidefinite, there is no guarantee for the convergence of the approximated solution to the optimal one. Therefore, we first demonstrate the performance of the randomization procedure proposed to solve Problem 3.6 (single sensor case) for this example scenario when the tolerable Eve's constraint is given by $\delta = 0.5$. We plot the KL Divergence at the FC due to a single sensor in Figure 3.2 and show that the system performance in terms of KL Divergence at FC by our proposed approach improves with increasing number of random samples collected in the randomization procedure presented in Stage 2 of our design-algorithm. Note that the solid-line in Figure 3.2 corresponds to the average KL Divergence that the FC attains, while the dotted-line corresponds to the average maximum KL Divergence attained at the FC in the absence of the secrecy constraint, as given by Equation (3.13a). Since the performance of the dotted-line is the optimal KL Divergence attained at the FC in the absence of a secrecy constraint on Eve, it acts like an upper bound to the optimal solution to Problem 3.6. Note that the randomization approach works well in our problem since the approximate solution to $n = 1$ case converges to a value that is very close to the optimal solution.

Given that the randomization algorithm works well in the context of our problem framework, we next illustrate the performance of our proposed design algorithm in two different simulation results. In the first experiment, we plot $\mathscr{D}_{FC}$ and $\mathscr{D}_E$ for varying number of random samples in the randomization procedure in Figure 3.3. As per our intuition, we observe that the KL Divergences at both the FC and Eve increases with increasing number of random samples. More specifically, in Figure 3.3a, we consider the scenario where the FC has better channels than Eve (i.e. when $\sigma_{fc}^2 \leq \sigma_e^2$). Here, we observe that the difference

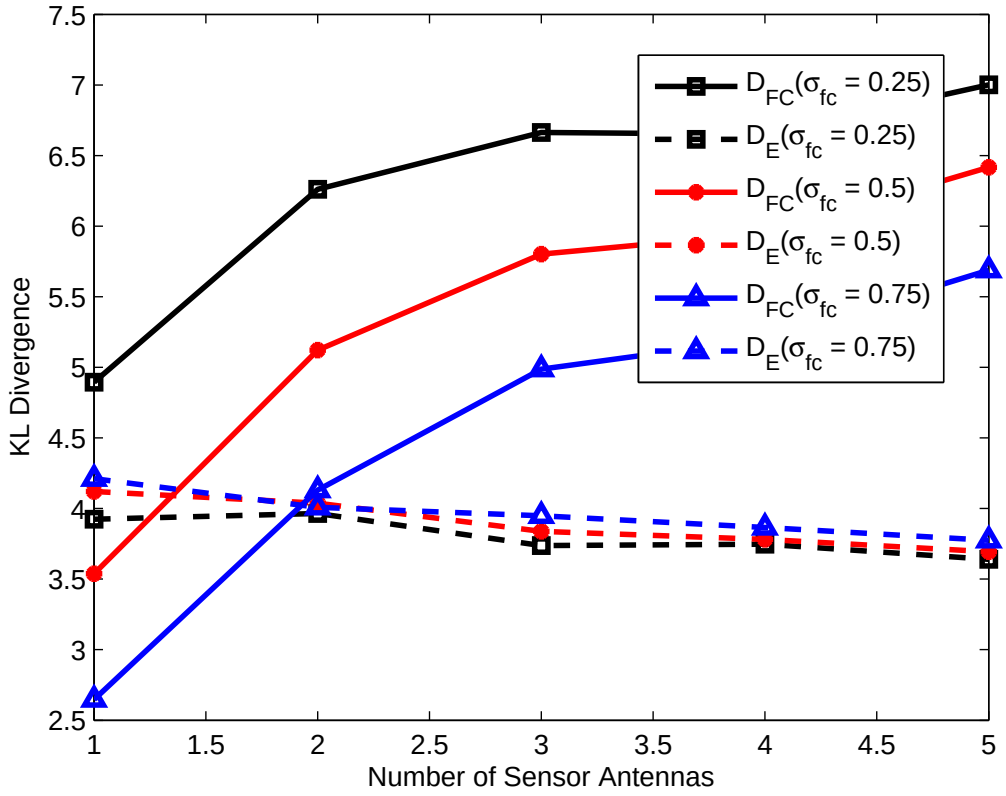


Figure 3.4: KL Divergences at both FC and Eve for increasing number of sensor antennas L, when $\sigma_e^2 = 0.5$, $N = 10$ and $\alpha = 5$.

between the KL Divergences at the FC and Eve increases with increasing number of sensor antennas. On the other hand, in Figure 3.3b, we consider the opposite scenario where the Eve has better channels than the FC (i.e. when $\sigma_{fc}^2 \leq \sigma_e^2$). In this scenario, we observe a tremendous improvement in the performance of the FC as the number of sensor antennas increases.

In the second scenario, in Figure 3.4, we plot the relationship between $\mathscr{D}_{FC}$ and $\mathscr{D}_E$ with respect to the number of sensor antennas L for different channel scenarios at the FC respectively. Here, we clearly observe that, while $\mathscr{D}_{FC}$ increases with increasing number of antennas, $\mathscr{D}_E$ decreases at a relatively slow rate. Furthermore, the intersection points between the $\mathscr{D}_{FC}$ and $\mathscr{D}_E$ curves corresponding to a given value of σ_{fc}^2 give us the minimum number of sensor antennas needed to ensure a greater KL Divergence at the FC than that

of Eve.

3.5 Summary

In summary, we have proposed a transmit-diversity framework for detection networks where the KL Divergence at the FC is maximized in the presence of a total energy budget at each sensor and a constraint on the Eve's KL Divergence. In this transmit-diversity framework, the sensors are allowed to construct a transmission signal by distributing the total energy between their observations and the artificial noise across multiple transmitting antennas. In this chapter, we have proposed a two-stage approximate algorithm to find efficient signaling at the sensors based on a greedy decomposition and random-sampling procedures. We have presented numerical results to illustrate the performance of the proposed design.

CHAPTER 4
BYZANTINE ATTACKS IN INFERENCE NETWORKS WITH M-ARY QUANTIZED DATA

Statistical inferences are reliable only when the data-collection process is reliable. If the sensing agents participating in the data-collection process are compromised, the inference performance can be deteriorated significantly. Therefore, in this chapter, we investigate the fundamental performance-limits of inference networks in the presence of Byzantine attacks, in addition to proposing an anomaly-detection scheme to detect the compromised agents in the network. We focus our attention on two inference problems, namely *detection* and *estimation*, when the sensors quantize their data to a more general M-ary symbols, with $M = 2$ (binary quantization) and $M \to \infty$ (centralized inference networks) being special cases.

The main contributions of the chapter are three-fold. First, in Section 4.2, we define a Byzantine attack model for a sensor network with individual sensors quantizing their observations into one of the M-ary symbols, when the attacker does not have complete knowledge about the true state of the POI and thresholds employed by the sensors. We

model the attack strategy as a flipping probability matrix, where $(i, j)^{th}$ entry represents the probability with which the i^{th} symbol is flipped into the j^{th} symbol. Second, we show that quantization into M-ary symbols at the sensors, as opposed to binary quantization, improves both inference as well as security performance simultaneously. As a function of the number of Byzantine nodes in the network, we derive the optimal flipping matrix for both ideal and non-ideal (discrete memoryless) channels in Sections 4.3 and 4.4 respectively. In Section 4.5, we investigate the optimal Byzantine attack in the context of distributed detection and estimation when the attacker is resource-constrained to compromise enough number of nodes in the network to *blind* the FC (to be defined in Section 4.2). Finally, in Section 4.6, we extend the mitigation scheme presented by Rawat *et al.* in [52] to the more general case where sensors generate M-ary symbols. We present numerical/simulation results to illustrate the performance of the proposed network-design.

4.1 Literature Survey

In the context of distributed inference networks, the sensing agents compress their observations by mapping them to one of the symbols in an alphabet set of size M, prior to transmission to the FC. In the context of sensing agents with binary quantization ($M = 2$) capabilities, a lot of work is done in the past to address Byzantine attacks in the context of distributed inference networks (see a recent survey [68] by Vempaty *et al.*).

Byzantine attacks (proposed by Lamport *et al.* in [27]) in general, are arbitrary and may refer to many types of malicious behavior. In this chapter, we focus only on the data-falsification aspect of the Byzantine attack wherein one or more compromised nodes of the network send false information to the FC in order to deteriorate the inference performance of the network. A well known example of this attack is the *man-in-the-middle* attack [44] where, on one hand, the attacker collects data from the sensors whose authentication process is compromised by the attacker emulating as the FC, while, on the other hand, the

attacker sends false information to the FC using the compromised sensors' identity. In summary, if the i^{th} sensor's authentication is compromised, the attacker remains invisible to the network, accepts the true decision u_i from the i^{th} sensor and sends v_i to the FC in order to deteriorate the inference performance.

Marano *et al.*, in [35], analyzed the Byzantine attack on a network of sensors carrying out the task of distributed detection, where the attacker is assumed to have complete knowledge about the hypotheses. This represents the extreme case of Byzantine nodes having an extra power of knowing the true hypothesis. In their model, they assumed that the sensors quantized their respective observations into M-ary symbols, which are later fused at the FC. The Byzantine nodes pick symbols using an optimal probability distribution that are conditioned on the true hypotheses, and transmit them to the FC in order to maximally degrade the detection performance. Rawat *et al.*, in [52], also considered the problem of distributed detection in the presence of Byzantine attacks with binary quantizers at the sensors in their analysis. Unlike the authors in [35], Rawat *et al.* did not assume complete knowledge of the true hypotheses at the Byzantine attacker. Instead, they assumed that the Byzantine nodes derive the knowledge about the true hypotheses from their own sensing observations. In other words, a Byzantine node potentially flips the local decision made at the node. It does not modify the thresholds at the sensor quantizers. Rawat *et al.* also analyzed the performance of the network in the presence of independent and collaborative Byzantine attacks and modeled the problem as a zero-sum game between the sensor network and the Byzantine attacker. In addition to the analysis of distributed detection in the presence of Byzantine attacks, a reputation-based scheme was proposed by Rawat *et al.* in [52] for identifying the Byzantine nodes by accumulating the deviations between each sensor's decision and the FC's decision over a time window of duration T. If the accumulated number of deviations is greater than a prescribed threshold for a given node, then the FC tags it as a Byzantine node. In order to mitigate the attack, the FC removes nodes which are tagged Byzantine node from the fusion rule. Another mitigation scheme was proposed

by Vempaty *et al.* [66], where each sensor's behavior is learnt over time and compared to the known behavior of the honest nodes. Any node with significant deviation in the learnt behavior from the expected honest behavior is labeled a Byzantine node. Having learnt their parameters, the authors also proposed the use of this information to adapt their fusion rule so as to maximize the performance of the FC. In contrast to the parallel topology in sensor networks, Kailkhura *et al.* in [21] investigated the problem of Byzantine attacks on distributed detection in a hierarchical sensor network. They presented the optimal Byzantine strategy when the sensors communicate their decisions to the FC in multiple hops of a balanced tree. They assumed that the cost of compromising sensors at different levels of the tree varies, and found the optimal Byzantine strategy that minimizes the cost of attacking a given hierarchical network.

Soltanmohammadi *et al.* in [57] investigated the problem of distributed detection in the presence of different types of Byzantine nodes. Each Byzantine node type corresponds to a different operating point, and, therefore, the authors considered the problem of identifying different Byzantine nodes, along with their operating points. The problem of maximum-likelihood (ML) estimation of the operating points was formulated and solved using the expectation-maximization (EM) algorithm. Once the Byzantine node operating points are estimated, this information was utilized at the FC to mitigate the malicious activity in the network, and also to improve global detection performance.

Distributed target localization in the presence of Byzantine attacks was addressed by Vempaty *et al.* in [67], where the sensors quantize their observations into binary decisions, which are transmitted to the FC. Similar to Rawat *et al.*'s approach in [52], the authors in [67] investigated the problem of distributed target localization from both the network's and Byzantine attacker's perspectives, first by identifying the optimal Byzantine attack and second, mitigating the impact of the attack with the use of non-identical quantizers at the sensors.

In this chapter, we extend the framework of Byzantine attacks when Byzantine nodes

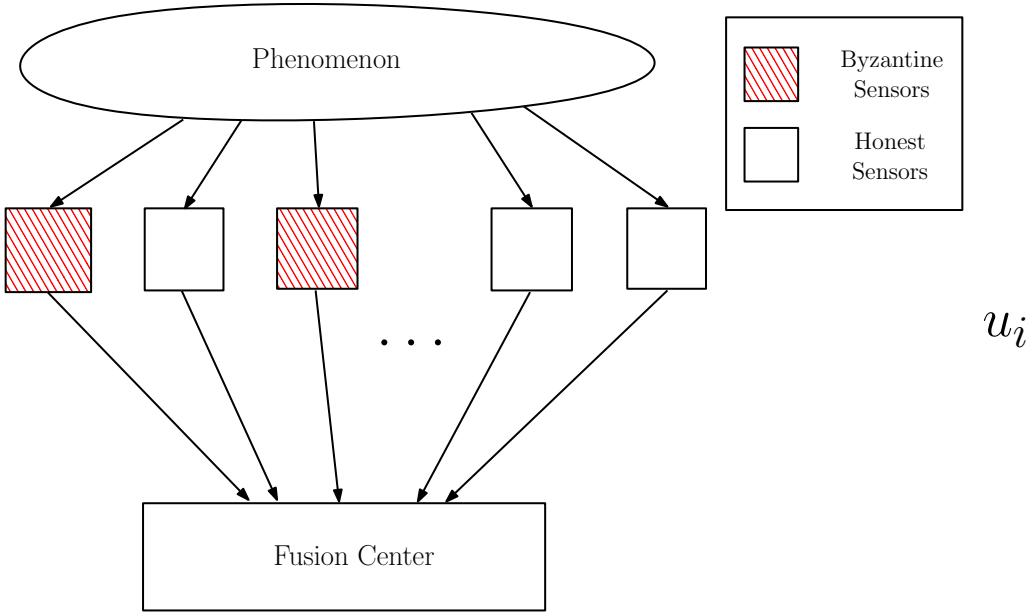
do not have complete knowledge about the true state of the phenomenon-of-interest (POI), and when the sensors generate M-ary symbols instead of binary symbols. We also assume that the Byzantine attacker is ignorant about the quantization thresholds used at the sensors to generate the M-ary symbols.[1] Under these assumptions, we address two inference problems: binary hypotheses-testing and parameter estimation.

4.2 System Model

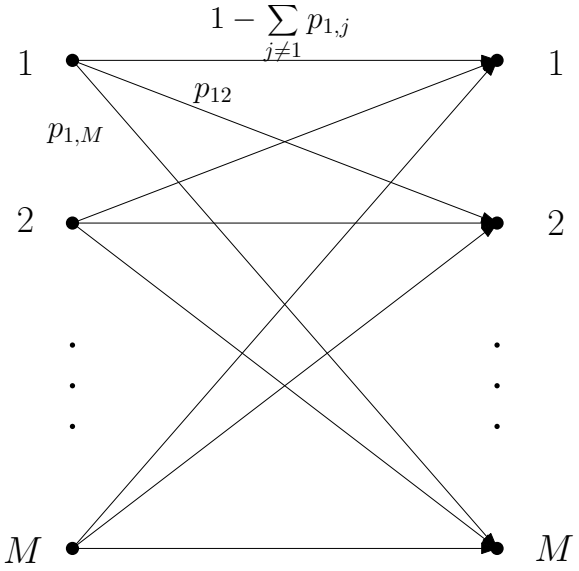
Consider an inference (sensor) network with N sensors, where α fraction of the nodes in the network are assumed to be compromised (Refer to Figure 4.1a). These compromised sensors transmit false data to the fusion center (FC) in order to deteriorate the inference performance of the network. We assume that the network is designed to infer about a particular phenomenon, regarding which sensors acquire *conditionally-independent* observations. We denote the observation of the i^{th} sensor as r_i. This observation r_i is mapped to one of the M symbols, $u_i \in \{1, \cdots, M\}$. In a compromised inference network, since the Byzantine sensors do not transmit their true quantized data, we denote the transmitted symbol as v_i at the i^{th} sensor. If the node i is honest, then $v_i = u_i$. Otherwise, we assume that the Byzantine sensor modifies $u_i = l$ to $v_i = m$ with a probability p_{lm}, as shown in Figure 4.1b. For the sake of compactness, we denote the transition probabilities depicted in the graph in Figure 4.1b using a row-stochastic matrix $\mathbb{P}$, as follows:

$$\mathbb{P} = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1M} \\ p_{21} & p_{22} & \cdots & p_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ p_{M1} & p_{M2} & \cdots & p_{MM} \end{bmatrix}. \tag{4.1}$$

[1] The well-known attacker-in-the-middle is one such example.



(a) Sensor Network Model



(b) Byzantine Attack Model

Figure 4.1: Distributed Inference Network in the Presence of Byzantine Attacks

Since the attacker has no knowledge of quantization thresholds employed at each sensor, we assume that $\mathbb{P}$ is independent of the sensor observations. The messages $\mathbf{v} = \{v_1, v_2, \cdots, v_N\}$ are transmitted to the fusion center (FC) where a global inference is made about the phenomenon of interest based on $\mathbf{v}$.

In order to consider the general inference problem, we assume that $\theta \in \Theta$ is the parameter that denotes the phenomenon of interest in the received signal r_i at the i^{th} sensor. If we are considering a detection/classification problem, θ is discrete (finite or countably infinite). In the case of parameter estimation, Θ is a continuous set. Without any loss of generality, we assume $\Theta = \{0, 1, \cdots, K-1\}$ if the problem of interest is classification. Hence, detection is a special case of classification with $K = 2$. In the case of estimation, we assume that $\Theta = \mathbb{R}$.

Based on this system model, we investigate the optimal Byzantine attack under different scenarios in the remaining chapter. Furthermore, we also propose a mitigation scheme where the FC computes a reputation index for each sensing agent to identify and remove the compromised nodes from the fusion rule.

4.3 Optimal Byzantine Attacks: Noiseless Channels

Given the conditional distribution of r_i, $p(r_i|\theta)$, and the sensor quantization thresholds, λ_j for $0 \leq j \leq M$, the conditional distribution of u_i can be found as

$$P(u_i = m|\theta) = \int_{\lambda_{m-1}}^{\lambda_m} p(r_i|\theta) dr_i \tag{4.2}$$

for all $m = 1, 2, \cdots, M$.

If the true quantized symbol at the i^{th} node is $u_i = m$, a compromised node will modify it into $v_i = l$ as depicted in Figure 4.1b, and transmit it to the FC. Since the FC is not aware of the type of the node (honest or Byzantine), it is natural to assume that node i is compromised with probability α, where α is the fraction of nodes in the network that are

compromised. Therefore, we find the conditional distribution of v_i at the FC as follows.

$$\begin{aligned} P(v_i = m|\theta) &= \alpha P(v_i = m|i = Byzantine, \theta) + (1-\alpha)P(v_i = m|i = Honest, \theta) \\ &= \alpha \sum_{l=1}^{M} P(u_i = l|\theta) \cdot P(v_i = m|u_i = l, \theta) + (1-\alpha)P(u_i = m|\theta) \\ &= \alpha \sum_{l=1}^{M} p_{lm} P(u_i = l|\theta) + (1-\alpha)P(u_i = m|\theta) \\ &= \alpha \sum_{l \neq m} p_{lm} P(u_i = l|\theta) + [(1-\alpha) + \alpha p_{mm}]P(u_i = m|\theta) \\ &= [(1-\alpha) + \alpha p_{mm}] + \sum_{l \neq m} \{\alpha p_{lm} - [(1-\alpha) + \alpha p_{mm}]\} P(u_i = l|\theta). \end{aligned} \tag{4.3}$$

The goal of a Byzantine attack is to blind the FC with the least amount of effort (minimum α). To totally blind the FC is equivalent to making $P(v_i = m|\theta) = 1/M$ for all $0 \leq m \leq M-1$. In Equation (4.3), the RHS consists of two terms. The first one is based on prior knowledge and the second term conveys information based on the observations. In order to blind the FC, the attacker should make the second term equal to zero. Since the attacker does not have any knowledge regarding $P(u_i = l|\theta)$, it can make the second term of Equation (4.3) equal to zero by setting

$$\alpha p_{lm} = (1-\alpha) + \alpha p_{mm}, \qquad \forall\, l \neq m. \tag{4.4}$$

Then the conditional probability $P(v_i = m|\theta) = (1-\alpha) + \alpha p_{mm}$ becomes independent of the observations r_i (or its quantized version u_i), resulting in equiprobable symbols at the FC. In other words, the received vector $\mathbf{v} = \{v_1, v_2, \cdots, v_N\}$ does not carry any informa-

tion about θ and, therefore, results in the most degraded performance at the FC. So, the FC now has to solely depend on its prior information about θ in making an inference.

Having identified the condition in Equation (4.4) under which the Byzantine attack makes the greatest impact on the performance of the network, we identify the strategy that the attacker should employ in order to achieve this condition as follows. Since we need

$$P(v_i = m|\theta) = (1-\alpha) + \alpha p_{mm} = 1/M,$$

$\alpha = \frac{M-1}{(1-p_{mm})M}$. To minimize α, one needs to make $p_{mm} = 0$. In this chapter, we denote the α corresponding to this optimal strategy that minimizes the Byzantine attacker's resources required to blind the FC as α_{blind}. Hence,

$$\alpha_{blind} = \frac{M-1}{M}.$$

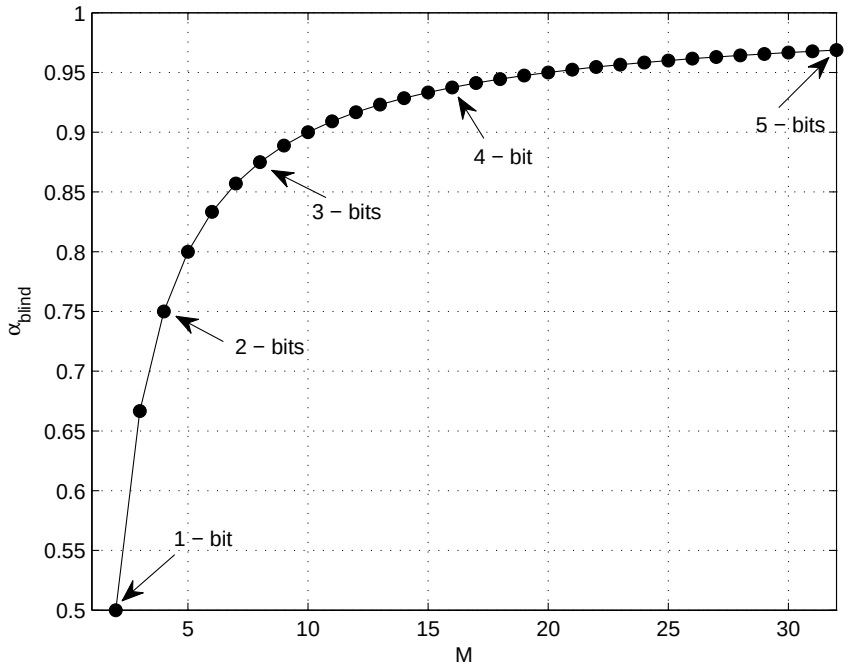
Rearranging Equation (4.4), we have

$$\frac{1}{\alpha} = 1 + (p_{lm} - p_{mm}) = 1 + p_{lm} \qquad \forall\, l \neq m. \tag{4.5}$$

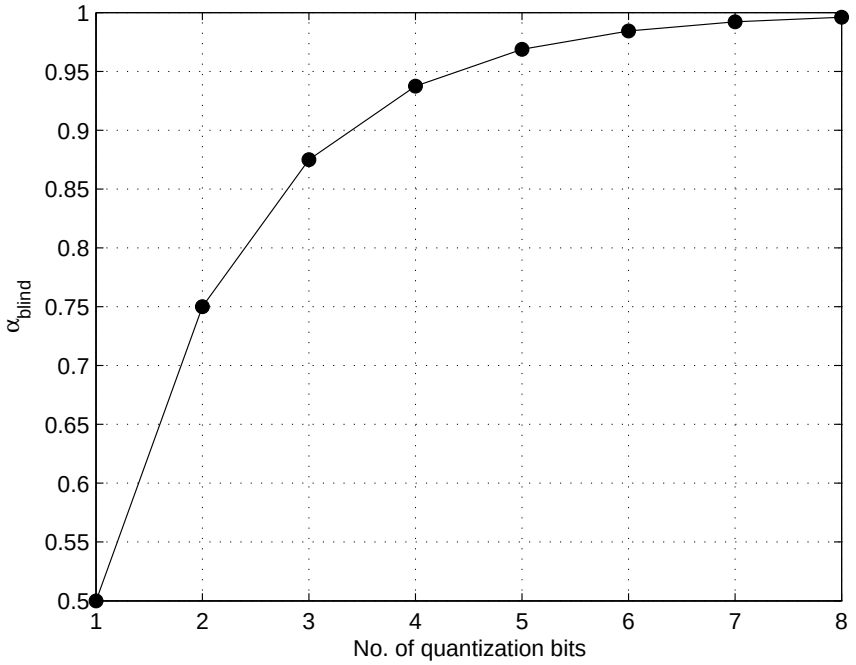
By setting α to α_{blind}, we have $p_{lm} = 1/(M-1), \forall\, l \neq m$. That is, the transition probability $\mathbb{P}$ is a highly-symmetric matrix. We summarize the result as a theorem as follows.

Theorem 4.1. *If the Byzantine attacker has no knowledge of the quantization thresholds employed at each sensor, then the optimal Byzantine attack is given as*

$$\begin{aligned} p_{lm} &= \begin{cases} \dfrac{1}{M-1} & ;\ \textit{if}\ l \neq m \\ 0 & ;\ \textit{otherwise} \end{cases} \\ \alpha_{blind} &= \frac{M-1}{M}. \end{aligned} \tag{4.6}$$



(a) α_{blind} vs. M



(b) α_{blind} vs. Number of quantization bits ($\log_2 M$)

Figure 4.2: Improvement in α_{blind} with increasing number of quantization levels

Quantization bits $(\log_2 M)$	**Quantization levels** (M)	**Blinding fraction** (α_{blind})
1	$2^1 = 2$	0.5
2	$2^2 = 4$	0.75
3	$2^3 = 8$	0.875
4	$2^4 = 16$	0.9375
5	$2^5 = 32$	0.9688
6	$2^6 = 64$	0.9844
7	$2^7 = 128$	0.9922
8	$2^8 = 256$	0.9961

Table 4.1: Improvement in α_{blind} with increasing number of quantization levels M, and quantization bits, $\log_2 M$

We term Equation (4.6) as the optimal Byzantine attack, since the FC does not get any information from the data $\mathbf{v}$ it receives from the sensors to perform an inference task. Therefore, the FC has to rely on prior information about the parameter θ, if available.

In Figure 4.2, we show how α_{blind} scales with increasing quantization alphabet size, M. Since the quantized symbols are encoded into bits, we also show an exponential increase in α_{blind} as the number of bits needed to encode the M symbols, i.e., $\log_2 M$, increases. This is also shown in Table 4.1. Note that, if the sensors use one additional quantization-bit (2-bit quantization) in their quantization scheme instead of 1-bit quantization (binary quantization), then the α_{blind} increases from 0.5 to 0.75. This trend is observed with increasing number of quantization bits, and when the sensors employ an 8-bit quantizer, then

the attacker needs to compromise at least 99.6% of the sensors in the network to blind the FC. Obviously, the improvement in security performance is not free as the sensors incur a communication cost in terms of energy and bandwidth as the number of quantization bits increases. Therefore, in a practical world, the network designer faces a trade-off between the communication cost and the security guarantees.

Also, note that, when $M = 2$ (1-bit quantization), our results coincide with those of Rawat *et al.* in [52], where the focus was on the problem of binary hypotheses testing in a distributed sensor network. On the other hand, our results are more general as they address any inference problem - detection, estimation or classification in a distributed sensor network. Another extreme case to note is when $M \to \infty$, in which case, $\alpha_{blind} \to 1$. This means that the Byzantine attacker cannot blind the FC unless all the sensors are compromised.

4.4 Optimal Byzantine Attacks: Discrete Memoryless Channels

Given that the messages $\mathbf{v} = \{v_1, v_2, \cdots, v_N\}$ are transmitted to the fusion center (FC), we assume a discrete noise channel $\mathbb{Q} = [q_{mn}]$ between the sensors and the FC, where q_{mn} is the probability with which $v_i = m$ is transformed to symbol $z_i = n$ at the i^{th} sensor. Based on the received $\mathbf{z}$ at the FC, a global inference is made about the phenomenon of interest. In this section, we assume that the row-stochastic channel matrix $\mathbb{Q}$ is invertible for the sake of tractability.

Given the transition probability matrix $\mathbb{Q}$ for the channel between the sensors and the FC, we assume that the FC receives $z_i = n$ when the the i^{th} sensor transmits $v_i = m$, with a probability q_{mn}. The conditional distribution of $z_i = n$ under a given phenomenon θ, is

given as

$$P(z_i = n|\theta) = \sum_{m=1}^{M} q_{mn} P(v_i = m|\theta). \tag{4.7}$$

Note that if $\mathbb{Q}$ is a doubly stochastic matrix, since $\sum_{m=1}^{M} q_{mn} = 1$, it is sufficient for the Byzantine attacker to ensure $P(v_i = m|\theta) = \frac{1}{M}$. Thus, by Theorem 4.1, we have the following theorem when $\mathbb{Q}$ is a doubly stochastic matrix.

Theorem 4.2. *If the channel matrix $\mathbb{Q}$ is doubly-stochastic, and if the Byzantine attacker has no knowledge about the sensors' quantization thresholds, then the optimal Byzantine attack is given as*

$$\begin{aligned} p_{lm} &= \begin{cases} \dfrac{1}{M-1} & ;\ \textit{if } l \neq m \\ 0 & ;\ \textit{otherwise} \end{cases} \\ \alpha_{blind} &= \frac{M-1}{M}. \end{aligned} \tag{4.8}$$

Therefore, we focus our attention to any general row-stochastic channel matrix $\mathbb{Q}$, where $\sum_{m=1}^{M} q_{mn}$ need not necessarily sum to unity for all $n = 1, \cdots, M$. In other words, the Byzantine attacker has to find an alternative strategy to blind the FC, where $P(z_i = n|\theta) = \frac{1}{M}$. Substituting Equation (4.3) in Equation (4.7) and rearranging the terms, we have

$$\begin{aligned} P(z_i = n|\theta) &= \sum_{m=1}^{M} q_{mn} P(v_i = m|\theta) \\ &= \sum_{m=1}^{M} q_{mn}[(1-\alpha) + \alpha p_{mm}] \\ &\quad + \sum_{m=1}^{M} q_{mn} \left\{ \sum_{l \neq m} \{\alpha p_{lm} - [(1-\alpha) + \alpha p_{mm}]\} P(u_i = l|\theta) \right\} \\ &= \sum_{m=1}^{M} q_{mn}[(1-\alpha) + \alpha p_{mm}] \\ &\quad + \sum_{l=1}^{M} \left[\sum_{m \neq l} q_{mn} \{\alpha p_{lm} - [(1-\alpha) + \alpha p_{mm}]\} \right] P(u_i = l|\theta). \end{aligned} \tag{4.9}$$

The goal of a Byzantine attack is to blind the FC with the least amount of effort (minimum α). To totally blind the FC is equivalent to making $P(z_i = n|\theta) = 1/M$ for all $0 \leq n \leq M-1$. In Equation (4.9), the RHS consists of two terms. The first one is based on prior knowledge and the second term conveys information based on the observations. In order to blind the FC, the attacker should make the second term equal to zero. Since the attacker does not have any knowledge regarding $P(u_i = l|\theta)$, it can make the second term of Equation (4.9) equal to zero by setting

$$\sum_{m \neq l} q_{mn}\{\alpha p_{lm} - [(1-\alpha) + \alpha p_{mm}]\} = 0 \text{ for all } 1 \leq n, l \leq M. \tag{4.10}$$

Then the conditional probability $P(z_i = n|\theta) = \sum_{m=1}^{M} q_{mn}[(1-\alpha) + \alpha p_{mm}]$ becomes independent of the observations r_i (or its quantized version u_i), resulting in equiprobable symbols at the FC. In other words, the received vector $\mathbf{z} = \{z_1, z_2, \cdots, z_N\}$ does not carry any information about $\mathbf{u} = \{u_1, u_2, \cdots, u_N\}$, thus making FC solely dependent on its prior information about θ in making an inference.

In order to identify the strategy that the attacker should employ to achieve the condition in Equation (4.10), for all $n = 1, \cdots, M$, we need

$$\begin{aligned} P(z_i = n|\theta) &= \frac{1}{M}, \\ \text{or, } \sum_{m=1}^{M} q_{mn}\left\{(1-\alpha) + \alpha p_{mm}\right\} &= \frac{1}{M}. \end{aligned} \tag{4.11}$$

In matrix form, we can rewrite Equation (4.11) as

$$(1-\alpha)\mathbf{1}^T\mathbb{Q} + \alpha\mathbf{p}^T\mathbb{Q} = \frac{1}{M}\mathbf{1}^T,$$

where $\mathbf{1}$ is an all-one column-vector and $\mathbf{p} = [p_{11}, \cdots, p_{MM}]^T$ is the column-vector of all

diagonal elements of $\mathbb{P}$. In other words,

$$\alpha(\mathbf{1}-\mathbf{p}) = \mathbf{1} - \frac{1}{M}\left(\mathbb{Q}^T\right)^{-1}\mathbf{1} \tag{4.12}$$

Note that every element in the LHS of Equation (4.12) always lies between 0 and 1. Therefore, the existence of the Byzantine's optimal strategy relies on the following condition. In other words,

$$\mathbf{0} \quad \leq \quad \left(\mathbb{Q}^T\right)^{-1}\mathbf{1} \quad \leq \quad M\,\mathbf{1}. \tag{4.13}$$

If (4.13) does not hold, there does not exist an optimal strategy. Given that the condition in Equation (4.13) holds, the minimum α can be found as follows.

$$\begin{aligned} \alpha_{blind} &= \min\left\{\mathbf{1} - \frac{1}{M}\left(\mathbb{Q}^T\right)^{-1}\mathbf{1}\right\} \\ &= 1 - \frac{1}{M}\max\left\{\left(\mathbb{Q}^T\right)^{-1}\mathbf{1}\right\}. \end{aligned} \tag{4.14}$$

Therefore, $\mathbf{p}$ can be calculated as

$$\begin{aligned} \mathbf{p} &= \mathbf{1} - \frac{1}{\alpha_{blind}}\left(\mathbf{1} - \frac{1}{M}\left(\mathbb{Q}^T\right)^{-1}\mathbf{1}\right) \\ &= \frac{1}{\alpha_{blind}M}\left(\mathbb{Q}^T\right)^{-1}\mathbf{1} - \frac{1-\alpha_{blind}}{\alpha_{blind}}\mathbf{1}. \end{aligned} \tag{4.15}$$

Next, in order to find the rest of the $\mathbb{P}$ matrix, let us consider Equation (4.10). Adding $q_{ln}\left\{\alpha p_{ll} - [1-\alpha+\alpha p_{ll}]\right\}$ on both sides to Equation (4.10), for all $1 \leq n, l \leq M$, we have

$$\begin{aligned} \sum_{m=1}^{M} q_{mn}\{\alpha p_{lm} - [(1-\alpha)+\alpha p_{mm}]\} &= -q_{ln}(1-\alpha) \\ \text{or,} \quad \alpha\sum_{m=1}^{M} q_{mn}p_{lm} &= \frac{1}{M} - q_{ln}(1-\alpha). \end{aligned} \tag{4.16}$$

In matrix form, we have

$$\alpha \mathbb{P}\mathbb{Q} = \frac{1}{M}\mathbb{1} - (1-\alpha)\mathbb{Q}, \tag{4.17}$$

where $\mathbb{1}$ is an all-one matrix. Equivalently, we have

$$\mathbb{P} = \frac{1}{\alpha M}\mathbb{1}\mathbb{Q}^{-1} - \frac{1-\alpha}{\alpha}\mathbb{I}, \tag{4.18}$$

where $\mathbb{I}$ is the identity matrix. Note that the vector $\mathbf{p}$ (comprising the diagonal elements of $\mathbb{P}$) obtained from Equation (4.18) is verified to be same as that from Equation (4.15).

In summary, we have the following theorem that provides the optimal Byzantine strategy in the presence of noisy FC channels:

Theorem 4.3. *Let the Byzantine attacker have no knowledge about the sensors' quantization thresholds, and, the FC's channel matrix be $\mathbb{Q}$. If $\mathbb{Q}$ is non-singular, and, if $\mathbf{0} \leq \left(\mathbb{Q}^T\right)^{-1}\mathbf{1} \leq M\mathbf{1}$, then the optimal Byzantine attack is given as*

$$\begin{aligned} \alpha_{blind} &= 1 - \frac{1}{M}\max\left\{\left(\mathbb{Q}^T\right)^{-1}\mathbf{1}\right\} \\ \mathbb{P} &= \frac{1}{\alpha_{blind} M}\mathbb{1}\mathbb{Q}^{-1} - \frac{1-\alpha_{blind}}{\alpha_{blind}}\mathbb{I}. \end{aligned} \tag{4.19}$$

Note that, if the channel matrix $\mathbb{Q}$ is doubly-stochastic, we have $\mathbb{Q}\mathbf{1} = \mathbf{1}$ and $\mathbb{Q}^T\mathbf{1} = \mathbf{1}$. Substituting these conditions in Equation (4.19), Theorem 4.3 reduces to Theorem 4.2.

Having identified the optimal Byzantine attack, one can observe that the attacker needs to compromise a huge number of sensors ($\alpha_{blind} = 1 - \frac{1}{M}\max\left\{\left(\mathbb{Q}^T\right)^{-1}\mathbf{1}\right\}$) in the network to blind the FC. Therefore, it is obvious that, in the case of a resource-constrained attacker, the attacker compromises a fixed fraction of nodes $\alpha \leq \alpha_{blind}$ in such a way that the performance degradation at the FC is maximized. In our future work, we will investigate the problem of finding the optimal strategy in the context of resource-constrained Byzantine attacks in the presence of noisy FC channels.

4.5 Optimal Byzantine Attacks: Constrained Resources

4.5.1 Distributed Detection

In this section, we consider a resource-constrained Byzantine attack on binary hypotheses testing in a distributed sensor network where the phenomenon of interest is denoted as θ and is modeled as follows:

$$\theta = \begin{cases} 0; \text{ if } H_0 \\ \\ 1; \text{ if } H_1 \end{cases} . \tag{4.20}$$

In order to characterize the performance of the FC, we consider Kullback-Leibler Divergence (KLD) as the performance metric. Note that KLD can be interpreted as the error exponent in the Neyman-Pearson detection framework [13], which means that the probability of missed detection goes to zero exponentially with the number of sensors at a rate equal to KLD computed at the FC. We denote KLD at the FC by $\mathbb{D}_{FC}$ and define it as follows:

$$\begin{aligned} \mathbb{D}_{FC} &= \mathbb{E}_{H_0}\left[\log\left(\frac{P(\mathbf{v}|H_0)}{P(\mathbf{v}|H_1)}\right)\right] \\ &= \sum_{\mathbf{m}\in\{1,\cdots,M\}^N} P(\mathbf{v}=\mathbf{m}|H_0)\cdot\log\left(\frac{P(\mathbf{v}=\mathbf{m}|H_0)}{P(\mathbf{v}=\mathbf{m}|H_1)}\right) \end{aligned} \tag{4.21}$$

Since we have assumed that the sensor observations are conditionally independent,[2] KLD can be expressed as

$$\mathbb{D}_{FC} = ND_{FC}, \tag{4.22}$$

where

$$D_{FC} = \sum_{m=1}^{M} P(v=m|H_0)\cdot\log\left(\frac{P(v=m|H_0)}{P(v=m|H_1)}\right).$$

Note that the optimal Byzantine attack, as given in Equation (4.6), results in equiprobable

[2]For notational convenience, sensor index i is ignored in the rest of the section.

symbols at the FC irrespective of the hypotheses. Therefore, $\mathbb{D}_{FC} = 0$ under optimal Byzantine attack, resulting in the blinding of the FC.

On the other hand, if the attacker does not have enough resources to compromise α_{blind} fraction of sensors in the network (i.e. $\alpha < \alpha_{blind}$), an optimal strategy for the Byzantine node is to use an appropriate $\mathbb{P}$ matrix that deteriorates the performance of the sensor network to the maximal extent. In this section, we restrict our search to finding the optimal $\mathbb{P}$ within a space of highly symmetric row-stochastic matrices, as given in Equation (4.23).

$$p_{jk} = \begin{cases} p & \text{if } j \neq k \\ 1-(M-1)p & \text{otherwise.} \end{cases} \tag{4.23}$$

Thus, we formulate the problem as follows.

Problem 4.1. *Given the value of $\alpha < \alpha_{blind}$, find the optimal $\mathbb{P}$ within a space of highly symmetric row-stochastic matrices, as given in Equation* (4.23)*, such that*

$$\begin{aligned} \underset{p}{\textit{minimize}} \quad & D_{FC} \\ \textit{subject to} \quad & 0 \leq p \leq \frac{1}{M-1} \end{aligned}$$

Theorem 4.4 presents the optimal flipping probability that provides the solution to Problem 4.1. Note that this result is independent of the design of the sensor network and, therefore, can be employed when the Byzantine has no knowledge about the network.

Theorem 4.4. *Given a fixed $\alpha < \dfrac{M-1}{M}$, the probability p that optimizes $\mathbb{P}$ within a space of highly symmetric row-stochastic matrices, as given in Equation* (4.23)*, such that D_{FC} is minimized, is given by*

$$p^* = \frac{1}{M-1}. \tag{4.24}$$

Proof. For the sake of notational simplicity, let us denote $x_m = P(u = m|H_0)$ and $y_m =$

$P(u = m|H_1)$. Similarly, $\tilde{x}_m = P(v = m|H_0)$ and $\tilde{y}_m = P(v = m|H_1)$.

Rewriting Equation (4.3) in our new notation, we have

$$\tilde{x}_m = \alpha \sum_{l \neq m} p x_l + (1 - \alpha(M-1)p)x_m = \alpha p + (1 - M\alpha p)x_m \tag{4.25}$$

and

$$\tilde{y}_m = \alpha \sum_{l \neq m} p y_l + (1 - \alpha(M-1)p)y_m = \alpha p + (1 - M\alpha p)y_m. \tag{4.26}$$

Therefore, the KLD at the FC can be rewritten as

$$D_{FC} = \sum_{m=1}^{M} \tilde{x}_m \log\left(\frac{\tilde{x}_m}{\tilde{y}_m}\right). \tag{4.27}$$

On partially differentiating D_{FC} with respect to p, we have

$$\begin{aligned}
\frac{\partial D_{FC}}{\partial p} &= \frac{\partial}{\partial p} \sum_{m=1}^{M} \tilde{x}_m \log\left(\frac{\tilde{x}_m}{\tilde{y}_m}\right) \\
&= \alpha \sum_{m=1}^{M} \left[(1 - Mx_m)\left(1 + \log\frac{\tilde{x}_m}{\tilde{y}_m}\right) - (1 - My_m)\frac{\tilde{x}_m}{\tilde{y}_m}\right] \\
&= \alpha \sum_{m=1}^{M}(1 - Mx_m) + \alpha \sum_{m=1}^{M}(1 - Mx_m)\log\frac{\tilde{x}_m}{\tilde{y}_m} - \alpha \sum_{m=1}^{M}(1 - My_m)\frac{\tilde{x}_m}{\tilde{y}_m}.
\end{aligned} \tag{4.28}$$

Consider the first term in the RHS of Equation (4.28). Note that, since $\mathbf{x} = \{x_1, \cdots, x_M\}$ is a probability mass function, we have

$$\sum_{m=1}^{M}(1 - Mx_m) = M - M\sum_{m=1}^{M} x_m = M - M = 0.$$

Therefore, Equation (4.28) reduces to

$$\frac{\partial D_{FC}}{\partial p} = \alpha \sum_{m=1}^{M} (1 - Mx_m) \log \frac{\tilde{x}_m}{\tilde{y}_m} - \alpha \sum_{m=1}^{M} (1 - My_m) \frac{\tilde{x}_m}{\tilde{y}_m}. \tag{4.29}$$

Rearranging the terms in Equation (4.29), we have

$$\frac{\partial D_{FC}}{\partial p} = \alpha \sum_{m=1}^{M} \left[\log \frac{\tilde{x}_m}{\tilde{y}_m} - \frac{\tilde{x}_m}{\tilde{y}_m} \right] - \alpha M \sum_{m=1}^{M} x_m \log \frac{\tilde{x}_m}{\tilde{y}_m} + \alpha M \sum_{m=1}^{M} y_m \frac{\tilde{x}_m}{\tilde{y}_m}. \tag{4.30}$$

Let us denote the first term as T_1. In other words,

$$T_1 = \alpha \sum_{m=1}^{M} \left[\log \frac{\tilde{x}_m}{\tilde{y}_m} - \frac{\tilde{x}_m}{\tilde{y}_m} \right].$$

Let us now focus our attention on the other terms in the RHS of Equation (4.30). Substituting Equations (4.25) and (4.26) in the second and third terms of the RHS of Equation (4.30), we have

$$\begin{aligned} \frac{\partial D_{FC}}{\partial p} &= T_1 - \frac{M\alpha}{1 - M\alpha p} \sum_{m=1}^{M} (\tilde{x}_m - \alpha p) \log \frac{\tilde{x}_m}{\tilde{y}_m} + \frac{M\alpha}{1 - M\alpha p} \sum_{m=1}^{M} (\tilde{y}_m - \alpha p) \frac{\tilde{x}_m}{\tilde{y}_m} \\ &= T_1 - \frac{M\alpha}{1 - M\alpha p} D(\tilde{\mathbf{x}} || \tilde{\mathbf{y}}) + \frac{M\alpha}{1 - M\alpha p} \left\{ \sum_{m=1}^{M} \alpha p \log \frac{\tilde{x}_m}{\tilde{y}_m} - \sum_{m=1}^{M} \alpha p \frac{\tilde{x}_m}{\tilde{y}_m} + \sum_{m=1}^{M} \tilde{x}_m \right\}, \end{aligned} \tag{4.31}$$

where $D(\tilde{\mathbf{x}} || \tilde{\mathbf{y}})$ is the KLD between $\tilde{\mathbf{x}}$ and $\tilde{\mathbf{y}}$ and is, therefore, non-negative. Also, note that in Equation (4.31), since $\tilde{\mathbf{x}} = \{\tilde{x}_1, \cdots, \tilde{x}_M\}$ is a probability mass function, $\sum_{m=1}^{M} \hat{x}_m = 1$.

Therefore, Equation (4.31) reduces to

$$\frac{\partial D_{FC}}{\partial p} = T_1 - \frac{M\alpha}{1 - M\alpha p} D(\tilde{\mathbf{x}} || \tilde{\mathbf{y}}) + \frac{M\alpha}{1 - M\alpha p} + \frac{M\alpha^2 p}{1 - M\alpha p} \sum_{m=1}^{M} \left[\log \frac{\tilde{x}_m}{\tilde{y}_m} - \frac{\tilde{x}_m}{\tilde{y}_m} \right]. \tag{4.32}$$

Note that the last term in the RHS of Equation (4.32),

$$\frac{M\alpha^2 p}{1-M\alpha p}\sum_{m=1}^{M}\left[\log\frac{\tilde{x}_m}{\tilde{y}_m}-\frac{\tilde{x}_m}{\tilde{y}_m}\right]=\frac{M\alpha p}{1-M\alpha p}T_1.$$

In other words,

$$\begin{aligned}\frac{\partial D_{FC}}{\partial p} &= \left(1+\frac{M\alpha p}{1-M\alpha p}\right)T_1-\frac{M\alpha}{1-M\alpha p}D(\tilde{\mathbf{x}}||\tilde{\mathbf{y}})+\frac{M\alpha}{1-M\alpha p} \\ &= \frac{1}{1-M\alpha p}T_1-\frac{M\alpha}{1-M\alpha p}D(\tilde{\mathbf{x}}||\tilde{\mathbf{y}})+\frac{M\alpha}{1-M\alpha p}.\end{aligned} \tag{4.33}$$

Rearranging the terms in Equation (4.33) and expanding T_1, we have

$$\begin{aligned}\frac{\partial D_{FC}}{\partial p} &= -\frac{M\alpha}{1-M\alpha p}D(\tilde{\mathbf{x}}||\tilde{\mathbf{y}})+\frac{M\alpha}{1-M\alpha p}+\frac{\alpha}{1-M\alpha p}\sum_{m=1}^{M}\left[\log\frac{\tilde{x}_m}{\tilde{y}_m}-\frac{\tilde{x}_m}{\tilde{y}_m}\right] \\ &= -\frac{M\alpha}{1-M\alpha p}D(\tilde{\mathbf{x}}||\tilde{\mathbf{y}})+\frac{\alpha}{1-M\alpha p}\sum_{m=1}^{M}\left[\log\frac{\tilde{x}_m}{\tilde{y}_m}-\left(\frac{\tilde{x}_m}{\tilde{y}_m}-1\right)\right].\end{aligned} \tag{4.34}$$

Since $\log x \leq x-1$ for all x, we find that the second term in the RHS of Equation (4.29) is negative. Therefore, we have

$$\frac{\partial D_{FC}}{\partial p}\leq 0. \tag{4.35}$$

Since D_{FC} is a non-increasing function of p, the optimal p, p^*, takes the maximum value $1/(M-1)$. □

Note that this solution is of particular interest to the Byzantine attacker since the solution does not require any knowledge about the sensor network design. Also, the attacker's strategy is very simple to implement.

Numerical Results

For illustration purposes, let us consider the following example, where the inference network is deployed to aid the opportunistic spectrum access for a cognitive radio network (CRN). In other words, the CRs are sensing a licensed spectrum band to find the vacant band for the operation of the CRN.

Let the observation model at the i^{th} sensor be defined as follows.

$$r_i = s(\theta) + n_i, \tag{4.36}$$

where $\theta \in \{0, 1\}$, $s(\theta) = \mu \cdot (-1)^{1+\theta}$ is a BPSK-modulated symbol transmitted by the licensed (or the primary) user transmitter, and the noise n_i is the AWGN at the i^{th} sensor with probability distribution $\mathcal{N}(0, \sigma^2)$.

Therefore, the conditional distribution of r_i under H_0 and H_1 can be given as $\mathcal{N}(-\mu, \sigma^2)$ and $\mathcal{N}(\mu, \sigma^2)$ respectively. The range of r_i spans the entire real line ($\mathbb{R}$). However, we assume that the quantizer restricts the support by limiting the range of output values to a smaller range, say $[-A, A]$. This parameter A is called the *overloading* parameter [51] because the choice of A dictates the amount of overloading distortion caused by the quantizer. Within this restricted range of observations, we assume a uniform quantizer with a step size (called the *granularity* parameter) given by $\Delta = \frac{2}{M-2}$, which dictates the granularity distortion of the quantizer. In other words, the observation r_i is quantized using the following quantizer:

$$u_i = \begin{cases} 0; & \text{if } -\infty < r_i \leq \lambda_1 \\ 1; & \text{if } \lambda_1 < r_i \leq \lambda_2 \\ \vdots & \\ M-1; & \text{if } \lambda_{M-1} < r_i \leq \infty \end{cases}, \tag{4.37}$$

where

$$\lambda_i = A \cdot \left[\frac{2(i-1)}{M-2} - 1 \right].$$

Note that, $\lambda_1 = -A$ and $\lambda_{M-1} = A$ represent the restricted range of the quantizer, as discussed earlier. The i^{th} sensor transmits a symbol v_i to the FC, where $v_i = u_i$ if it is honest. In the case of the i^{th} sensor being a Byzantine node, the decision u_i is modified into v_i using the flipping probability matrix $\mathbb{P}$ as given in Equation (4.6).

Although the performance of a given sensor network is quantified by the probability of error at the FC, we use a surrogate metric, as described earlier, called the KLD at the FC (Refer to Equation (4.21)) for the sake of tractability. In an asymptotic sense, Stein's Lemma [13] states that the KLD is the rate at which the probability of missed detection converges to zero under a constrained probability of false alarm. Therefore, in our numerical results, we present how KLD at the FC varies with the fraction of Byzantine nodes α, in the network.

For the above sensor network, we assume that $\mu = 1$, $\sigma^2 = 1$ and $A = 2$. In Figure 4.3, we plot the contribution of each sensor in terms of KLD at the FC as a function of α, for 1-bit, 2-bit, 3-bit and 4-bit quantizations, i.e., $M = 2$, 4, 8 and 16 respectively, at the sensors. As per our intuition, we observe an improvement in both the detection performance (KLD) as well as security performance (α_{blind}). Therefore, for a given α, the Byzantine attack can be mitigated by employing finer quantization at the sensors. Of course, the best that the designer can do is to let the sensors transmit unquantized data to the FC, whether in the form of observation samples or their sufficient statistic (likelihood ratio). In this case, we can see that $\alpha_{blind} = 1$, since $\lim_{M \to \infty} \frac{M-1}{M} = 1$.

4.5.2 Distributed Estimation

In this section, we consider the problem of estimating a scalar parameter of interest, denoted by $\theta \in \mathbb{R}$, in a distributed sensor network. As described in the system model, we

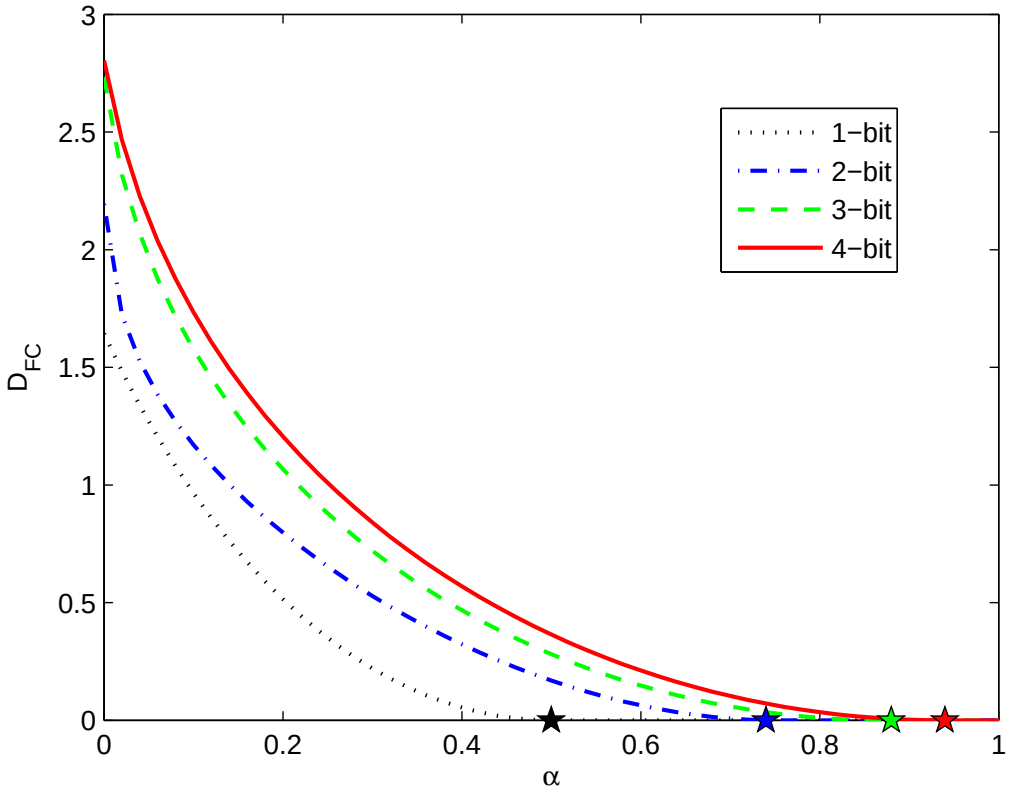


Figure 4.3: Contribution of a sensor to the overall KLD at the fusion center as a function of α, for different number of quantization levels. The pentagrams on the x-axis correspond to the α_{blind} for 1-bit, 2-bit, 3-bit and 4-bit quantizations respectively from left to right.

assume that the i^{th} sensor quantizes its observation r_i into an M-ary symbol u_i, and transmits v_i to the FC. If the i^{th} node is honest, then $v_i = u_i$. Otherwise, we assume that the sensor is compromised and flips u_i into v_i using a flipping probability matrix $\mathbb{P}$. Under the assumption that the FC receives the symbols $\mathbf{v}$ over an ideal channel, the estimation performance at the FC depends on the probability mass function $P(\mathbf{v}|\theta)$.

The performance of a distributed estimation network can be expressed in terms of the mean-squared error, defined as $\mathbb{E}\left[(\hat{\theta}-\theta)^2\right]$. In the case of unbiased estimators, this mean-squared error is lower bounded by the *Cramer-Rao lower bound* (CRLB) [53], which provides a benchmark for the design of an estimator at the FC. We present this result in Equation (4.38):

$$\mathbb{E}\left[(\hat{\theta}(\mathbf{v})-\theta)^2\right] \geq \frac{1}{I_{FC}}, \tag{4.38}$$

where

$$I_{FC} = \mathbb{E}\left[\left(\frac{\partial \log P(\mathbf{v},\theta)}{\partial \theta}\right)^2\right]. \tag{4.39}$$

The term I_{FC} is well known as the Fisher information (FI), and is, therefore, a performance metric that captures the performance of the optimal estimator at the FC. Note that, as shown in Equation (4.40), I_{FC} can be further decomposed into two parts, one corresponding to the prior knowledge about θ at the FC, and the other (denoted as J_{FC}) representing the information about θ, in the sensor transmissions $\mathbf{v}$:

$$I_{FC} = J_{FC} + \mathbb{E}\left[\left(\frac{\partial \log p(\theta)}{\partial \theta}\right)^2\right], \tag{4.40}$$

where

$$J_{FC} = \mathbb{E}\left[\left(\frac{\partial \log P(\mathbf{v}|\theta)}{\partial \theta}\right)^2\right]. \tag{4.41}$$

In most cases, a closed form expression for the mean-squared error is intractable and, therefore, conditional Fisher information (FI) is used as a surrogate metric to quantify the performance of a distributed estimation network. In this chapter, we also use conditional FI of the received data $\mathbf{v}$ as the performance metric. Since the sensor observations are conditionally independent resulting in independent $\mathbf{v}$, we denote the conditional FI as $\mathbb{J}_{FC}$ and is defined as follows:

$$\mathbb{J}_{FC} = N J_{FC}, \tag{4.42}$$

where

$$J_{FC} = \mathbb{E}\left[\frac{\partial}{\partial \theta}\log P(\mathbf{v}|\theta)\right]^2 = -\mathbb{E}\left[\frac{\partial^2}{\partial \theta^2}\log P(\mathbf{v}|\theta)\right]. \tag{4.43}$$

Following the same approach as in Section 4.5.1, we consider the problem of finding an

optimal resource-constrained Byzantine attack when $\alpha < \alpha_{blind}$, by finding the symmetric transition matrix $\mathbb{P}$ that minimizes the conditional FI at the FC. This can be formulated as follows.

Problem 4.2. *Given the value of α, determine the optimal $\mathbb{P}$ within a space of highly symmetric row-stochastic matrices, as given in Equation* (4.23)*, such that*

$$\begin{array}{ll} \underset{p}{\textit{minimize}} & J_{FC} \\ \textit{subject to} & 0 \leq p \leq \dfrac{1}{M-1} \end{array}.$$

Theorem 4.5 presents the optimal flipping probability that provides a solution to Problem 4.2. Note that this result is independent of the design of the sensor network and, therefore, can be employed when the Byzantine has no knowledge about the network.

Theorem 4.5. *Given a fixed $\alpha < \dfrac{M-1}{M}$, the flipping probability p that optimizes $\mathbb{P}$ over a space of highly symmetric row-stochastic matrices, as given in Equation* (4.23)*, by minimizing J_{FC} is given by*

$$p^* = \frac{1}{M-1}.$$

Proof. For the sake of notational simplicity, we let $z_m = P(u = m|\theta)$. Similarly, $\tilde{z}_m = P(v = m|\theta)$. Using this notation in Equation (4.43), we have

$$\begin{aligned} J_{FC} &= \sum_{m=1}^{M} P(v=m|\theta) \left(\frac{\partial \log P(v=m|\theta)}{\partial \theta} \right)^2 = \sum_{m=1}^{M} \tilde{z}_m \left(\frac{\partial \log \tilde{z}_m}{\partial \theta} \right)^2 \\ &= (1 - M\alpha p)^2 \sum_{m=1}^{M} \frac{1}{\tilde{z}_m} \left(\frac{\partial z_m}{\partial \theta} \right)^2. \end{aligned} \tag{4.44}$$

Partially differentiating J_{FC} with respect to p, we have

$$\begin{aligned}
\frac{\partial J_{FC}}{\partial p} &= 2(1-M\alpha p)(-M\alpha)\sum_{m=1}^{M}\frac{1}{\tilde{z}_m}\left(\frac{\partial z_m}{\partial\theta}\right)^2 \\
&\quad + (1-M\alpha p)^2\sum_{m=1}^{M}\left(-\frac{1}{\tilde{z}_m^2}\right)(\alpha - M\alpha z_m)\left(\frac{\partial z_m}{\partial\theta}\right)^2 \\
&= -(1-M\alpha p)\left[2M\alpha\sum_{m=1}^{M}\tilde{z}_m\left(\frac{1}{\tilde{z}_m}\frac{\partial z_m}{\partial\theta}\right)^2 + (1-M\alpha p)\sum_{m=1}^{M}\alpha\left(\frac{1}{\tilde{z}_m}\frac{\partial z_m}{\partial\theta}\right)^2\right. \\
&\quad \left. -(1-M\alpha p)\sum_{m=1}^{M}M\alpha z_m\left(\frac{1}{\tilde{z}_m}\frac{\partial z_m}{\partial\theta}\right)^2\right] \\
&= -(1-M\alpha p)\left[\alpha(1-M\alpha p)\sum_{m=1}^{M}\left(\frac{1}{\tilde{z}_m}\frac{\partial z_m}{\partial\theta}\right)^2 + M\alpha(1+M\alpha p)\sum_{m=1}^{M}z_m\left(\frac{1}{\tilde{z}_m}\frac{\partial z_m}{\partial\theta}\right)^2\right].
\end{aligned} \tag{4.45}$$

In Equation (4.45), we have a negative term multiplied by a non-negative term, and hence we have

$$\frac{\partial J_{FC}}{\partial p} \leq 0. \tag{4.46}$$

Since J_{FC} is a non-increasing function of p, $p^* = \dfrac{1}{M-1}$, being the maximum value, is the optimal solution to Problem 4.2. □

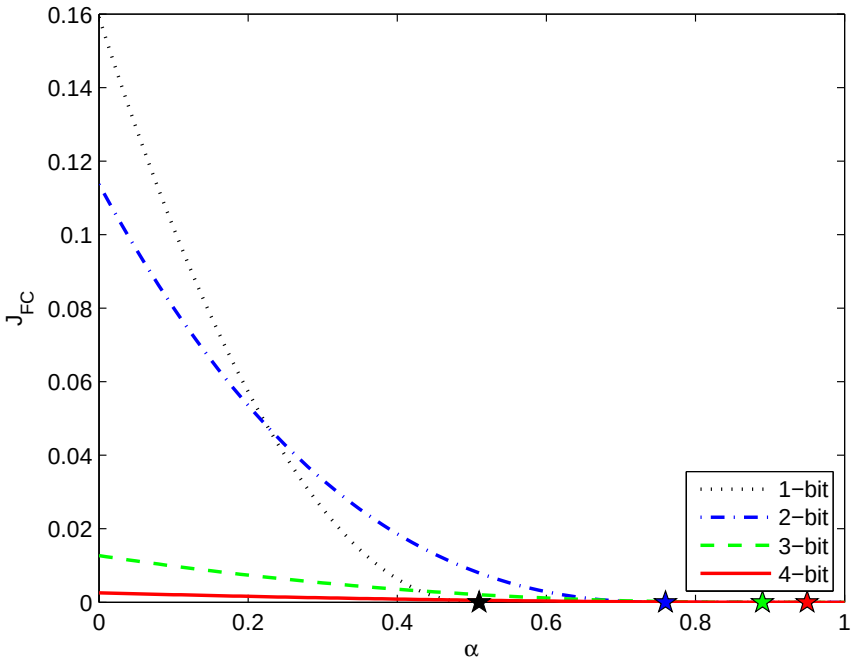
Numerical Results

As an illustrative example, we consider the problem of estimating $\theta = 1$ at the FC based on all the sensors' transmitted messages. Let the observation model at the i^{th} sensor be defined as follows:

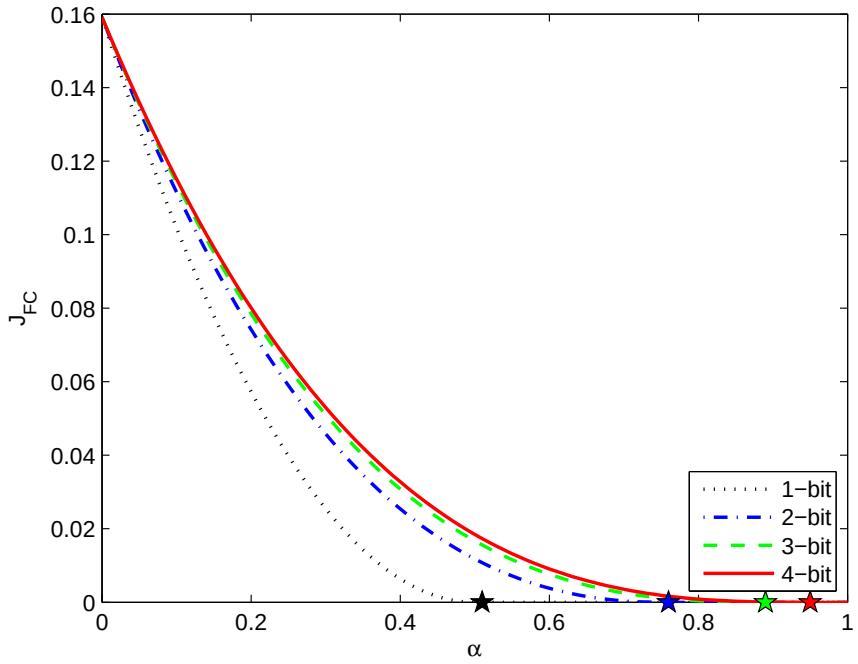
$$r_i = \theta + n_i, \tag{4.47}$$

where the noise n_i is the AWGN at the i^{th} sensor with probability distribution $\mathcal{N}(0, \sigma^2)$. The sensors employ the same quantizer as the one presented in Equation (4.37). The quantized symbol, denoted as u_i at the i^{th} sensor, is then modified into v_i using the flipping probability matrix $\mathbb{P}$, as given in Equation (4.6).

Figure 4.4 plots the conditional FI corresponding to one sensor, for different values of α and M, when the uniform quantizer is centered around the true value of θ. Note that as SNR increases ($\sigma \to 0$), we observe that it is better for the network to perform as much finer quantization as possible to mitigate the Byzantine attackers. On the other hand, if SNR is low, coarse quantization performs better for lower values of α. This phenomenon of coarse quantization performing better under low SNR scenarios, can be attributed to the fact that more noise gets filtered as the quantization gets coarser (decreasing M) than the signal itself. On the other hand, in the case of high SNR, since the signal level is high, coarse quantization cancels out the signal component significantly, thereby resulting in a degradation in performance.



(a) Low SNR case: $\sigma = 1$



(b) High SNR case: $\sigma = 0.01$

Figure 4.4: Contribution of a sensor to the overall conditional FI at the FC as a function of α, for different number of quantization levels when $\theta = 0$ and $A = 2$. The pentagrams on the x-axis correspond to the α_{blind} for 1-bit, 2-bit, 3-bit and 4-bit quantizations respectively from left to right.

4.6 Reputation-based Detection of Byzantine Nodes

Given that the distributed inference network is under Byzantine attack, we showed that the performance of the network can be improved by increasing the quantization alphabet size of the sensors. Obviously, in a bandwidth-constrained distributed inference network, the sensors can only transmit with the maximum possible M, which is finite. In this section, we assume that the network cannot further increase the quantization alphabet size due to this bandwidth constraint. Therefore, we present a reputation-based Byzantine identification/mitigation scheme, which is an extension of the one proposed by Rawat *et al.* in [52], in order to improve the inference performance of the network.

4.6.1 Reputation-Tagging at the Sensors

As proposed by Rawat *et al.* in [52], the FC identifies the Byzantine nodes by iteratively updating a reputation-tag for each node as time progresses. We extend the scheme to include fine quantization scenarios, i.e., $M > 2$, and analyze its performance through simulation results.

As mentioned earlier in the chapter, the FC receives a vector $\mathbf{v}$ of received symbols from the sensors and fuses them to yield a global decision, denoted as $\hat{\theta}$. We assume that the observation model is known to the network designer, and is given as follows:

$$r_i = f_i(\theta) + n_i, \tag{4.48}$$

where $f_i(\cdot)$ denotes the known observation model. We denote the quantization rule employed at the sensor as γ. Therefore, the quantized message at the sensor is given by $u_i = \gamma(r_i)$. As discussed earlier, the i^{th} sensor flips u_i into v_i using a flipping probability matrix $\mathbb{P}$. Since the FC makes a global inference $\hat{\theta}$, it can calculate the squared-deviation

d_i of each sensor from the expected message that it is to nominally transmit as follows:

$$d_i = \left(\gamma^{-1}(v_i) - f_i(\hat{\theta})\right)^2, \tag{4.49}$$

where $\gamma^{-1}(v_i)$ is the inverse of the sensor quantizer $\gamma(v_i)$ and it is assumed to be the centroid of the corresponding decision region of the quantizer v_i.

Note that v_i is the received symbol which characterizes the behavior (honest or Byzantine) of the i^{th} sensor, while $f_i(\hat{\theta})$ is the signal that the FC expects the sensor to observe. If the i^{th} sensor is honest, we expect the mean of d_i to be small. On the other hand, if the i^{th} sensor is a compromised node, then the mean of d_i is expected to be large. Therefore, we accumulate the squared-deviations $\mathbf{d_i} = \{d_i(1), \cdots, d_i(T)\}$ over T time intervals and compute a reputation tag $\Lambda_i(\mathbf{d_i})$, as a time-average for the i^{th} node as follows:

$$\Lambda_i = \frac{1}{T}\sum_{t=1}^{T} d_i(t). \tag{4.50}$$

The i^{th} sensor is declared honest/Byzantine using the following threshold-based tagging rule

$$\Lambda_i \underset{\text{Honest}}{\overset{\text{Byzantine}}{\gtrless}} \eta. \tag{4.51}$$

The performance of the above tagging rule depends strongly on the choice of η. Note that the threshold η should be chosen based on two factors. Firstly, η should be chosen in such a way that the probability with which a malicious node is tagged Byzantine is high. Higher the value of η, lower is the chance of tagging a node to be Byzantine and vice-versa. This results in a tradeoff between the probability of detecting a Byzantine vs. the probability of falsely tagging an honest node as a Byzantine. Secondly, the value of M also plays a role in the choice of η, and therefore, the performance of the tagging rule. We illustrate this phenomenon in our simulation results.

4.6.2 Optimal Choice of the Tagging Threshold as $T \to \infty$

In this chapter, we denote the true type of the i^{th} node as $\mathscr{T}_i$, where $\mathscr{T}_i = H$ corresponds to honest behavior, while $\mathscr{T}_i = B$ corresponds to Byzantine behavior, for all $i = 1, \cdots, N$. Earlier, in this section, we presented Equation (4.51) which allows us to make inferences about the true type. But, the performance of the Byzantine tagging scheme corresponding the i^{th} sensor is quantified by the conditional probabilities $P(\Lambda_i \geq \eta | \mathscr{T}_i = \mathscr{T})$, for both $\mathscr{T} = H, B$. In order to find the optimal choice of η in Equation (4.51), we continue with the Neyman-Pearson framework even in the context of Byzantine identification, where the goal is to maximize $P(\Lambda_i \geq \eta | \mathscr{T}_i = B)$, subject to the condition that $P(\Lambda_i \geq \eta | \mathscr{T}_i = H) \leq \xi$.

To find these two conditional probabilities $P(\Lambda_i \geq \eta | \mathscr{T}_i = H)$ and $P(\Lambda_i \geq \eta | \mathscr{T}_i = B)$, we need a closed form expression of the conditional distributions, $P(\Lambda_i | \mathscr{T}_i = H)$ and $P(\Lambda_i | \mathscr{T}_i = B)$ respectively. In practice, where T is finite, it is intractable to determine the conditional distribution of Λ_i, which is necessary to come up with the optimal choice of η. Therefore, in this section, we assume that $T \to \infty$ and present an asymptotic choice of the tagging threshold η used in Equation (4.51).

As $T \to \infty$, since $d_i(t)$ is independent across $t = 1, \cdots, T$, due to central-limit theorem, $(\Lambda_i | \mathscr{T}_i = \mathscr{T}) \sim \mathscr{N}(\mu_{i,\mathscr{T}}, \sigma_{i,\mathscr{T}})$, where

$$\begin{aligned} \mu_{i,\mathscr{T}} &= \mathbb{E}(\Lambda_i \mid \mathscr{T}_i = \mathscr{T}) \\ &= \mathbb{E}\left[\left(\gamma^{-1}(v_i(t)) - \hat{\theta}(t) \right)^2 \mid \mathscr{T}_i = \mathscr{T} \right] \end{aligned} \tag{4.52}$$

and

$$\begin{aligned} \sigma^2_{i,\mathscr{T}} &= \mathrm{Var}(\Lambda_i \mid \mathscr{T}_i = \mathscr{T}) \\ &= \frac{1}{T} \mathrm{Var}\left[\left(\gamma^{-1}(v_i(t)) - \hat{\theta}(t) \right)^2 \mid \mathscr{T}_i = \mathscr{T} \right] \end{aligned} . \tag{4.53}$$

In this section, we do not present the final form of $\mu_{i,\mathscr{T}}$ and $\sigma_{i,\mathscr{T}}$ in order to preserve

generality. Assuming that $v_i(t)$ is independent across sensors as well as time, the moments of d_i can be computed for any given FC's inference $\hat{\theta}(t)$ at time t about a given phenomenon. Although the final form of $\mu_{i,\mathscr{T}}$ and $\sigma_{i,\mathscr{T}}$ is not presented, since $d_i(t)$ is a function of $\mathbf{v}$, we present the conditional probability of $(v_j|\mathscr{T}_i = \mathscr{T})$ in Equation (4.54), which is necessary for the computation of $\mu_{i,\mathscr{T}}$ and $\sigma_{i,\mathscr{T}}$.

$$P(v_j|\mathscr{T}_i = \mathscr{T}) = \int P(v_j|\theta, \mathscr{T}_i = \mathscr{T})p(\theta)d\theta, \tag{4.54}$$

where $P(v_j|\theta, \mathscr{T}_i = \mathscr{T})$ can be calculated as follows:

$$P(v_j = m|\theta, \mathscr{T}_i = H) = \begin{cases} P(u_j = m|\theta), & \text{if } j = i \\ (1 - \pi_{BH})P(u_j = m|\theta) + & \\ \quad \pi_{BH} \sum_{k=1}^{M} p_{km} P(u_j = k|\theta), & \text{if } j \neq i \end{cases} \tag{4.55}$$

and

$$P(v_j = m|\theta, \mathscr{T}_i = B) = \begin{cases} \sum_{k=1}^{M} p_{km} P(u_j = k|\theta), & \text{if } j = i \\ (1 - \pi_{BB})P(u_j = m|\theta) + & \\ \quad \pi_{BB} \sum_{k=1}^{M} p_{km} P(u_j = k|\theta), & \text{if } j \neq i \end{cases}, \tag{4.56}$$

where $\pi_{BH} = P(\mathscr{T}_j = B|\mathscr{T}_i = H)$ and $\pi_{BB} = P(\mathscr{T}_j = B|\mathscr{T}_i = B)$ are conditional probabilities of the j^{th} node's type, given the type of the i^{th} node. Since there are α fraction of nodes in the network, given that the FC knows the type of i^{th} node as H, the conditional probability of the j^{th} node belonging to a type $\mathscr{T}$ is given by $\pi_{BH} = \dfrac{N\alpha}{N-1}$ and $\pi_{BB} =$

$\frac{N\alpha - 1}{N - 1}$.

Given the conditional distributions $P(\Lambda_i|\mathscr{T}_i = H)$ and $P(\Lambda_i|\mathscr{T}_i = B)$, we find the performance of the Byzantine identification scheme as follows:

$$\begin{aligned} P(\Lambda_i \geq \eta|\mathscr{T}_i = H) &= Q\left(\frac{\eta - \mu_{i,H}}{\sigma_{i,H}}\right) \\ P(\Lambda_i \geq \eta|\mathscr{T}_i = B) &= Q\left(\frac{\eta - \mu_{i,B}}{\sigma_{i,B}}\right) \end{aligned} . \tag{4.57}$$

Under the NP framework, the optimal η can be chosen by letting $P(\Lambda_i \geq \eta|i = H) = \beta$, when Λ_i is normally distributed conditioned on the true type of a given node. In other words,

$$Q\left(\frac{\eta - \mu_{i,H}}{\sigma_{i,H}}\right) = \xi \tag{4.58}$$

or equivalently,

$$\eta_{optimal} = \mu_{i,H} + \sigma_{i,H}Q^{-1}(\xi). \tag{4.59}$$

Note that, since $P(v_i|\mathscr{T}_i = H)$ is a function of α, it follows that both $\mu_{i,H}$ and $\sigma_{i,H}$ are functions of α. Although we do not provide a closed-form expression for η as a function of α, we provide the following example to portray how η varies with different values of α.

Example: Variation of η as a function of α

Consider a distributed estimation network with $N = 5$ identical nodes. Let the prior distribution of the true phenomenon θ be the uniform distribution $\mathscr{U}(0, 1)$. We assume that the sensing channel is an AWGN channel where the sensor observations is given by $r_i = \theta + n_i$. Therefore, the conditional distribution of the sensor observations is $\mathscr{N}(\theta, \sigma^2)$, when conditioned on θ. We assume that the sensors employ the quantizer rule shown in Equation (4.37) on their observations r_i. At the FC, we let $\gamma^{-1}(\cdot)$ be defined as the centroid function that returns $c_i = \frac{\lambda_{i-1} + \lambda_i}{2}$. Let $\hat{\theta} = \frac{1}{M}\sum_{i=1}^{N} \gamma^{-1}(v_i(t))$ be the fusion rule employed at the

FC to estimate θ.

Since the network comprises of identical nodes, without any loss of generality, we henceforth focus our attention on the reputation-based identification rule at sensor-1. Substituting the above mentioned fusion rule in the squared-deviation d_1 corresponding to sensor-1 in Equation (4.49), we have

$$\begin{aligned} d_1 &= \left(\gamma^{-1}(v_1) - \frac{1}{M}\sum_{i=1}^{5}\gamma^{-1}(v_i(t))\right)^2 \\ &= \left(\frac{M-1}{M}\gamma^{-1}(v_1) - \frac{1}{M}\sum_{i=2}^{5}\gamma^{-1}(v_i(t))\right)^2. \end{aligned} \tag{4.60}$$

Let us denote

$$\phi_{ij} = \mathbb{E}\left\{\left(\gamma^{-1}(v_i)\right)^j | \mathscr{T}_1 = H\right\} = \sum_{v_i=1}^{M}\left[\left(\gamma^{-1}(v_i)\right)^j P\left(v_i|\mathscr{T}_1 = H\right)\right],$$

for all $i = 1, \cdots, 5$ and $j = 1, 2, \cdots, \infty$. Here, $P\left(v_i|\mathscr{T}_1 = H\right)$ can be computed using Equation (4.55) as follows:

$$\begin{aligned} P\left(v_i = m|\mathscr{T}_1 = H\right) &= \int_{-\infty}^{\infty} P\left(v_i = m|\theta, \mathscr{T}_1 = H\right) p(\theta) d\theta \\ &= \int_{0}^{1} P\left(v_i = m|\theta, \mathscr{T}_1 = H\right) d\theta \\ &= \begin{cases} a_{1,m} & \text{if } i = 1 \\ \dfrac{N\alpha}{(N-1)(M-1)} + \left(1 - \dfrac{MN\alpha}{(N-1)(M-1)}\right) a_{i,m} & \text{otherwise.} \end{cases} \end{aligned} \tag{4.61}$$

where $a_{i,m} = \int_0^1 P\left(u_i = m|\theta\right) d\theta$, for all $i = 1, \cdots, N$. Note that, since all the nodes

in the network are identical, $P(u_i|\theta)$ is independent of the node-index i, and therefore, $\phi_{ij} = \phi_{2j}$, for all $i \neq 1$.

Thus, the conditional mean and variance, μ_{1H} and σ^2_{1H}, are given as follows for the special case of $N = 5$:

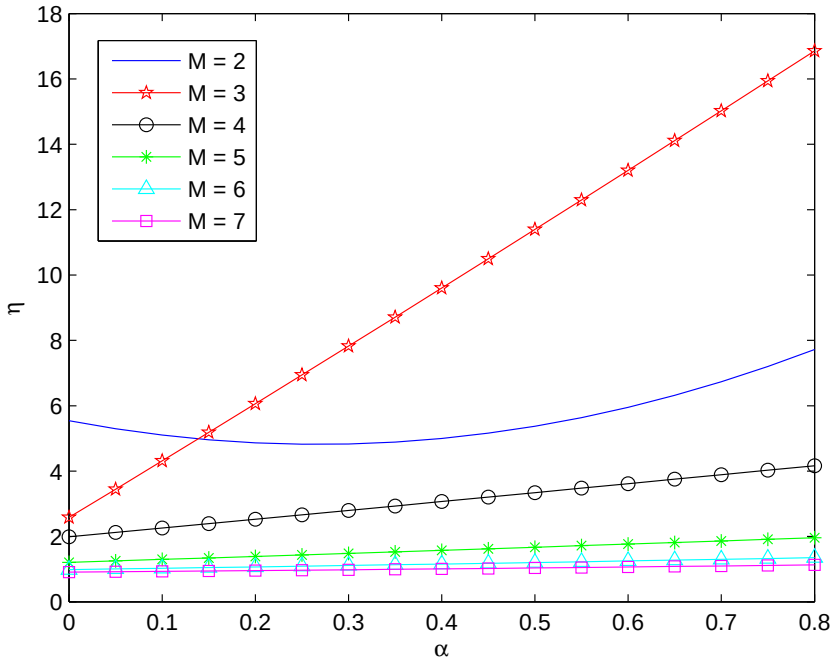
$$\begin{aligned}
\mu_{1H} &= \mathbb{E}\left[\left(\frac{M-1}{M}\gamma^{-1}(v_1) - \frac{1}{M}\sum_{i=2}^{5}\gamma^{-1}(v_i(t))\right)^2 \mid \mathscr{T}_i = H\right] \\
&= \frac{1}{M^2}\mathbb{E}\left[\left((M-1)\gamma^{-1}(v_1) - \sum_{i=2}^{5}\gamma^{-1}(v_i(t))\right)^2 \mid \mathscr{T}_i = H\right] \\
&= \frac{1}{M^2}\left[(M-1)^2\phi_{12} + 4\phi_{22} + 12\phi_{21}^2 - 8(M-1)\phi_{11}\phi_{21}\right]
\end{aligned} \tag{4.62}$$

and

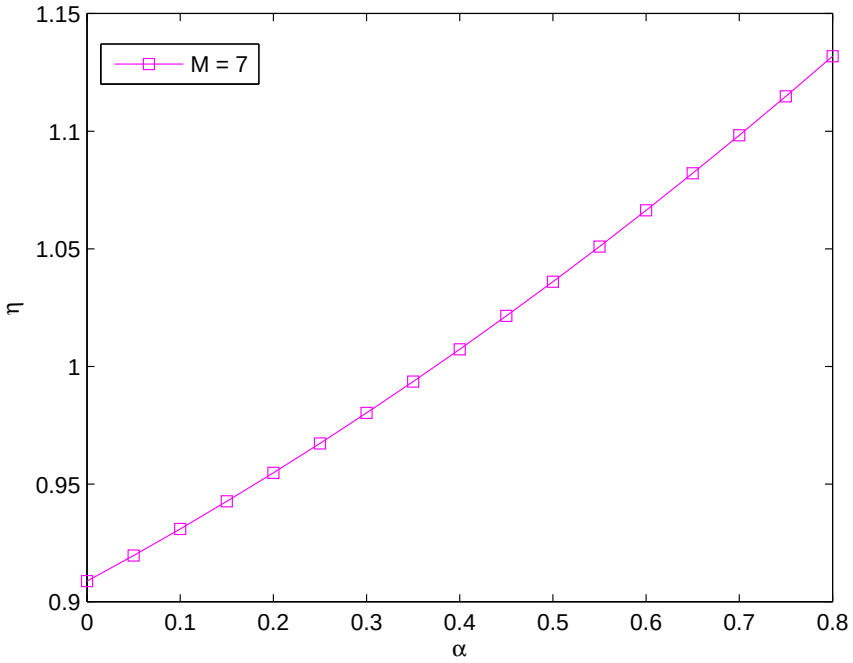
$$\begin{aligned}
\sigma^2_{1H} &= \frac{1}{T}\,\mathrm{Var}\left[\left(\gamma^{-1}(v_i(t)) - \hat{\theta}(t)\right)^2 \mid \mathscr{T}_i = H\right] \\
&= \frac{1}{T}\left\{\Delta - \mu^2_{1H}\right\},
\end{aligned} \tag{4.63}$$

where

$$\begin{aligned}
\Delta &= \mathbb{E}\left[\left(\frac{M-1}{M}\gamma^{-1}(v_1) - \frac{1}{M}\sum_{i=2}^{5}\gamma^{-1}(v_i(t))\right)^4 \mid \mathscr{T}_i = H\right] \\
&= \frac{1}{M^4}\left[(M-1)^4\phi_{14} - 16(M-1)^3\phi_{13}\phi_{21} + 6(M-1)^2\phi_{12}\{4\phi_{22} + 12\phi_{21}^2\}\right. \\
&\qquad -4(M-1)\phi_{11}(4\phi_{23} + 36\phi_{22}\phi_{21} + 24\phi_{21}^3) + 4\phi_{24} + 12\phi_{23}\phi_{21} \\
&\qquad \left. +36(\phi_{23}\phi_{21} + \phi_{22}^2 + 2\phi_{22}\phi_{21}^2) + 24(\phi_{21}^4 + 3\phi_{22}\phi_{21}^2)\right].
\end{aligned} \tag{4.64}$$



(a) $M = 2, \cdots, 7$



(b) $M = 7$

Figure 4.5: Variation of the optimal tagging threshold η (in the asymptotic sense, where $T \to \infty$) as a function of α

Thus, for $\xi = 0.01$, we compute the tagging threshold η numerically as shown in Equation (4.59), and plot the variation of η as a function of α in Figure 4.5. Note that, in our numerical results, we observe that the optimal choice of η is a convex function of α, where the curvature of the convexity decreases with increasing M. This can be clearly seen from Figure 4.5b, where we only plot the case of $M = 7$. We observe a similar behavior for all the other values of M, and therefore, present the case of $M = 7$ to illustrate the convex behavior of η. In other words, for very large values of M, the choice of η becomes independent of α, for any fixed $\alpha \leq \alpha_{blind}$.

4.6.3 Simulation Results

In order to illustrate the performance of the proposed reputation-based scheme, we consider a sensor network with a total of 100 sensors in the network, out of which 20 are Byzantine sensors. Let the sensor quantizers be given by Equation (4.37) and the fusion rule at the FC be the MAP rule, given as follows:

$$\sum_{i=1}^{N} \log\left(\frac{P(v_i|H_1)}{P(v_i|H_0)}\right) \underset{\hat{\theta}=0}{\overset{\hat{\theta}=1}{\gtrless}} \log\frac{p_0}{p_1}. \tag{4.65}$$

Figure 4.6 plots the rate of identification of the number of Byzantine nodes in the network for the proposed reputation-based scheme for different sizes of the quantization alphabet set. Note that the convergence rate deteriorates as M increases. This is due to the fact that the Byzantine nodes have increasing number of symbol options to flip to, because of which a greater number of time-samples are needed to identify the malicious behavior. In addition, we also simulate the evolution of mislabelling an honest node as a Byzantine node in time, and plot the probability of the occurrence of this event in Figure 4.7. Just as the convergence deteriorates with increasing M, we observe a similar behavior in the evolution of the probability of mislabelling honest nodes. Another important observation in Figure 4.7 is that the probability of mislabelling a node always converges to zero in time.

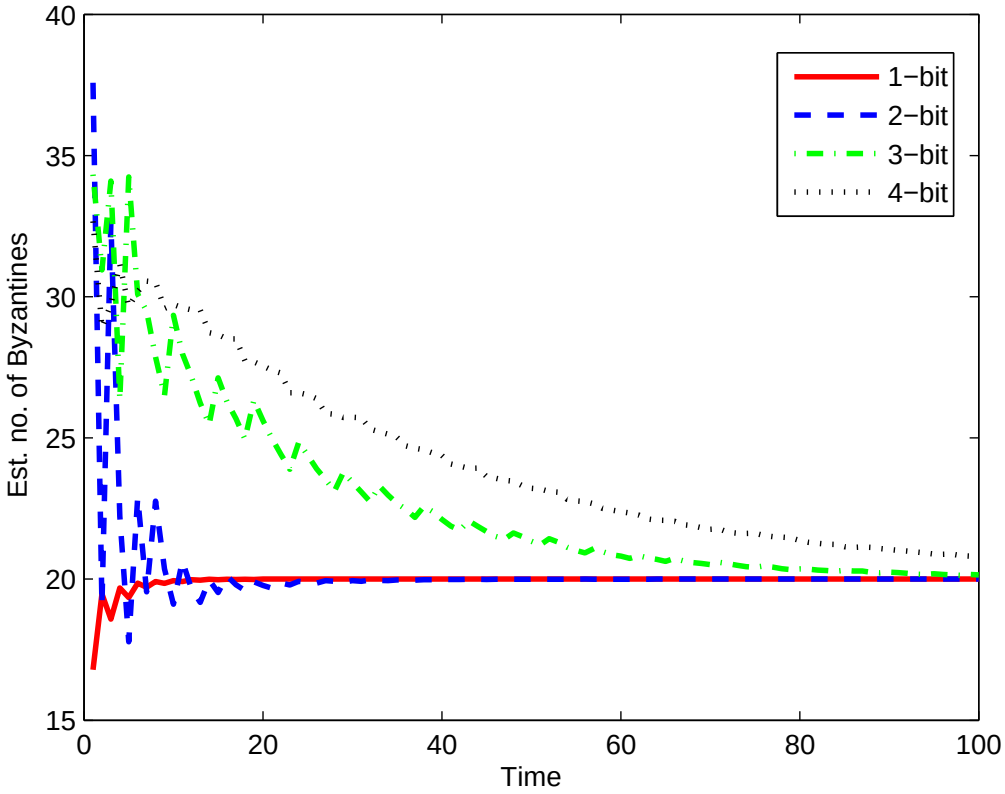


Figure 4.6: Rate of identification of the number of Byzantine nodes in time for different number of quantization levels

Similarly, we simulate the evolution of mislabelling a Byzantine node as an honest one in time in Figure 4.8. We observe similar convergence of the probability of mislabelling a Byzantine node as an honest node to zero, with a rate that decreases with increasing number of quantization levels, M. Therefore, Figures 4.6, 4.7 and 4.8 demonstrate that, after a sufficient amount of time, the reputation-based scheme always identifies the true behavior of a node within the network, with negligible number of mislabels.

4.7 Summary

In summary, we have modeled the problem of distributed inference with M-ary quantized data in the presence of Byzantine attacks, under the assumption that the attacker does not

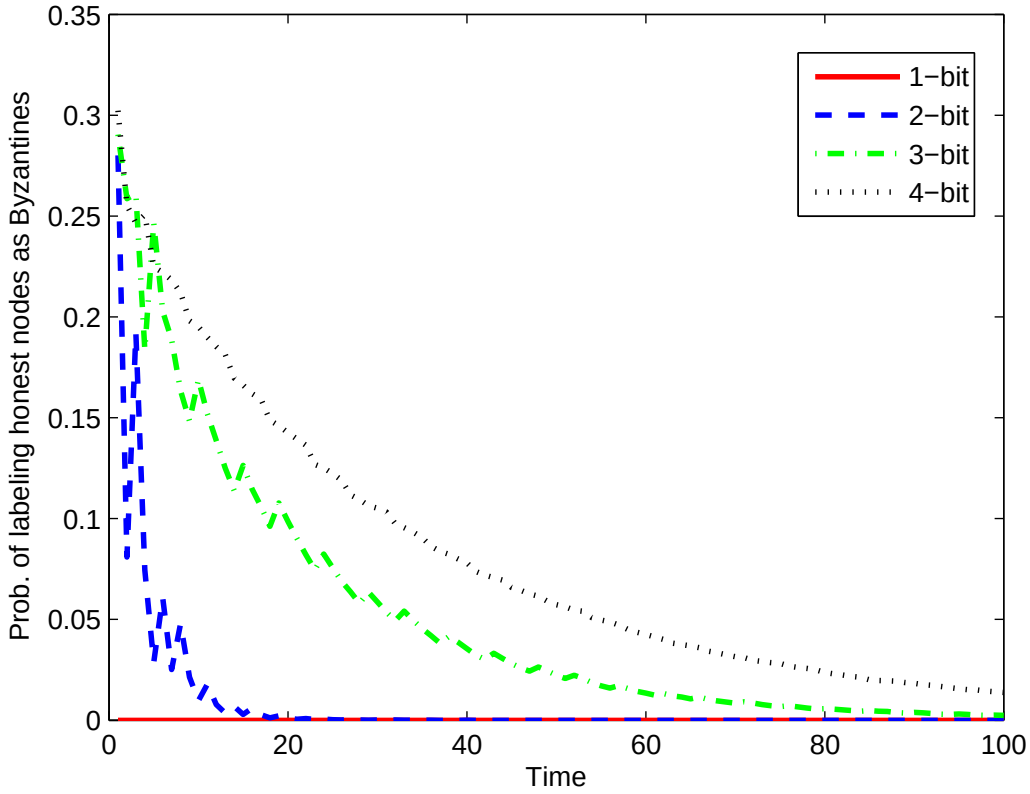


Figure 4.7: Evolution of the probability of mislabelling an honest node as a Byzantine in time for different number of quantization levels

have knowledge about either the true hypotheses or the quantization thresholds at the sensors. We found the optimal Byzantine attack that *blinds* the FC in the case of any inference task for both noiseless and noisy FC channels. We have also considered the problem of resource-constrained Byzantine attack ($\alpha < \alpha_{blind}$) for distributed detection and estimation in the presence of resource-constrained Byzantine attacker for the special case of highly symmetric attack strategies in the presence of noiseless channels at the FC. From the inference network's perspective, we have presented a mitigation scheme that identifies the Byzantine nodes through reputation-tagging. We have also shown how the optimal tagging threshold can be found when the time-window $T \to \infty$. Finally, we have also investigated the performance of our reputation-based scheme in our simulation results, and showed that our scheme always converges to finding all the compromised nodes, given sufficient amount

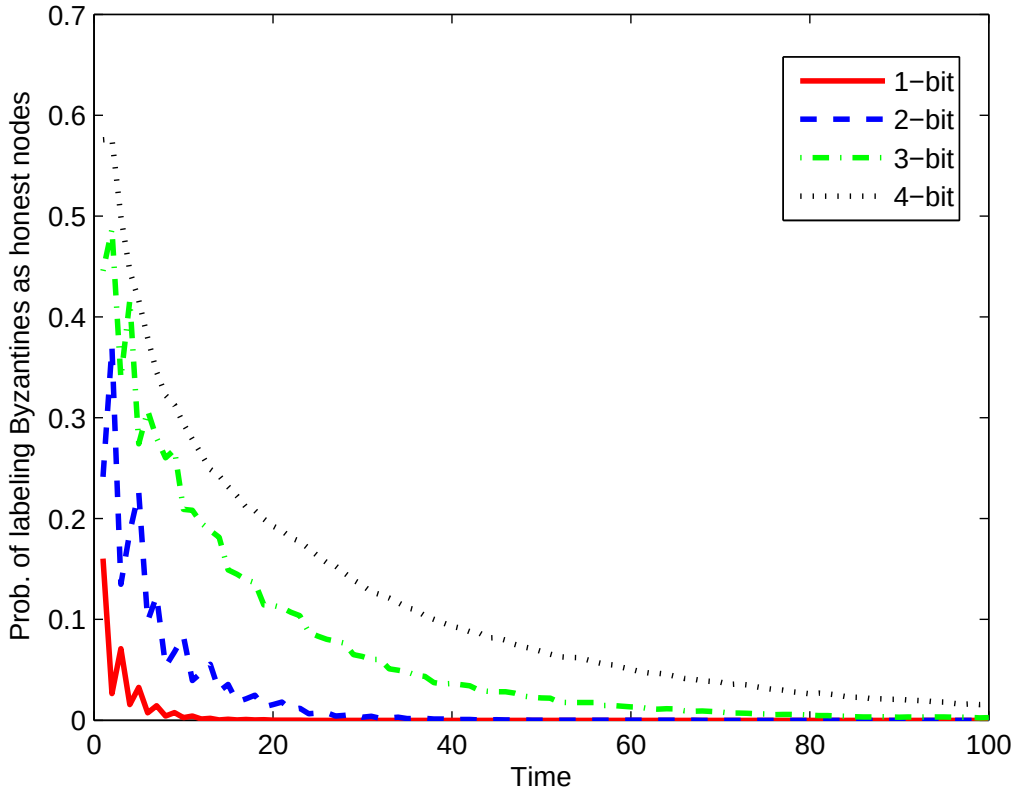


Figure 4.8: Evolution of the probability of mislabelling a Byzantine node as an honest node in time for different number of quantization levels

of time.

CHAPTER 5
JAMMING ATTACKS IN DISTRIBUTED DETECTION: POWER-ALLOCATION AND PLACEMENT

Interference has always been a nuisance in the design of any electronic system. Most of the past literature had addressed noise-like interference which disrupts the system unintentionally. In contrast, jamming attacks are designed to introduce interference intentionally so as to cause maximal degradation in their performance. In this chapter, we assume that the sensors use a multiple access channel (MAC) to communicate their messages to the FC. Given that there are fundamentally two types of channels in an inference network, namely the *sensing* channel and the *communication* channel (MAC), we consider a general jamming attack model which allows the jammer to distribute its energy across these two channels. Such attack models are particularly useful in some practical applications such as radar networks and cooperative spectrum sensing in cognitive radio (CR) networks where the sensing channel can be jammed using an electromagnetic signal[1].

[1] In other applications where the PoI does not emit an electromagnetic signal, we can let the energy deployed in the sensing channel to be zero.

5.1 Literature Survey

Jamming attacks have traditionally been addressed in communication systems, where several mitigation schemes have been proposed based on low probability of intercept (LPI) techniques such as spectrum spectrum technology [38, 50] and adaptive filtering mechanisms [72]. Recently, a few authors have modeled the interaction between decision-theoretic systems and jammers in a game-theoretic framework [2, 4, 55].

In the context of detection networks, there have been a few papers in the past which addressed the problem of jamming attacks. In particular, in the context of ad-hoc wireless sensor networks, Wood and Stankovic have discussed several denial-of-service (DoS) attacks in [74]. In [31], Li *et al.* presented optimal ad-hoc sensor network and jammer designs under perfect knowledge of the channel-state information of all the channels. In the context of spectrum sensing in cognitive radio networks, Li *et al.* [29] proposed a channel-hopping design for multi-band spectrum sensing in cognitive radio networks under a game-theoretic framework, where the radio tries to move from one channel to another in order to evade interferers. In this chapter, we consider the problem of finding the optimal jamming attack, which maximizes the error probability in a simple distributed detection network where there is one sensor and one FC.

5.2 System Model

Consider a simple detection network in a one-dimensional field as shown in Figure 5.1, where all the entities in our model lie on a straight line. We assume that our network model consists of a single sensing agent located at x_s and the FC located at $x_{fc} = 0$. Let the PoI be located at x_t. Let the two hypotheses corresponding to the absence and presence of PoI be denoted as H_0 and H_1 respectively, each with a prior probability $P(H_0) = p_0$ and $P(H_1) = p_1 = 1 - p_0$.

Consider a random jammer located at x_J, trying to maximally degrade the performance

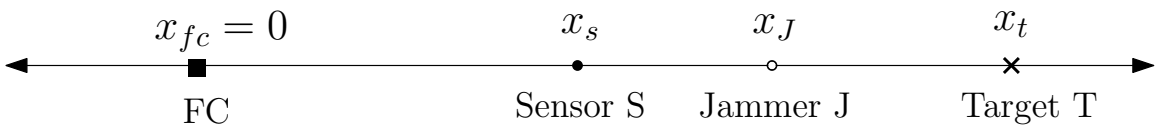


Figure 5.1: Detection Network Model

of the network under a power constraint P_J. Therefore, the jammer distributes the available power between the two channels: the sensing channel and the communication channel. Let the two jamming signals be denoted as w_s and w_{fc} corresponding to the sensing and the communication channels respectively. We assume that these two jamming signals follow $w_s \sim \mathcal{N}(0, \sigma^2_{W_s})$ and $w_{fc} \sim \mathcal{N}(0, \sigma^2_{W_{fc}})$ respectively. Thus, the power constraint on the jammer can be formally stated as $\sigma^2_{W_s} + \sigma^2_{W_{fc}} \leq P_J$.

In this chapter, we assume that the observation at the i^{th} sensing agent is modeled as follows.

$$r_s = h_s \cdot \theta + g_s \cdot w_s + n_s \tag{5.1}$$

where $h_s = \dfrac{1}{1 + \alpha(x_i - x_t)^n}$ and $g_s = \dfrac{1}{1 + \alpha(x_i - x_J)^n}$ are the path-loss coefficients to the PoI-sensor channel and the jammer-sensor channel respectively, both with exponent n and attenuation factor α. We assume that the PoI's state is modeled using a binary variable θ, which takes the value 0 under hypothesis H_0, and 1 otherwise. Furthermore, we assume that $n_i \sim \mathcal{N}(0, \sigma^2_s)$.

The sensor processes its observation r_s into a binary antipodal symbol u using the following quantizer rule.

$$r_s \underset{u=-1}{\overset{u=+1}{\gtrless}} \lambda_s. \tag{5.2}$$

Let P_F and P_D be the probabilities of false alarm and detection respectively, at the sensor. Assuming that the path-loss coefficients for all the channels are known, we have

$$P_F = P(r_s \geq \lambda_s | H_0) = Q\left(\frac{\lambda_s}{\sqrt{\sigma^2_s + g^2_s \sigma^2_{W_s}}}\right) \tag{5.3a}$$

$$P_D = P(r_s \geq \lambda_s | H_1) = Q\left(\frac{\lambda_s - h_s}{\sqrt{\sigma_s^2 + g_s^2 \sigma_{W_s}^2}}\right) \tag{5.3b}$$

When the sensor transmits this binary symbol u, the FC receives a signal

$$r_{fc} = h_{fc} u + g_{fc} w_{fc} + n_{fc}, \tag{5.4}$$

where h_{fc} and g_{fc} are path-loss coefficients for the communication channel and the jammer-to-FC channel respectively. The FC makes a global inference regarding the PoI using the following decision rule:

$$r_{fc} \underset{H_0}{\overset{H_1}{\gtrless}} \lambda_{fc}. \tag{5.5}$$

We consider error probability at the FC as the performance metric, which is defined as

$$P_E = p_0 Q_F + p_1 (1 - Q_D) \tag{5.6}$$

where

$$Q_F = P(r_{fc} \geq \lambda_{fc} | H_0) = P_F Q\left(\frac{\lambda_{fc} - h_{fc}}{\sqrt{\sigma_{fc}^2 + g_{fc}^2 \sigma_{W_{fc}}^2}}\right) + (1 - P_F) Q\left(\frac{\lambda_{fc} + h_{fc}}{\sqrt{\sigma_{fc}^2 + g_{fc}^2 \sigma_{W_{fc}}^2}}\right), \tag{5.7a}$$

$$Q_D = P(r_{fc} \geq \lambda_{fc} | H_1) = P_D Q\left(\frac{\lambda_{fc} - h_{fc}}{\sqrt{\sigma_{fc}^2 + g_{fc}^2 \sigma_{W_{fc}}^2}}\right) + (1 - P_D) Q\left(\frac{\lambda_{fc} + h_{fc}}{\sqrt{\sigma_{fc}^2 + g_{fc}^2 \sigma_{W_{fc}}^2}}\right) \tag{5.7b}$$

are the false alarm and detection probabilities at the FC respectively.

In this chapter, we assume that the thresholds λ_s and λ_{fc} are the optimal thresholds that minimize the error probability of the detection network in the absence of the jammer. In the following subsection, we compute these optimal thresholds at both the sensor and the FC.

5.2.1 Network Design in the Absence of Jammer

In this subsection, we find the optimal thresholds λ_s and λ_{fc}, such that the error probability P_E is minimized in the absence of a jammer. Since there is no jamming signal in our model, we have $w_s = w_{fc} = 0$. Consequently, the average power of the jamming signals reduces to $\sigma^2_{W_s} = \sigma^2_{W_{fc}} = 0$. Therefore, the design of a detection network reduces to the following problem.

Problem 5.1. *Determine the thresholds λ_s and λ_{fc} such that*

$$\underset{\lambda_s, \lambda_{fc}}{\arg\min} \quad P_E$$

where P_E can be found by substituting $\sigma^2_{W_s} = \sigma^2_{W_{fc}} = 0$ in Equation (5.6). Assuming that all the nodes' locations are known, all the channel-coefficients are treated as constants in the above problem. In such a case, the above problem of designing the detection rules jointly at the sensors and the FC is NP-Hard in general [63]. Therefore, we provide closed-form expressions for the thresholds λ_s and λ_{fc} by addressing the above problem in a person-by-person-optimization (PBPO) framework.

Theorem 5.1. *The optimal thresholds λ_s and λ_{fc} that minimize P_E in a PBPO manner, in the absence of the jammer are given by*

$$\lambda_s = \frac{h_s}{2} - \frac{\sigma_s^2}{h_s} \ln \frac{p_1}{p_0} \tag{5.8a}$$

$$\lambda_{fc} = \frac{\sigma_{fc}^2}{2h_{fc}} \ln \left(1 + \frac{p_1 - p_0}{p_0 Q\left(\frac{\lambda_s}{\sigma_s}\right) - p_1 Q\left(\frac{\lambda_s - h_s}{\sigma_s}\right)} \right) \tag{5.8b}$$

Proof. First, we consider the optimization at the sensing agents. The optimal detector at the local sensing agents can be calculated from its likelihood ratio test (LRT), which is

defined as follows.

$$\frac{p(r_s|H_1)}{p(r_s|H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \frac{p_0}{p_1} \tag{5.9}$$

Substituting the conditional distributions $p(r_s|H_0)$ and $p(r_s|H_1)$, and applying logarithms on both sides, we have

$$\frac{1}{2\sigma_s^2}\left[r_s^2 - (r_s - h_s)^2\right] \underset{H_0}{\overset{H_1}{\gtrless}} \ln\frac{p_0}{p_1} \tag{5.10}$$

On further simplification, we find that the optimal local detection rule is given by

$$r_s \underset{u=0}{\overset{u=1}{\gtrless}} \left(\frac{h_s}{2} - \frac{\sigma_s^2}{h_s}\ln\frac{p_0}{p_1}\right) \triangleq \lambda_s. \tag{5.11}$$

Similarly, the optimal decision rule at the FC is an LRT, which is given as follows.

$$\frac{p(r_{fc}|H_1)}{p(r_{fc}|H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \frac{p_0}{p_1}. \tag{5.12}$$

On simplification, we get

$$r_{fc} \underset{H_0}{\overset{H_1}{\gtrless}} \frac{\sigma_{fc}^2}{2h_{fc}}\left(1 + \frac{p_1 - p_0}{p_0 Q\left(\frac{\lambda_s}{\sigma_s}\right) - p_0 Q\left(\frac{\lambda_s - h_s}{\sigma_s}\right)}\right) \triangleq \lambda_{fc}. \tag{5.13}$$

□

Given that the network employs the thresholds presented in Theorem 5.1, we investigate optimal strategies at the jammer in terms of its location and power distribution across sensing and communication channels under the assumption that the jammer has complete information regarding the detection network. Note that the impact due to such a genie-aided jammer serves as an upper-bound on the performance loss at the FC.

5.3 Numerical Study of Optimal Jamming Attack

In this section, we investigate the problem in which the jammer's goal is to optimize its attack within the total available power by inflicting maximum performance deterioration to the detection network. We assume that the jammer has three control parameters - one being its location x_J, and the other two being the energy distributions to the sensing and the communication channels, namely $\sigma^2_{W_s}$ and $\sigma^2_{W_{fc}}$ respectively. The jammer distributes its total available power optimally between the sensing and the communication channels, while simultaneously finding its optimal location so that the error probability at the FC is maximized.

We state this problem formally, as follows.

Problem 5.2. *Given the detection network as stated in Theorem 5.1, determine the optimal power distribution at the jammer as follows.*

$$\begin{aligned} &\underset{\sigma^2_{w_s},\sigma^2_{w_{fc}},x_J}{\textit{maximize}} && P_E \\ &\textit{subject to} && \textit{1.}\ \ \mathbb{E}(w_s^2) + \mathbb{E}(w_{fc}^2) \leq P_J. \end{aligned}$$

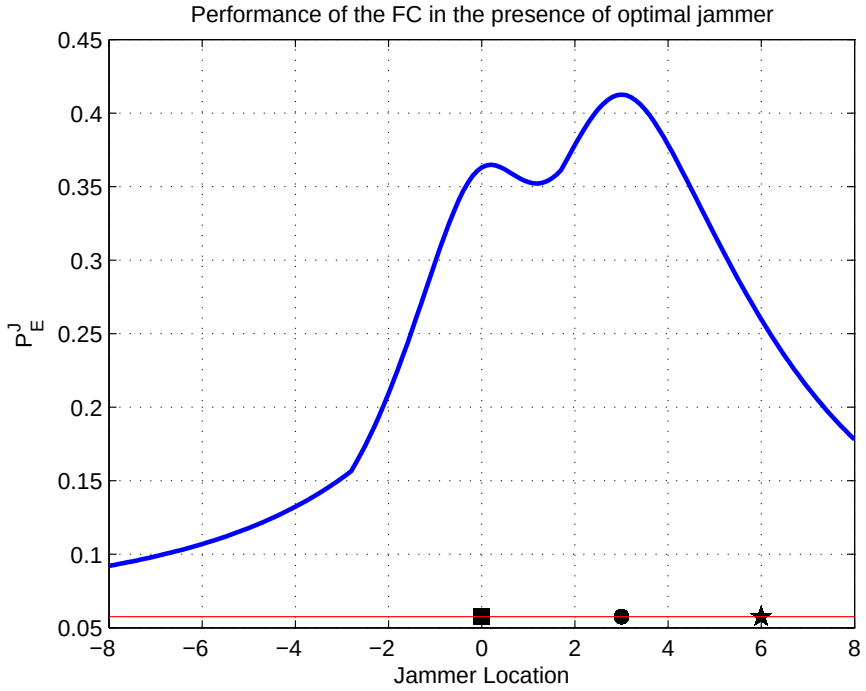
where Condition 1 represents the total power constraint at the jammer. Furthermore, we also assume that the jammer has complete knowledge about the detection network - the decision rules, node locations and also the prior information about the PoI. Therefore, this genie-aided scheme, although not feasible in practice at the jammer, serves as an upper bound on the impact that the jammer can cause on the CR network.

Due to the complicated structure of P_E with respect to the three jammer's parameters x_J, $\sigma^2_{W_s}$ and $\sigma^2_{W_{fc}}$, the problem is analytically intractable. Therefore, we investigate the optimal jammer's strategy numerically in the following subsection.

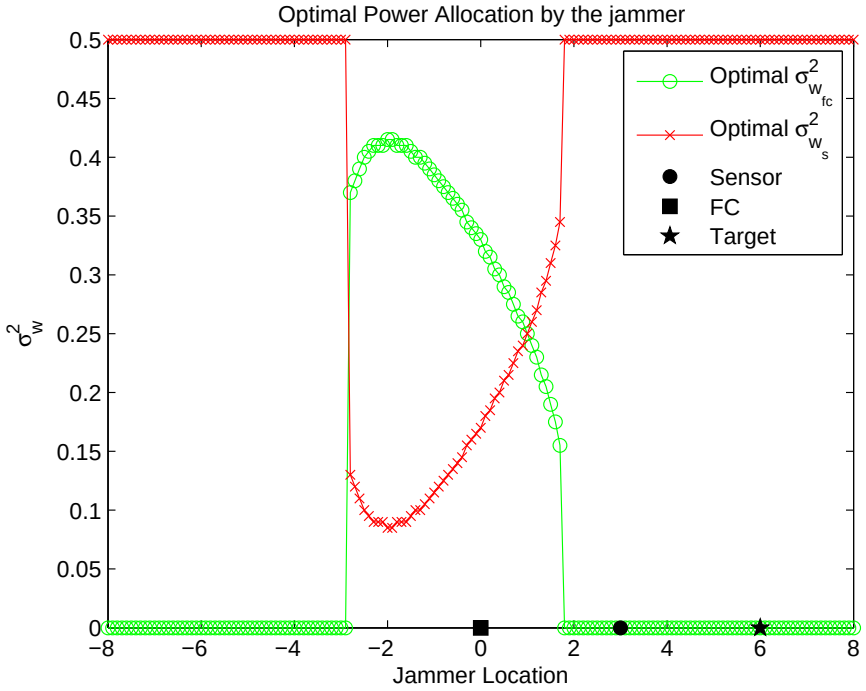
5.3.1 Results and Discussion

In our numerical results, we assume isotropic signal power attenuation models for all the channels' gains where the decaying exponent n varies between 2 and 3 as in free-space RF propagation models. We further assume $\sigma_s = \sigma_{fc} = 0.1$ in our results in Figures 5.2, 5.3, 5.4 and 5.5. In Figures 5.2a, 5.3a, 5.4a and 5.5a, the blue curve indicates the error-probability at the FC as a function of jammer's location, and the red curve represents the the performance of the cognitive radio network in the absence of the jammer. Note that, as the jammer moves away from the network, we observe that the impact of jammer's attack becomes weaker and approaches the red curve even though the jammer employs optimal $\sigma^2_{W_s}$ and $\sigma^2_{W_{fc}}$. On the other hand, Figures 5.2b, 5.3b, 5.4b and 5.5b show how the power is optimally distributed by the jammer between the sensing (red curve) and the communication (green curve) channels.

Note that in Figures 5.2, 5.3 and 5.4, we observe that the optimal jamming attack against the network is to employ the total available energy to either jam the sensor's channel or the FC's channel. The attack is more severe if the jammer is located closer to the CR. As the observations are processed in the CR, noise margin also increases, not allowing the jammer to maximally degrade the transmission channels at the FC. Since the maximum useful information about the PU activity is available at the CR receptions (data-processing inequality from information-theory [16]), the jammer tries to invest more resources on jamming the CR receptions if it is located close to the CR. Note that if the jammer is far away from the FC, then the optimal attack is to direct all its energy to jam the sensor itself. If the jammer is closer to the FC, then the jammer has to distribute its energy between the two channels to bring maximal impact to the CR network.

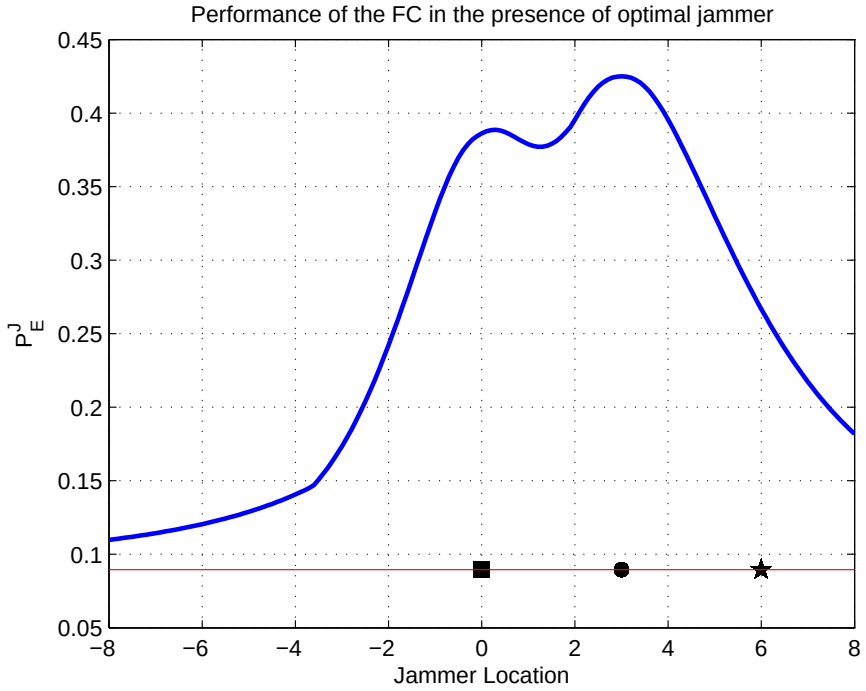


(a)

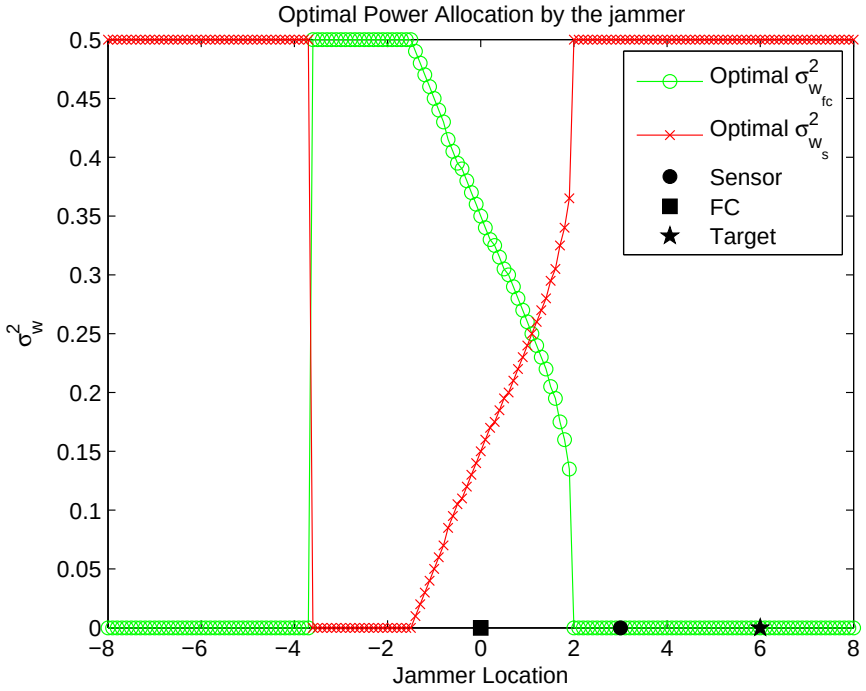


(b)

Figure 5.2: Optimal Jamming attack when $x_s = 3$, $x_t = 6$, $p_0 = 0.5$, $\alpha = 1$, $P_J = 0.5$ and $n = 2$

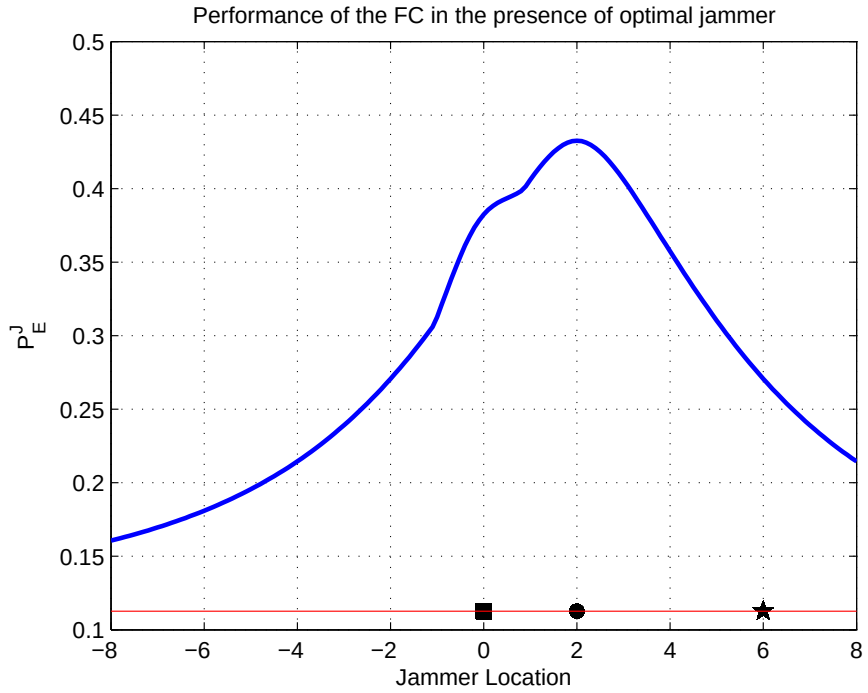


(a)

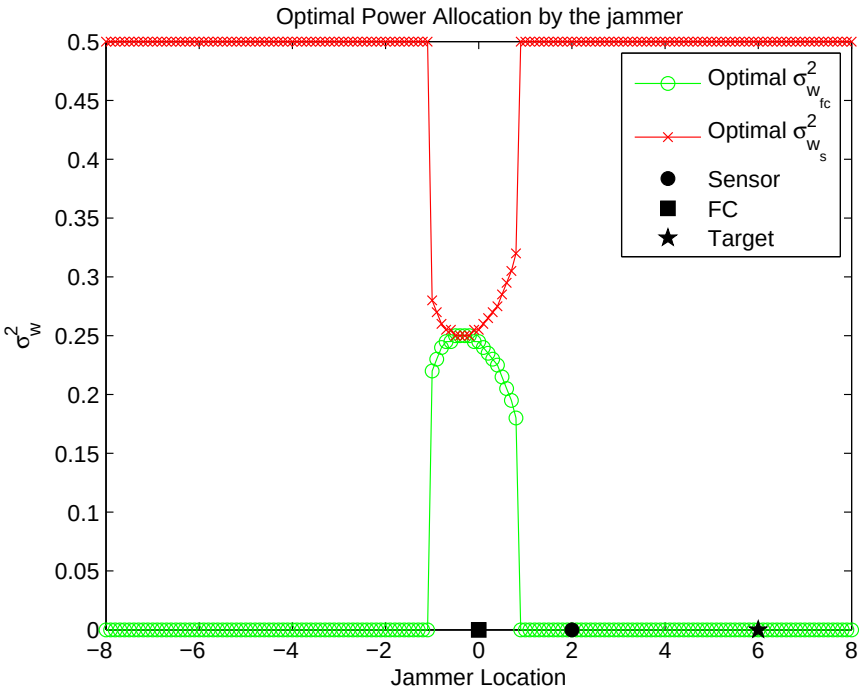


(b)

Figure 5.3: Optimal Jamming attack when $x_s = 3$, $x_t = 6$, $p_0 = 0.5$, $\alpha = 1$, $P_J = 0.5$ and $n = 2.3$

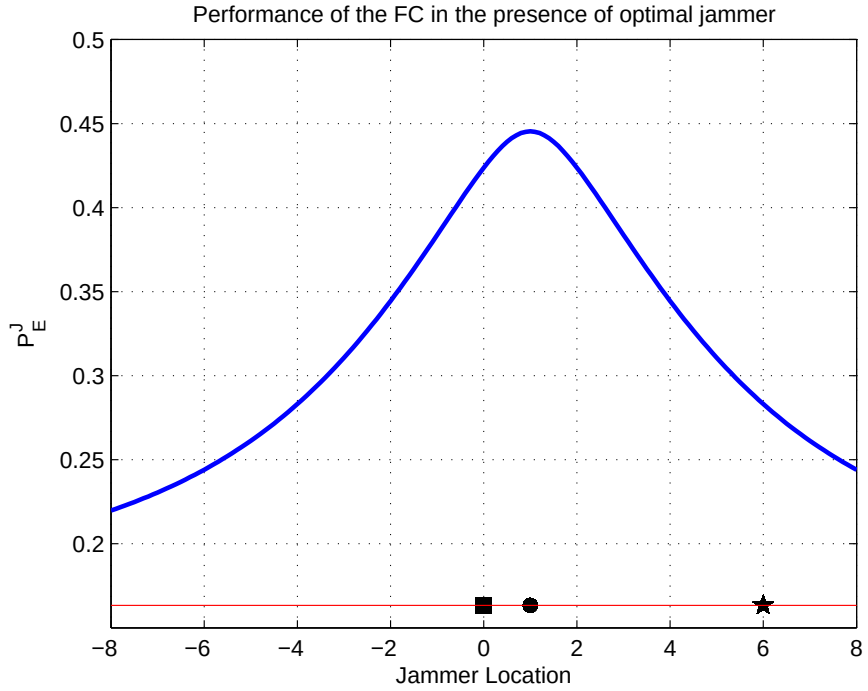


(a)

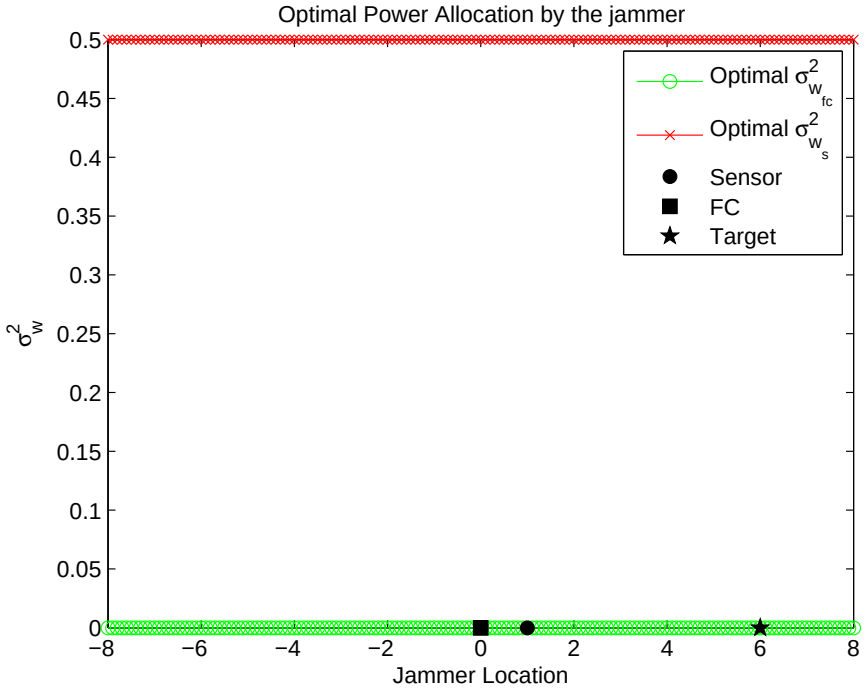


(b)

Figure 5.4: Optimal Jamming attack when $x_s = 2$, $x_t = 6$, $p_0 = 0.5$, $\alpha = 1$, $P_J = 0.5$ and $n = 2$



(a)



(b)

Figure 5.5: Optimal Jamming attack when $x_s = 1$, $x_t = 6$, $p_0 = 0.5$, $\alpha = 1$, $P_J = 0.5$ and $n = 2$

One can also observe from Figures 5.2 and 5.3 that as the decaying exponent increases, the performance of the network degrades. Also, the jammer has a choice to be closer to either the sensor or the fusion center due to the multimodal nature of P_E as a function of x_J. While in the case of Figures 5.4 and 5.5, as compared to Figure 5.2, the CR node is very close to the FC, while the PU transmitter is located far away from the CR network. Hence the jammer has greater impact when it is close to the network. Note that P_E is now a unimodal quasi-convex function of the location of the jammer, x_J since the jammer need not distribute its energy due to the close proximity of FC to the CR node.

5.4 Summary

In this chapter, we proposed a novel attacking scheme for the jammer that can distribute its limited resources over different possible channels utilized for communication in a cognitive radio network. A specific example was considered for the network design where the CR network has only one CR node, to illustrate how a jammer can attack a given network. We presented the optimal distribution of jammer's energy resource between the sensing and the communication channels, and also illustrated how the jammer chooses an optimal location in our simulation results. We have also presented an interesting scenario where the jammer totally focuses its attack on the sensor's reception alone, irrespective of its location.

CHAPTER 6
JAMMING ATTACKS IN CENTRALIZED DETECTION: STRATEGIC GAMES

In contrast to Chapter 5, we model a complete-information zero-sum game between a *centralized* detection network and the jammer in this chapter, where the jammer's strategy is to design its interfering signals rather than energy distribution across the sensing and the communication channels. Furthermore, in this chapter, we consider a more powerful jamming attack than the one in Chapter 5 [42, 43] by assuming multiple antennas at the jammer for transmitting its interfering signals in both sensing and communication channels. We choose the error probability at the FC as the performance metric (utility) in this game, which the network tries to minimize by appropriately choosing the threshold in its fusion rule, while the jammer tries to maximize it by choosing an appropriate jamming signal. We find closed-form expressions for the optimal pure strategies and show that the jammer has no impact on the error probability at the FC due to pure-strategies. We also prove that the network and the jammer converge to one of these pure-strategy equilibria when they play best-response strategies iteratively from any initial point within the space of all possible strategy-profiles. In other words, the jammer has no incentive to employ pure strategies since the network can nullify its impact completely. Therefore, we investigate the impact

of mixed strategies on the network performance and show that the jammer is more effective when it employs mixed strategies.

6.1 Literature Survey

In the past, several efforts have been made to model the interaction between jammers and communication systems in a game-theoretic setting. In particular, a seminal paper [4] by Basar addressed this framework for the first time in 1983, where the author modeled the interaction between a point-to-point communication system and an intelligent jammer as a complete-information zero-sum game in a decision-theoretic framework. While the communication system is designed to minimize the mean-square error at the receiver, the jammer tries to maximize this distortion. Assuming that the jammer is equipped to wire-tap the communication channel, the paper presents optimal (equilibrium) strategies under different channel conditions (low, mid and high signal-to-noise ratio at the receiver). Later, several papers have been published in this topic with the label "correlated jamming" under different scenarios and networked systems [6, 25, 56, 73]. For more details about this line of work, the reader may refer to a well-written survey by Sagduyu *et al.* in [55].

In the context of inference networks, a few efforts have been made to study jamming attacks in a game-theoretic setting. Apart from our work presented in Chapter 5 and [42] within the context of detection networks, Akyol *et al.* in [2] have studied the interaction between a Gaussian sensor network (distributed estimation network with a Gaussian source) and a jammer. While the network is designed to minimize the receiver's mean square error, the jammer employs a strategy to maximize the distortion. In this paper, the authors assume that the jammer can also acquire observations regarding the PoI (Gaussian source), and only jams the multiple access communication channel. They have shown that the optimal network strategy requires sensors to collaborate in order to enable identical realization of a randomized encoding scheme. The authors have also shown that the optimal strate-

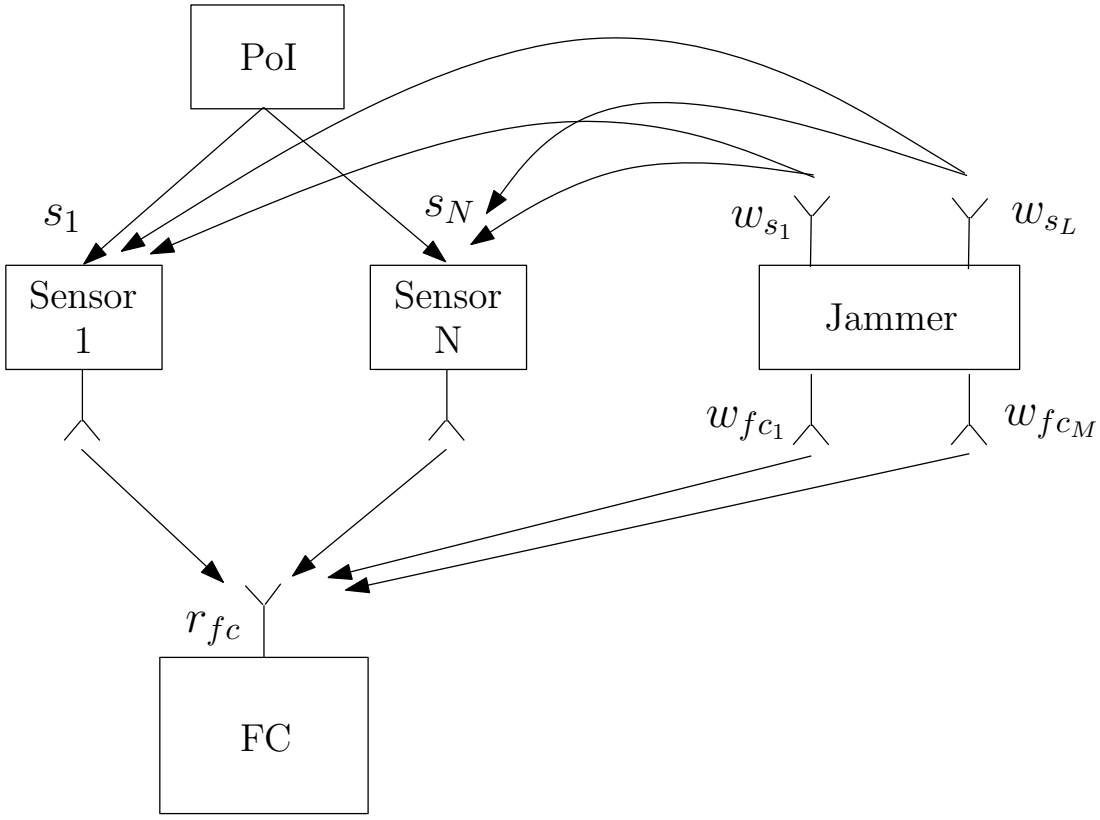


Figure 6.1: Detection Network in the Presence of a Jammer

gies are uncoded in one-shot games, and that a Stackelberg game (sequential interaction between the network and the jammer) does not admit an equilibrium solution.

6.2 System Model

Consider a detection network with N sensing agents and a fusion center (FC) which makes a global decision regarding the presence/absence of the phenomenon-of-interest (PoI) in the presence of a disruptive jammer, as shown in Figure 6.1. Let H_1 denote the hypothesis when PoI is present, and H_0 otherwise, with prior probabilities π_1 and π_0 respectively. We assume that the PoI's signal is modeled as $\theta = 1$ under H_1, and $\theta = 0$ otherwise.

In this chapter, we denote the channel between the PoI and any given sensor as a *sensing*

channel, and the channel between the sensors and the FC as a *communication channel*. We assume that the communication channel at the FC is a multiple access channel (MAC), where all the sensors' messages are superimposed into one received signal at each antenna at the FC. The disruptive jammer interferes with both the sensing and the communication channels by introducing the jamming symbols $\boldsymbol{w}_s$ and $\boldsymbol{w}_{fc}$ respectively. For the sake of notational convenience, we stack these jamming symbols together into a super-symbol $\boldsymbol{w} = \{\boldsymbol{w}_s, \boldsymbol{w}_{fc}\}$.

If α_i and β_{il} denote the known channel-gains at the i^{th} sensing channel due to the PoI signal and the l^{th} antenna at the jammer respectively, the i^{th} sensor acquires an observation s_i as

$$s_i = \alpha_i \theta + \sum_{l=1}^{L} \beta_{il} w_{s_l} + n_i, \tag{6.1}$$

where n_i is a zero-mean AWGN noise with variance σ_s^2.

We assume that the i^{th} sensor transmits its raw observation s_i over the MAC. The FC receives the combined signal

$$\begin{aligned} r_{fc} &= \sum_{i=1}^{N} \phi_i s_i + \sum_{m=1}^{M} \psi_m w_{fc_m} + n_{fc} \\ &= \sum_{i=1}^{N} \phi_i \left(\alpha_i \theta + \sum_{l=1}^{L} \beta_{il} w_{s_l} + n_i \right) + \sum_{m=1}^{M} \psi_m w_{fc_m} + n_{fc} \\ &= a\theta + \boldsymbol{b}^T \boldsymbol{w} + z, \end{aligned} \tag{6.2}$$

where

$$a = \sum_{i=1}^{N} \phi_i \alpha_i, \qquad z = \sum_{i=1}^{N} \phi_i n_i + n_{fc}, \tag{6.3a}$$

$$\boldsymbol{b}^T = \left[\sum_{i=1}^{N} \phi_i \beta_{i1} \quad \cdots \quad \sum_{i=1}^{N} \phi_i \beta_{iL} \quad \psi_1 \quad \cdots \quad \psi_M \right]. \tag{6.3b}$$

We assume that the FC employs a decision rule[1]

$$r_{fc} \underset{H_0}{\overset{H_1}{\gtrless}} \lambda, \tag{6.4}$$

where $\lambda \in \Lambda$[2] is a real-valued threshold designed to minimize the FC's error probability

$$P_E \;\; = \;\; \pi_0 Q_F + \pi_1 (1 - Q_D). \tag{6.5}$$

while the jammer simultaneously attempts to maximize P_E by employing an appropriate jamming signal $\boldsymbol{w}$.

6.3 Jamming Games with Strict Power Constraints

In this section, we assume that the jammer has a strict power constraint, i.e., $||\boldsymbol{w}||_2^2 \leq P$. We denote the set of all possible strategies at the jammer as

$$\mathscr{W} \triangleq \{\boldsymbol{w} \in \mathbb{R}^{L+M} \mid ||\boldsymbol{w}||_2^2 \leq P\}.$$

Since r_{fc} is a superposition of the PoI's signal with several Gaussian random variables, the conditional distributions of the received signal at the FC are given by $r_{fc}|H_0 \sim \mathcal{N}(\boldsymbol{b}^T\boldsymbol{w}, \sigma^2)$ and $\boldsymbol{r}_{fc}|H_1 \sim \mathcal{N}(a + \boldsymbol{b}^T\boldsymbol{w}, \sigma^2)$, where $\sigma^2 = \sigma_{fc}^2 + \sigma_s^2 \sum_{i=1}^{N} \phi_i^2$ is the variance of the noise signal z.

[1] Since this is a likelihood ratio test, all the other rules are dominated. Therefore, their removal does not result any loss in network performance.

[2] Although λ can be any real number in practice, for the sake of tractability, we assume that $\Lambda \triangleq [-R, R]$, where R is a sufficiently large real number. For more details, the reader may refer to Theorem 5, Page 168 in [5] which guarantees the existence of a mixed strategy equilibrium.

Consequently, the error probability P_E at the FC stated in Equation (6.5), is given by

$$\begin{aligned} P_E &= \pi_0 Q_F + \pi_1(1 - Q_D) \\ &= \pi_0 Q\left(\frac{\lambda - \boldsymbol{b}^T\boldsymbol{w}}{\sigma}\right) + \pi_1\left[1 - Q\left(\frac{\lambda - \boldsymbol{b}^T\boldsymbol{w} - a}{\sigma}\right)\right]. \end{aligned} \tag{6.6}$$

Within this framework, we investigate pure-strategy equilibria when the jammer has strict power constraints. We also study the convergence of these pure-strategy equilibria when the detection network and the jammer interact in a repeated game setting. In the latter part of the section, we analyze the effectiveness of mixed strategies at the jammer in comparison to the pure strategy equilibria.

6.3.1 Evaluation of Pure Strategy Equilibria

While the FC employs a strategy λ^* that minimizes P_E, the jammer employs a counter strategy $\boldsymbol{w}^*$ that maximizes P_E. We model this interaction formally as a zero-sum game between the FC and the jammer in the following problem statement.

Problem 6.1. *Find the Nash equilibria* $\{\lambda^*, \boldsymbol{w}^*\} \in \Lambda \times \mathscr{W}$ *that satisfy the following inequality:*

$$P_E(\lambda^*, \boldsymbol{w}) \leq P_E(\lambda^*, \boldsymbol{w}^*) \leq P_E(\lambda, \boldsymbol{w}^*)$$

$$\forall\ \lambda \in \Lambda,\ \boldsymbol{w} \in \mathscr{W}.$$

Before we solve the above problem statement, we investigate some important properties of P_E. These properties of P_E guarantee the existence of pure-strategy Nash equilibria.

Lemma 6.1. *For a given* $\boldsymbol{b}$*,* $\boldsymbol{w}$ *and* σ*,* P_E *is a quasiconvex function of* λ*.*

Proof. In order to prove quasiconvexity of P_E, we adopt an approach, similar to that employed in Lemma 1 in [76]. For a fixed $\boldsymbol{b}$, $\boldsymbol{w}$ and σ, we first differentiate P_E with respect to λ as follows.

$$\begin{aligned} \frac{\partial P_E}{\partial \lambda} &= \pi_0 \frac{\partial Q_F}{\partial \lambda} - \pi_1 \frac{\partial Q_D}{\partial \lambda} \\ &= f_1(\lambda) \cdot [\pi_1 f_2(\lambda) - \pi_0] \end{aligned} \tag{6.7}$$

where

$$f_1(\lambda) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(\lambda - \boldsymbol{b}^T\boldsymbol{w})^2}{2\sigma^2}\right), \tag{6.8a}$$

$$f_2(\lambda) = \exp\left(\frac{2a(\lambda - \boldsymbol{b}^T\boldsymbol{w}) - a^2}{2\sigma^2}\right). \tag{6.8b}$$

Note that $f_1(\lambda) \geq 0$. Therefore, the value of $f_2(\lambda)$ decides the behavior of P_E. One can easily observe that $f_2(\lambda)$ is an exponential function of λ and is, therefore, a monotonically increasing function of λ. Hence, there is only one value of $\lambda = \lambda_0$ at which $f_2(\lambda) = 0$. As a result, we have $\frac{\partial P_E}{\partial \lambda} \geq 0$ whenever $\lambda \geq \lambda_0$, and $\frac{\partial P_E}{\partial \lambda} < 0$, otherwise. In other words, P_E is a quasi-convex function of λ. □

Note that channel models with non-negative channel gains ensure that every element in the vector $\boldsymbol{b}$ is non-negative. Since many practical channel models such as path-loss model and Rayleigh fading model have non-negative channel gains, we assume that $\boldsymbol{b}$ is a non-negative vector in the rest of this section.

Lemma 6.2. *For a given λ, $\boldsymbol{b}$ and σ, P_E is jointly quasiconcave in $\boldsymbol{w}$, if every entry in $\boldsymbol{b}$ is non-negative.*

Proof. Given any two points $\boldsymbol{w}_1, \boldsymbol{w}_2 \in \mathscr{W}$, P_E is jointly quasiconcave [10] if and only if

$$P_E(\boldsymbol{w}_1) \leq P_E(\boldsymbol{w}_2) \quad \Rightarrow \quad \nabla_{\boldsymbol{w}} P_E(\boldsymbol{w}_1) \cdot (\boldsymbol{w}_1 - \boldsymbol{w}_2) \leq 0. \tag{6.9}$$

Therefore, we first consider the necessary condition $P_E(\boldsymbol{w}_1) \leq P_E(\boldsymbol{w}_2)$ and expand it

as follows:

$$\begin{aligned} & P_E(\boldsymbol{w}_1) - P_E(\boldsymbol{w}_2) \ \leq \ 0 \\ \Leftrightarrow \quad & \pi_1 \left[Q\left(\frac{\lambda - \boldsymbol{b}^T \boldsymbol{w}_1 - a}{\sigma} \right) - Q\left(\frac{\lambda - \boldsymbol{b}^T \boldsymbol{w}_2 - a}{\sigma} \right) \right] \\ & -\pi_0 \left[Q\left(\frac{\lambda - \boldsymbol{b}^T \boldsymbol{w}_1}{\sigma} \right) - Q\left(\frac{\lambda - \boldsymbol{b}^T \boldsymbol{w}_2}{\sigma} \right) \right] \ \geq \ 0 \\ \Leftrightarrow \quad & \int_{y_2}^{y_1} g(y) dy \ \geq \ 0. \end{aligned} \tag{6.10}$$

where

$$\begin{aligned} g(y) \ = \ & \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{ -\frac{(y-\lambda)^2}{2\sigma^2} \right\} \\ & \cdot \left[\pi_1 \exp\left(\frac{2a(\lambda - y) - a^2}{2\sigma^2} \right) - \pi_0 \right], \end{aligned} \tag{6.11}$$

and $y_1 = \boldsymbol{b}^T \boldsymbol{w}_1$ and $y_2 = \boldsymbol{b}^T \boldsymbol{w}_2$ are the integral limits.

Given that the values of $\boldsymbol{b}$, λ and σ are fixed, we differentiate P_E with respect to $\boldsymbol{w}$ to have the following.

$$\begin{aligned} \nabla_{\boldsymbol{w}} P_E(\boldsymbol{w}_1) \ &= \ \pi_0 \nabla_{\boldsymbol{w}} Q_F(\boldsymbol{w}_1) - \pi_1 \nabla_{\boldsymbol{w}} Q_D(\boldsymbol{w}_1) \\ &= \ -\boldsymbol{b} \cdot g(y_1). \end{aligned} \tag{6.12}$$

In other words, whenever Equation (6.10) holds true, we need to show that the following condition holds true.

$$\nabla_{\boldsymbol{w}} P_E(\boldsymbol{w}_1) \cdot (\boldsymbol{w}_1 - \boldsymbol{w}_2) \ = \ -g(y_1) \cdot [y_1 - y_2] \ \leq \ 0. \tag{6.13}$$

Equivalently, we need to show that

$$g(y_1) \cdot [y_1 - y_2] \ \geq \ 0. \tag{6.14}$$

Before we prove the above condition, as given in Equation (6.14), we investigate the

behavior of the function $g(y)$. Note that the function $g(y)$ is of the following form.

$$g(y) = f_3(y) \cdot [\pi_1 f_4(y) - \pi_0], \tag{6.15}$$

where

$$f_3(y) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(\lambda - y)^2}{2\sigma^2}\right), \tag{6.16a}$$

$$f_4(y) = \exp\left(\frac{2a(\lambda - y) - a^2}{2\sigma^2}\right). \tag{6.16b}$$

Note that $f_3(y) \geq 0$. Since $f_4(y)$ is a monotonically decreasing function of y, we have $g(y) \geq 0$ whenever $y \leq y_0$, and $g(y) < 0$ whenever $y > y_0$, where y_0 is the unique zero-crossing point at which $f_4(y_0) = \dfrac{\pi_0}{\pi_1}$.

Therefore, we prove the theorem statement case-by-case as shown below.

CASE-1 [$y_0 \leq y_1, y_2$] Given that $y_0 \leq y_1, y_2$, we have $g(y) \leq 0$ for any y between y_1 and y_2. In such a case, the necessary condition given in Equation (6.10) holds true when $y_1 \leq y_2$. In other words, $g(y_1) \cdot [y_1 - y_2] \geq 0$ whenever Equation (6.10) holds true in this case.

CASE-2 [$y_1, y_2 \leq y_0$] Given that $y_1, y_2 \leq y_0$, we have $g(y) \geq 0$ for any y between y_1 and y_2. Therefore, the necessary condition in Equation (6.10) holds true when $y_2 \leq y_1$. As a result, $g(y_1) \cdot [y_1 - y_2] \geq 0$ whenever Equation (6.10) holds true in this case.

CASE-3 [$y_1 \leq y_0 \leq y_2$ or $y_2 \leq y_0 \leq y_1$] Note that this is a trivial case. This is because of the following. If $y_1 \leq y_0 \leq y_2$, both $g(y_1)$ and $(y_1 - y_2)$ are negative. On the other hand, if $y_2 \leq y_0 \leq y_1$, both $g(y_1)$ and $(y_1 - y_2)$ are positive. Either way, their product $g(y_1) \cdot [y_1 - y_2] \geq 0$ whether or not, the necessary condition in Equation (6.10) holds true. □

Given that P_E is quasi-concave-convex in nature, a pure strategy solution exists due

to the classic Debreu-Glicksberg-Fan existence theorem [5, 17]. Therefore, we start by investigating the necessary conditions that a pure-strategy equilibrium would satiate.

Proposition 6.1. *The optimal threshold* $\lambda^* = \arg\min_{\lambda} P_E(\lambda, \boldsymbol{w})$ *for a fixed jammer's strategy* $\boldsymbol{w}$ *is given by*

$$\lambda^* = \boldsymbol{b}^T \boldsymbol{w} + c \tag{6.17}$$

where $c = \dfrac{1}{2a}\left[a^2 + 2\sigma^2 \log\left(\dfrac{\pi_0}{\pi_1}\right)\right]$ *is a constant. Furthermore,* $P_E(\lambda = \lambda^*, \boldsymbol{w})$ *is independent of* $\boldsymbol{w}$.

Proof. We first consider the inner optimization in the max-min problem where we minimize P_E with respect to λ for a fixed jammer's strategy $\boldsymbol{w}$. The optimal $\lambda = \lambda^*$ satisfies

$$\frac{\partial P_E}{\partial \lambda} \;=\; f_1(\lambda) \cdot [\pi_1 f_2(\lambda) - \pi_0] = 0, \tag{6.18}$$

where $f_1(\lambda) \geq 0$. Thus, if $f_2(\lambda) = \dfrac{\pi_0}{\pi_1}$, we have $\dfrac{\partial P_E}{\partial \lambda} = 0$. Substituting Equation (6.8b) and rearranging terms, we have

$$\lambda^* = \boldsymbol{b}^T \boldsymbol{w} + c \tag{6.19}$$

where $c = \dfrac{1}{2a}\left[a^2 + 2\sigma^2 \log\left(\dfrac{\pi_0}{\pi_1}\right)\right]$ is independent of $\boldsymbol{w}$, and $\boldsymbol{b}$ is given in Equation (6.3b).

Given a fixed jammer's strategy $\boldsymbol{w}$, if the FC employs the optimal threshold λ^*, from Equation (6.19), the error probability at the FC is given by

$$P_E(\lambda^*, \boldsymbol{w}) \;=\; \pi_0 Q\left(\frac{c}{\sigma}\right) + \pi_1 \left[1 - Q\left(\frac{c-a}{\sigma}\right)\right]. \tag{6.20}$$

Note that $P_E(\lambda^*, \boldsymbol{w})$ is independent of the jammer's strategy $\boldsymbol{w}$, as stated in the proposition statement. □

Note that the best response strategy employed by the network, as shown in Equation (6.19), is unique for a fixed jammer's strategy $\boldsymbol{w}$. Furthermore, the jammer's signal introduces a linear shift to the point $\lambda = c$, which is optimal in the absence of the jammer.

On the other hand, when we investigate the optimal jammer's strategy $\boldsymbol{w}^*$ by considering the min-max framework, we have the following proposition.

Proposition 6.2. *The optimal jammer's strategy* $\boldsymbol{w}^* = \arg\max_{\boldsymbol{w}} P_E(\lambda, \boldsymbol{w})$ *for a fixed threshold* λ *satisfies*

$$\boldsymbol{b}^T \boldsymbol{w}^* = \lambda - c. \tag{6.21}$$

where $c = \dfrac{1}{2a}\left[a^2 + 2\sigma^2 \log\left(\dfrac{\pi_0}{\pi_1}\right)\right]$. *Such a pure-strategy solution exists only when*

$$c - \sqrt{P \cdot \boldsymbol{b}^T \boldsymbol{b}} \;\leq\; \lambda \;\leq\; c + \sqrt{P \cdot \boldsymbol{b}^T \boldsymbol{b}}. \tag{6.22}$$

Proof. A similar approach to the proof of Proposition 6.1 can be followed in finding Equation (6.21). Therefore, we focus our attention in finding the existence condition, given in Equation (6.22).

In order for a pure-strategy solution to exist, $\boldsymbol{w}^*$ should lie within the set of strategies that satisfy the jammer's total power budget. In other words, we need $(\boldsymbol{w}^*)^T \boldsymbol{w}^* \leq P$. Therefore, the affine function given in Equation (6.21) should be within the squared-distance of P units from the origin $\boldsymbol{w} = \boldsymbol{0}$. In other words, we have

$$\frac{(\lambda - c)^2}{\boldsymbol{b}^T \boldsymbol{b}} \leq P. \tag{6.23}$$

Note that this condition can also be equivalently stated as given in Equation (6.22). □

Note that the jammer's best response strategy is not unique, as shown in Equation (6.21). Indeed, there are infinite possibilities since the jammer can adopt any strategy on a line segment without any regret.

Combining the results from Propositions 6.1 and 6.2, we have the following main result of this section.

Theorem 6.1. *For every $-\boldsymbol{b} \leq \boldsymbol{\epsilon} \leq \boldsymbol{b}$,*

$$\lambda^* = c + \sqrt{\frac{P}{\boldsymbol{b}^T\boldsymbol{b}}}\boldsymbol{b}^T\boldsymbol{\epsilon}, \qquad \boldsymbol{w}^* = \sqrt{\frac{P}{\boldsymbol{b}^T\boldsymbol{b}}}\boldsymbol{\epsilon} \tag{6.24}$$

is a pure-strategy Nash equilibrium. At the above equilibrium point, the error probability at the FC is given by

$$P_E(\lambda^*, \boldsymbol{w}^*) \;\; = \;\; \pi_0 Q\left(\frac{c}{\sigma}\right) + \pi_1\left[1 - Q\left(\frac{c-a}{\sigma}\right)\right]. \tag{6.25}$$

Proof. As stated in Proposition 6.2, λ^* varies between $c - \sqrt{P \cdot \boldsymbol{b}^T\boldsymbol{b}}$ and $c + \sqrt{P \cdot \boldsymbol{b}^T\boldsymbol{b}}$. Therefore, we first investigate the extreme points $\lambda_1^* = c - \sqrt{P \cdot \boldsymbol{b}^T\boldsymbol{b}}$ and $\lambda_2^* = c + \sqrt{P \cdot \boldsymbol{b}^T\boldsymbol{b}}$.

We first consider the case where $\lambda_1^* = c - \sqrt{P \cdot \boldsymbol{b}^T\boldsymbol{b}}$. Comparing this threshold to the optimal threshold from Equation (6.19), we have $\lambda_1^* = \boldsymbol{b}^T\boldsymbol{w} + c = c - \sqrt{P \cdot \boldsymbol{b}^T\boldsymbol{b}}$. On simplification, we find that $\boldsymbol{w}_1^* = -\sqrt{\frac{P}{\boldsymbol{b}^T\boldsymbol{b}}}\boldsymbol{b}$ is the optimal jammer's strategy. Thus, $\lambda_1^* = \boldsymbol{b}^T\boldsymbol{w} + c = c - \sqrt{P \cdot \boldsymbol{b}^T\boldsymbol{b}}$ and $\boldsymbol{w}_1^* = -\sqrt{\frac{P}{\boldsymbol{b}^T\boldsymbol{b}}}\boldsymbol{b}$ form a pure-strategy equilibrium. Similarly, it is easy to show that $\lambda_2^* = c + \sqrt{P \cdot \boldsymbol{b}^T\boldsymbol{b}}$ and $\boldsymbol{w}_2^* = \sqrt{\frac{P}{\boldsymbol{b}^T\boldsymbol{b}}}\boldsymbol{b}$ is another pure-strategy equilibrium.

Given these two pure-strategy equilibria, we find a parametric representation of all possible pure-strategy Nash equilibria, as given below. Let

$$\boldsymbol{w}_{\boldsymbol{\epsilon}}^* = \sqrt{\frac{P}{\boldsymbol{b}^T\boldsymbol{b}}}\boldsymbol{\epsilon} \tag{6.26}$$

where $\boldsymbol{\epsilon}$ is the vector parameter that ranges from $-\boldsymbol{b}$ and $\boldsymbol{b}$. Note that the two solutions $\boldsymbol{w}_1^*$ and $\boldsymbol{w}_2^*$ both correspond to the parameter values $\boldsymbol{\epsilon}_1 = -\boldsymbol{b}$ and $\boldsymbol{\epsilon} = \boldsymbol{b}$ respectively. Furthermore, such a linear parameterization is valid because of the fact that $\boldsymbol{w}^*$ always lies on the line $\boldsymbol{b}^T\boldsymbol{w}^* = \lambda - c$, as given in Equation (6.21).

Substituting Equation (6.26) in Equation (6.19), we have

$$\lambda_{\boldsymbol{\epsilon}}^{*} = c + \sqrt{\frac{P}{\boldsymbol{b}^T\boldsymbol{b}}}\boldsymbol{b}^T\boldsymbol{\epsilon}. \tag{6.27}$$

Since the equilibrium point satiates the necessary conditions presented in Propositions 6.1 and 6.2, the error probability at the FC is given by Equation (6.20). □

6.3.2 Convergence in Repeated Games

In this section, we first investigate if the pure strategy equilibrium is attainable in practice in a repeated setting. Since the network and the jammer do not communicate to agree and play a pure strategy equilibria, it is necessary to analyze their convergence in a repeated game setting. Therefore, in this section, we first investigate the convergence of the players' strategies in a repeated game setting from any arbitrary strategy profile employed by the network and the jammer. We denote the initial pure strategy profile as $(\lambda_0, \boldsymbol{w}_0)$, where the total power of the initial jammer's strategy $\boldsymbol{w}_0$ is within the jammer's power budget P.

Lemma 6.3. *Given any pure strategy profile $(\lambda_0, \boldsymbol{w}_0)$, the players always converge to one of the equilibria presented in Theorem 6.1 in a perfectly-observable repeated-game irrespective of the order of their play.*

Proof. In proving this lemma, we make an assumption that the players' strategies are *perfectly observable*, i.e., the network makes noiseless observations regarding the jammer's strategy and vice-versa. Under such an assumption, we prove the lemma in two cases. In the first case, we assume that the network takes the lead, followed by the jammer and so on. In the latter case, we assume the opposite where the jammer takes the lead, followed by the network and so on.

CASE-1 [N-J-N-J-$\cdots$] In this case, we assume that the network takes the lead. Therefore, given the initial strategy profile $(\lambda_0, \boldsymbol{w}_0)$, the network chooses its best response from

Proposition 6.1, which is

$$\lambda_1 = \boldsymbol{b}^T \boldsymbol{w}_0 + c. \tag{6.28}$$

Given that $||\boldsymbol{w}_0||_2^2 \leq P$, without any loss of generality, we can represent $\boldsymbol{w}_0$ in the same form as shown in Theorem 6.1. As a result, λ_1 also has the form presented in Theorem 6.1. Thus, the repeated game converges to an equilibrium point $(\lambda_1, \boldsymbol{w}_0)$ within one iteration.

CASE-2 [J-N-J-N-$\cdots$] In this case, we assume that the jammer takes the lead. Therefore, given the initial strategy profile $(\lambda_0, \boldsymbol{w}_0)$, the jammer chooses its best response as stated in Proposition 6.2. In other words, if λ_0 lies between $c - \sqrt{P \cdot \boldsymbol{b}^T \boldsymbol{b}}$ and $c + \sqrt{P \cdot \boldsymbol{b}^T \boldsymbol{b}}$, the jammer chooses its best response $\boldsymbol{w}_{1a}$ such that

$$\boldsymbol{b}^T \boldsymbol{w}_{1a} = \lambda_0 - c. \tag{6.29}$$

Otherwise, the jammer employs a strategy $\boldsymbol{w}_{1b} = \pm \boldsymbol{b}$ where the sign of $\boldsymbol{w}_{1b}$ matches to $\text{sign}(\lambda_0 - c)$. In such a case, the network adopts a best response strategy

$$\lambda_1 = c \pm \sqrt{P \cdot \boldsymbol{b}^T \boldsymbol{b}}. \tag{6.30}$$

In summary, if λ_0 lies between $c - \sqrt{P \cdot \boldsymbol{b}^T \boldsymbol{b}}$ and $c + \sqrt{P \cdot \boldsymbol{b}^T \boldsymbol{b}}$, the repeated game converges to an equilibrium point $(\lambda_0, \boldsymbol{w}_{1a})$ in one iteration. Else, the repeated game converges to an equilibrium point $(\lambda_1, \boldsymbol{w}_{1b})$.

□

6.4 Effectiveness of a Gaussian Jammer with Average Power Constraint

Given that both the network and the jammer converge rationally to the pure strategy equilibrium presented in Theorem 6.1, pure-strategies are practically ineffective at the jammer. This is because the error probability at the FC under such equilibrium solutions is totally independent of the jammer's strategy. In fact, the error probability at the FC in the presence of a jammer is identical to the FC's performance in the absence of a jammer (i.e., $\boldsymbol{w} = \mathbf{0}$). In such a case, there is no incentive for the jammer to launch its attack as the network can easily mitigate its impact with very minimal effort.

Given that pure strategies are not beneficial to the jammer, we now investigate if mixed strategy equilibria can help deteriorate the network performance. For the sake of illustration and tractability, we further relax our problem by assuming that the jammer admits an average power constraint. In other words, if $W = \mathbb{E}(\boldsymbol{w}\boldsymbol{w}^T)$ denotes the covariance matrix of the jamming signal $\boldsymbol{w}$, then we have $\mathrm{Tr}(W) \leq P$. Furthermore, we assume that the jammer employs additive Gaussian noise such that $\boldsymbol{w} \sim \mathscr{N}(\mathbf{0}, W)$. In the following lemma, we demonstrate that a Gaussian jammer with an average power constraint, as stated above, has a greater impact than that of a pure-strategy equilibrium.

Lemma 6.4. *When the network employs its best response (mixed) strategy to the jammer's mixed strategy, the expected utility (average error probability) due to a Gaussian jammer with an average power constraint is always greater than the error probability under pure-strategy equilibrium.*

Proof. Given a fixed threshold λ at the FC, the error probability at the FC turns out to be

$$\tilde{P}_E(\lambda) \;=\; \pi_0 Q\left(\frac{\lambda}{\sqrt{\sigma^2 + \boldsymbol{b}^T W \boldsymbol{b}}}\right) + \pi_1 \left[1 - Q\left(\frac{\lambda - a}{\sqrt{\sigma^2 + \boldsymbol{b}^T W \boldsymbol{b}}}\right)\right]. \tag{6.31}$$

Note that $\tilde{P}_E(\lambda)$ is a quasiconvex[3] function of λ. In other words, if the network employs a mixed strategy, the optimal (best response) distribution is given by $p(\lambda) = \delta(\lambda^*)$, where $\lambda^* = c + \frac{1}{a}\boldsymbol{b}^T W \boldsymbol{b} \log \frac{\pi_0}{\pi_1}$ is the optimal threshold that minimizes $\tilde{P}_E(\lambda)$, and $\delta(x)$ is a Dirac delta function centered at x. Thus, the expected utility (minimum $\tilde{P}_E(\lambda)$) due to a Gaussian jammer is

$$U(W) \;=\; \pi_0 Q\left(\frac{c + \boldsymbol{b}^T W \boldsymbol{b}\frac{1}{a}\log\frac{\pi_0}{\pi_1}}{\sqrt{\sigma^2 + \boldsymbol{b}^T W \boldsymbol{b}}}\right) + \pi_1\left[1 - Q\left(\frac{c - a + \boldsymbol{b}^T W \boldsymbol{b}\frac{1}{a}\log\frac{\pi_0}{\pi_1}}{\sqrt{\sigma^2 + \boldsymbol{b}^T W \boldsymbol{b}}}\right)\right]. \tag{6.32}$$

Note that $U(W)$ is a quasiconvex[4] function of W, with its minimum at W being an all-zero matrix. In other words,

$$U(W) \geq P_E(\lambda^*, w^*), \tag{6.33}$$

where $P_E(\lambda^*, w^*)$ is given in Equation 6.25. Consequently, the jammer has every incentive to use a mixed strategy rather than employing a deterministic (pure) strategy. □

6.4.1 Illustrative Example

For the sake of illustration, we study the properties of saddle point equilibria in the following cognitive radio (CR) network example, where a Gaussian jammer (interferer) is equipped with one antenna each for the sensing and the communication channels, to inject a random Gaussian signal in each of these channels. We assume that there are $N = 20$ CRs in the network, whose locations are defined using a Binomial point process [71] over the 10×10 grid centered about the origin. Furthermore, we assume that the FC, primary user (PU) and the jammer are located at $\boldsymbol{x}_{fc} = (0, 0)$, $\mathbf{x_t} = (-3, -4)$ and $\mathbf{x_j} = (1, -2)$ respectively. We assume $\sigma_s = 0.1$, $\sigma_{fc} = 0.1$ and $P_J = 0.5$. Also, we assume free-space

[3]Proof is similar to our approach in Lemma 6.1.

[4]The proof is similar to our approach in Lemma 6.2.

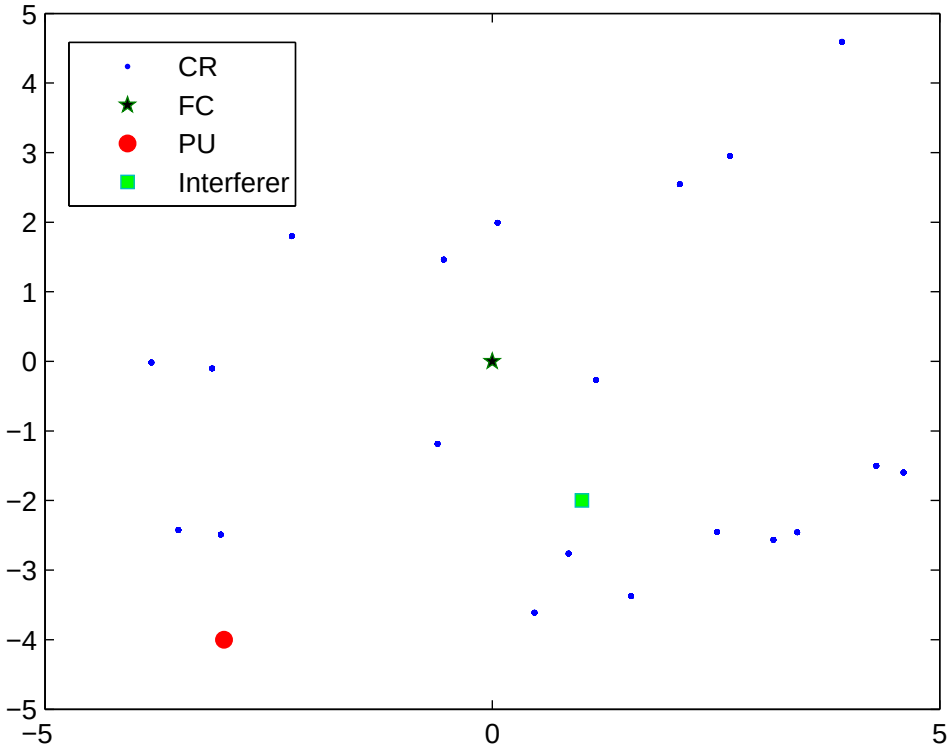
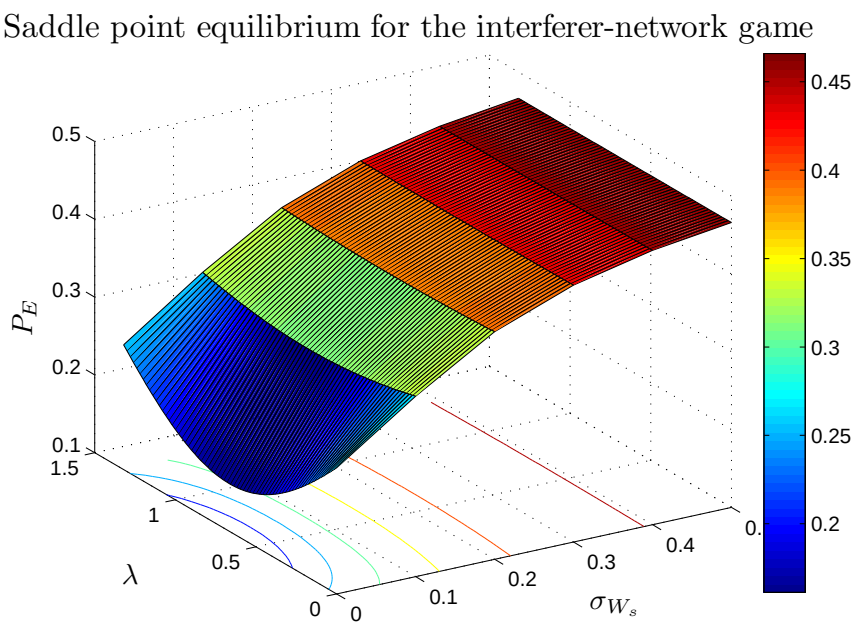
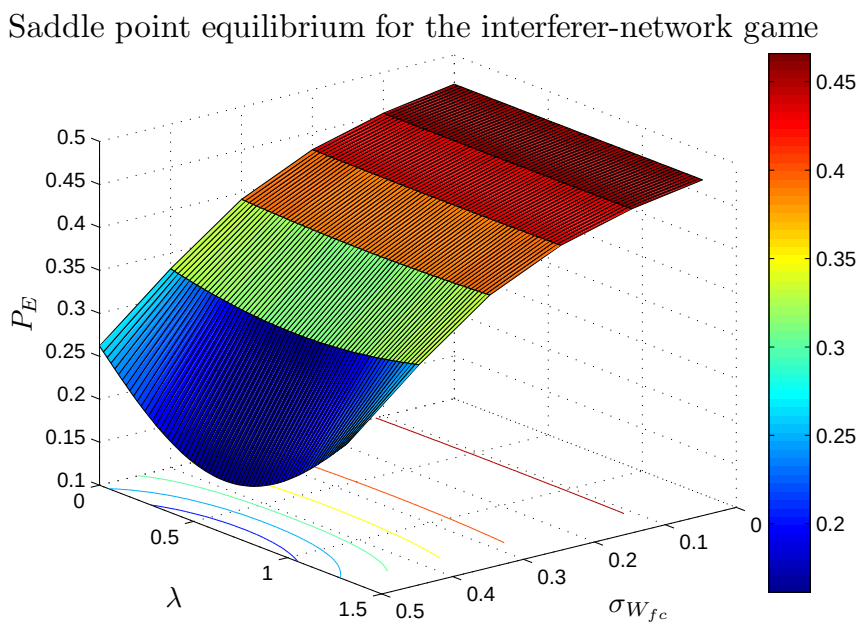


Figure 6.2: CR Network for $\pi_0 = 0.5$ case

path loss shadowing with each path loss coefficient taking the form $\sqrt{\frac{1}{(1+d_{\{.\}}^2)}}$, where $d_{\{.\}}$ is the propagation distance between the transmitter node and the receiver node in the above mentioned system-model.



(a) Error Probability in terms of λ and $\sigma^2_{W_s}$



(b) Error Probability in terms of λ and $\sigma^2_{W_{fc}}$

Figure 6.3: Performance of the CR network for $\pi_0 = 0.5$

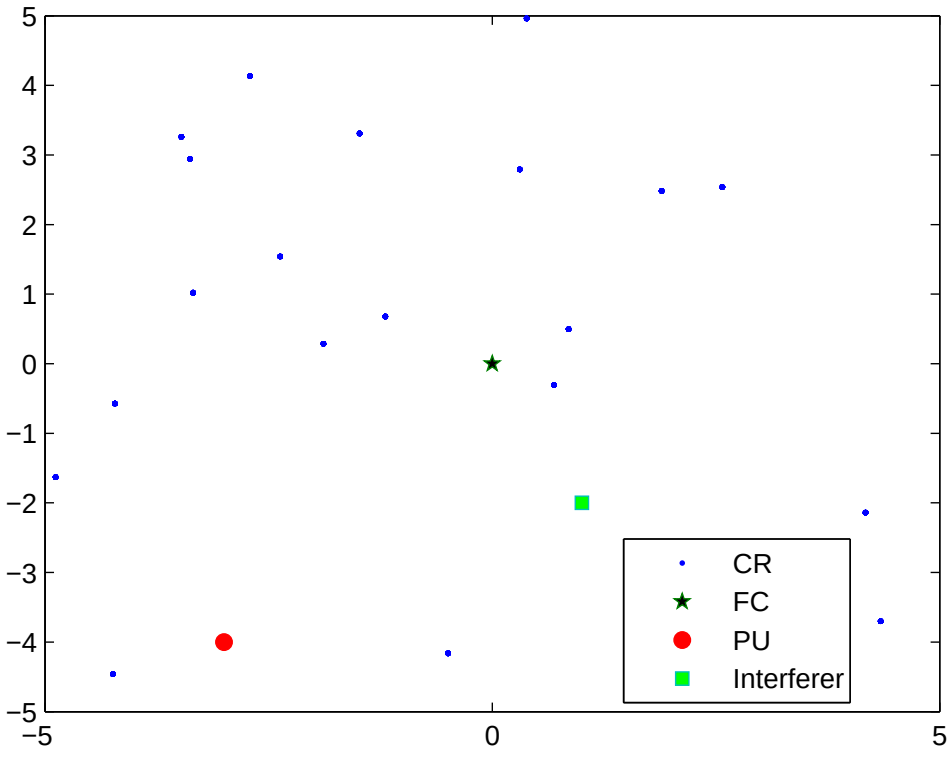
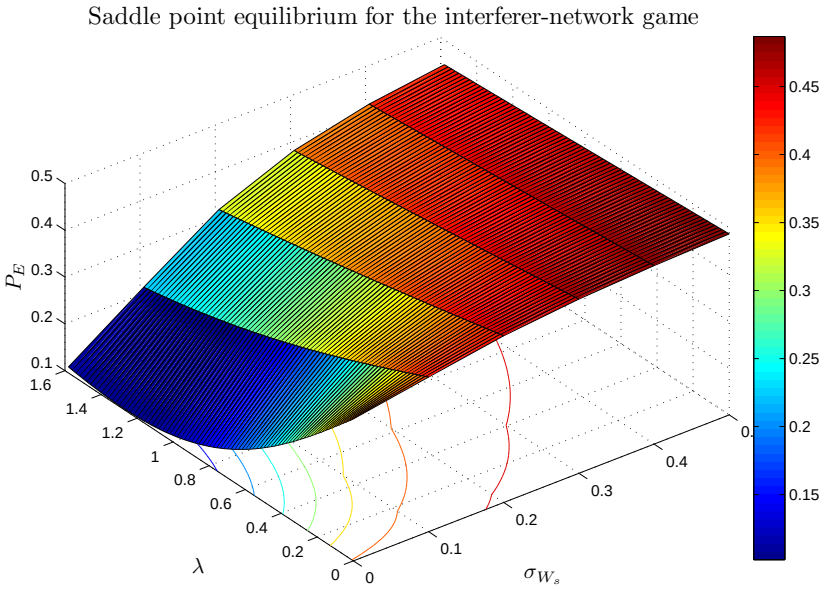


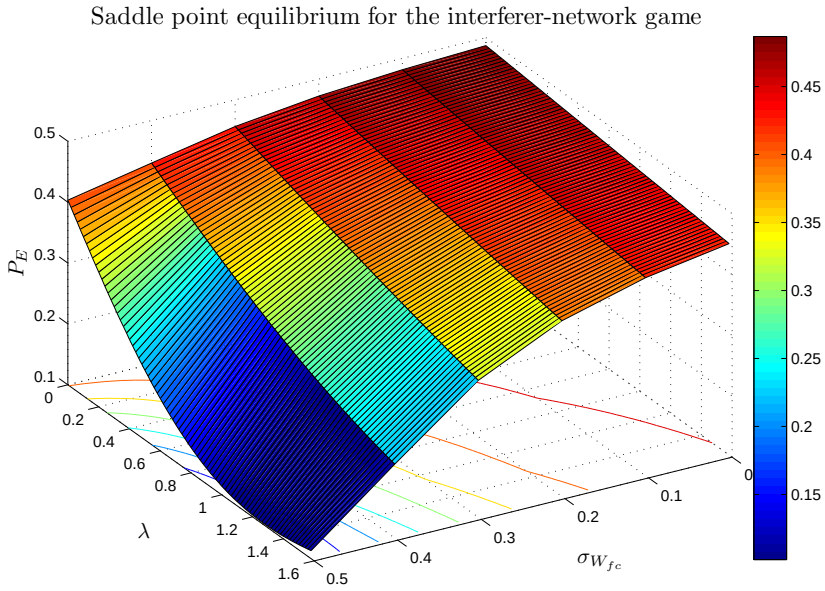
Figure 6.4: CR Network for $\pi_0 = 0.8$ case

First, in Figure 6.3, we present error probability as a function of λ, $\sigma_{W_{fc}}$ and σ_{W_s} for the CR network shown in Figure 6.2. In this case, we let $\pi_0 = 0.5$, which is the worst case performance scenario of the network. The plots depict clearly both quasiconvexity with respect to λ and monotonicity with respect to the interfering node's parameters, $\sigma_{W_{fc}}$ and σ_{W_s}, especially when $\lambda \geq 0$. In this case, as per our intuition, we numerically find $\lambda_{opt} = \frac{a}{2}$.

Figure 6.5, on the other hand, presents the error probability as a function of λ, $\sigma_{W_{fc}}$ and σ_{W_s} for $\pi_0 = 0.8$. We particularly present these results because $\pi_0 = 0.8$ is found in practice as pointed out by the FCC's survey on spectrum utilization of licensed bands [77]. One can clearly note that the NE of the game from the network's perspective has now moved away from $\lambda = \frac{a}{2}$, due to the bias in the prior probabilities.



(a) Error Probability in terms of λ and $\sigma^2_{W_s}$



(b) Error Probability in terms of λ and $\sigma^2_{W_{fc}}$

Figure 6.5: Performance of the CR network for $\pi_0 = 0.8$

One can also note from Figures 6.3 and 6.5 that in both the scenarios considered, the optimal jammer degrades the individual CRs' performance by allocating all the available power to the sensing channel ($\sigma^2_{W_s} = 0.5$ and $\sigma^2_{W_{fc}} = 0$). This is similar to our numerical results in Chapter 5, where we had studied the optimal jamming attack on a simple detection network. Such a strategy can also be justified as per our intuition, since the jammer will always invest all of its resources to interfere with the most vulnerable channel available (one with maximum information about the spectrum availability) in order to have the maximal impact on network performance. Given that the sensing channel carries the maximum amount of information regarding the true state of the PoI, the jammer employs all its power in the sensing channel to cause maximal impact on the network performance.

Note that all our results point to another important observation regarding the effectiveness of mixed strategies at the jammer. A Gaussian jammer with enough resources ($P_J = 0.5$ in this example) can bring the error probability P_E close to 0.5, which is the worse case performance at the FC. In other words, although the jammer has no incentive to employ pure strategies, it can simply inject Gaussian signals into the sensing channel to launch a very powerful denial-of-service attack on the detection network.

6.5 Summary

We have modeled the interaction between a centralized detection network and a jammer as a zero-sum game. We have obtained a family of pure strategy Nash equilibria in closed-form, and proved that the pure-strategy jamming attacks have no impact on the error probability at the FC. We have also shown that both the players will converge to one of the equilibrium points proposed, in a perfectly-observable repeated game irrespective of the order of their play. We also showed that the jammer has an incentive to employ a mixed strategy since the expected utility (average error probability due to mixed strategies) is always greater than that in the case of pure-strategy equilibrium.

CHAPTER 7
CONCLUDING REMARKS

7.1 Summary

With a broad range of applications, security threats in inference networks have a significant impact on several practical domains. This dissertation focused on the design and analysis of secure inference networks under three attack scenarios: (a) *eavesdropping* threats in detection networks in Chapters 2 and 3, (b) *Byzantine* attacks in distributed inference networks in Chapter 4, and, (c) *jamming* attacks in detection networks in Chapters 5 and 6. Following is a brief summary of this dissertation.

In Chapter 2, we have considered the design of binary quantizers for secure distributed detection networks in the presence of an eavesdropper. If the goal is to maximize the difference between the KL Divergences at the FC and Eve, we have shown that the optimal binary quantizers are the same as when maximizing the KL Divergence at the FC alone. In contrast, in the case of identical sensors and channels, we have proved that the optimal binary quantizers at the sensors are likelihood-ratio test-based, and have presented a numerical algorithm to find the optimal threshold. In the case of non-identical sensors and channels, we have presented a greedy algorithm to find efficient, near-optimal binary quantizers at the sensors. On the other hand, in Chapter 3, we have proposed an efficient

transmit-diversity mechanism at the sensing agents (which are equipped with multiple antennas) in the context of centralized detection networks. While the sensing agents amplify and forward their raw observations to the FC, they also inject artificial noise in order to confuse Eve. Since the problem of finding the optimal transmit-diversity mechanism is non-convex, we presented a near-optimal solution using semidefinite relaxation.

In Chapter 4, we have investigated optimal Byzantine attacks in the presence of both ideal and non-ideal (discrete, memoryless) communication channels, when the sensing agents quantize their observations into an M-ary symbol. We have also studied optimal resource-constrained Byzantine attacks when the attacker cannot compromise the blinding fraction of nodes in the network. Furthermore, we have also proposed a novel deviation-based reputation mechanism to identify Byzantine nodes in the network.

In Chapter 5, we have investigated an optimal jamming attack in a distributed detection network, where the goal of the jammer is to maximize error probability at the FC by optimizing its placement and power allocation between the sensing and the communication channels. Since the problem is non-convex, we have considered a simple network where there is only one sensing agent. For the sake of illustration, we have assumed that all the entities (sensing agent, FC, PoI and the jammer) lie on a straight line, and presented numerical results that throw light on the jammer's optimal strategy. In contrast, in Chapter 6, we have modeled the interaction between the jammer and a centralized detection network as a complete-information zero-sum game. We have found closed-form expressions for a family of pure-strategy equilibria when there is a strict power constraint on the jammer. In addition, we have also shown that the jammer has no incentive to employ a pure-strategy, but instead, chooses mixed strategies to alleviate detection performance at the FC. We have also investigated mixed strategy equilibria numerically in the presence of a Gaussian jammer.

7.2 Future Research Directions

Security in any domain/application is evolutionary, and demands novel designs and solutions as the attacker evolves in time. In this dissertation, we have investigated three basic security threats in inference networks, each belonging to a class within the *CIA* framework. Given the evolutionary nature of security attacks, several open problems still remain unsolved in this field, even within the three security threats addressed in this dissertation. We discuss some of these open problems for future work in the remaining section.

- *Eavesdropping Attack*: A direct extension to our current work is to explore other methods (e.g. convex-concave approximation) to solve the non-convex optimal transmit diversity design problem efficiently in order to further improve the detection performance of the inference network. Similar problems still remain open within the context of other inference networks which are designed to address inference problems such as statistical estimation, classification, prediction, tracking and so on. Given that there is a tradeoff between detection performance and security, it is necessary to find a methodology to detect the presence of eavesdroppers. In practice, this is a difficult problem especially when the eavesdropper remains passive and does not emit any electromagnetic radiation.

- *Byzantine Attack*: The problem of designing an optimal inference network still remains open under the non-asymptotic regime (finite number of sensors in the inference network) when the sensors transmit M-ary quantized data to the FC. Furthermore, in the case of resource-constrained Byzantine attacks, the problem of finding the optimal Byzantine attack in the space of all row-stochastic flipping probability matrices still remains open. Moreover, the problem of designing a secure inference network, along with mitigation techniques in the presence of heterogeneous sensing agents still remains open.

- *Jamming Attack*: Given that the problem of designing an optimal jamming attack

in distributed detection networks is a non-convex problem, one can explore efficient approximations to find near-optimal designs at both the jammer and the network. Also, the problem of finding mixed-strategy equilibria in our proposed game-theoretic framework still remains open under strict power constraints at the jammer. One interesting technique worth investigating, is to study the effects of diversity due to the presence of multiple receiving antennas at the FC, on the network performance in the presence of a jammer.

Note that, in Chapters 2, 3, 5 and 6, we have assumed complete channel-state information at both the network and the attacker. This may not be possible in practice, and therefore, security should be addressed in the presence of incomplete information about the channel gains at both the network and the attacker.

In addition to the open problems discussed in the context of three security threats discussed in this dissertation, the study of other security threats in parallel-topology inference networks still remains open. For example, several other attack models such as a Sybil attack [46] have already been proposed in the context of ad-hoc sensor networks. Security threats in other network topologies also remain open. For example, secure tree-topology inference networks in the presence of eavesdropping and jamming attacks remain open. More recently, there have been several efforts in designing optimal network topologies in inference networks where sensing agents collaborate with each other in order to maximize energy efficiency of the network [23, 32, 33]. Security in such inference networks is a very interesting topic, as it provides many venues (for example, denial-of-service attacks via skipping transmissions) for the attacker to bring down the network.

REFERENCES

[1] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "Wireless sensor networks: a survey," *Computer networks*, vol. 38, no. 4, pp. 393–422, 2002.

[2] E. Akyol, K. Rose, and T. Basar, "Gaussian sensor networks with adversarial nodes," in *Proceedings of IEEE International Symposium on Information Theory (ISIT),*, July 2013, pp. 539–543.

[3] T. C. Aysal and K. E. Barner, "Sensor data cryptography in wireless sensor networks," *IEEE Transactions on Information Forensics and Security*, vol. 3, no. 2, pp. 273–289, 2008.

[4] T. Basar, "The gaussian test channel with an intelligent jammer," *IEEE Transactions on Information Theory*, vol. 29, no. 1, pp. 152–157, 1983.

[5] T. Basar and G. J. Olsder, *Dynamic Noncooperative Game Theory*. Academic Press Inc., 1982.

[6] T. U. Basar, "Optimum linear causal coding schemes for gaussian stochastic processes in the presence of correlated jamming," *IEEE Transactions on Information Theory*, vol. 35, no. 1, pp. 199–202, Jan 1989.

[7] K.-J. Bathe and E. L. Wilson, "Solution methods for eigenvalue problems in structural mechanics," *International Journal for Numerical Methods in Engineering*, vol. 6, no. 2, pp. 213–226, 1973. [Online]. Available: http://dx.doi.org/10.1002/nme.1620060207

[8] R. Bellman, *Dynamic Programming*. Dover Publications, 2003.

[9] R. S. Blum, S. A. Kassam, and H. V. Poor, "Distributed detection with multiple sensors ii. advanced topics," *Proceedings of the IEEE*, vol. 85, no. 1, pp. 64–79, 1997.

[10] S. P. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[11] R. R. Brooks, P. Ramanathan, and A. M. Sayeed, "Distributed target classification and tracking in sensor networks," *Proceedings of the IEEE*, vol. 91, no. 8, pp. 1163–1171, 2003.

[12] J. Chamberland and V. Veeravalli, "Wireless sensors in distributed detection applications," *IEEE Signal Processing Magazine*, vol. 24, no. 3, pp. 16–25, 2007.

[13] J.-F. Chamberland and V. V. Veeravalli, "Decentralized detection in sensor networks," *IEEE Transactions on Signal Processing*, vol. 51, no. 2, pp. 407–416, 2003.

[14] B. Chen, L. Tong, and P. K. Varshney, "Channel-aware distributed detection in wireless sensor networks," *IEEE Signal Processing Magazine*, vol. 23, no. 4, pp. 16–26, 2006.

[15] L. Cheng, C. Wu, Y. Zhang, H. Wu, M. Li, and C. Maple, "A survey of localization in wireless sensor network," *International Journal of Distributed Sensor Networks*, vol. 2012, pp. 1–12, 2012.

[16] T. J. A. Cover, T. M., *Elements of Information Theory*. U.S.A: John Wiley and Sons Inc., 2006.

[17] D. Fudenberg and J. Tirole, *Game Theory*. The MIT Press, 1991.

[18] S. Goel and R. Negi, "Guaranteeing secrecy using artificial noise," *IEEE Transactions on Wireless Communications*, vol. 7, no. 6, pp. 2180–2189, June 2008.

[19] Y.-W. P. Hong, P.-C. Lan, and C.-C. J. Kuo, *Signal Processing Approaches to Secure Physical Layer Communications in Multi-Antenna Wireless Systems*, ser. SpringerBriefs in Electrical and Computer Engineering. Springer, 2014.

[20] H. Jeon, S. W. McLaughlin, and J. Ha, "Cooperative secure transmission for distributed detection in wireless sensor networks," in *Proc. IEEE 54th Int Circuits and Systems (MWSCAS) Midwest Symp*, 2011, pp. 1–4.

[21] B. Kailkhura, S. Brahma, and P. K. Varshney, "Optimal byzantine attack on distributed detection in tree based topologies," in *Proc. of International Conference on Computing, Networking and Communications Workshops (ICNC-CPS)*, San Diego, USA, January 2013.

[22] B. Kailkhura, V. S. S. Nadendla, and P. K. Varshney, "Distributed inference in the presence of eavesdroppers: A survey," *IEEE Communications Magazine*, vol. 53, no. 6, pp. 40–46, 2015.

[23] S. Kar and P. K. Varshney, "Linear coherent estimation with spatial collaboration," *IEEE Transactions on Information Theory*, vol. 59, no. 6, pp. 3532–3553, June 2013.

[24] C. Karlof, N. Sastry, and D. Wagner, "Tinysec: a link layer security architecture for wireless sensor networks," in *Proceedings of the 2nd international conference on Embedded networked sensor systems*, ser. SenSys '04. New York, NY, USA: ACM, 2004, pp. 162–175. [Online]. Available: http://doi.acm.org/10.1145/1031495.1031515

[25] A. Kashyap, T. Basar, and R. Srikant, "Correlated jamming on mimo gaussian fading channels," *IEEE Transactions on Information Theory*, vol. 50, no. 9, pp. 2119–2123, Sept 2004.

[26] O. Kosut, L. Jia, R. J. Thomas, and L. Tong, "Malicious data attacks on the smart grid," *IEEE Transactions on Smart Grid*, vol. 2, no. 4, pp. 645–658, Dec 2011.

[27] L. Lamport, R. Shostak, and M. Pease, "The byzantine generals problem," *ACM Trans. Program. Lang. Syst.*, vol. 4, no. 3, pp. 382–401, Jul. 1982. [Online]. Available: http://doi.acm.org/10.1145/357172.357176

[28] E. G. Larsson, "Mimo detection methods: How they work [lecture notes]," *IEEE Signal Processing Magazine*, vol. 26, no. 3, pp. 91–95, May 2009.

[29] H. Li and Z. Han, "Dogfight in spectrum: Jamming and anti-jamming in multichannel cognitive radio systems," in *Proc. IEEE Global Telecommunications Conf. GLOBECOM 2009*, 2009, pp. 1–6.

[30] J. Li and P. Stoica, "Mimo radar with colocated antennas," *IEEE Signal Processing Magazine*, vol. 24, no. 5, pp. 106–114, Sept 2007.

[31] M. Li, I. Koutsopoulos, and R. Poovendran, "Optimal jamming attack strategies and network defense policies in wireless sensor networks," *IEEE Transactions on Mobile Computing*, vol. 9, no. 8, pp. 1119–1133, Aug 2010.

[32] L. Liu, X. Zhang, and H. Ma, "Dynamic node collaboration for mobile target tracking in wireless camera sensor networks," in *INFOCOM 2009, IEEE*, April 2009, pp. 1188–1196.

[33] S. Liu, S. Kar, M. Fardad, and P. K. Varshney, "Sparsity-aware sensor collaboration for linear coherent estimation," *IEEE Transactions on Signal Processing*, vol. 63, no. 10, pp. 2582–2596, May 2015.

[34] Z.-Q. Luo, W.-K. Ma, A.-C. So, Y. Ye, and S. Zhang, "Semidefinite relaxation of quadratic optimization problems," *IEEE Signal Processing Magazine*, vol. 27, no. 3, pp. 20–34, May 2010.

[35] S. Marano, V. Matta, and L. Tong, "Distributed detection in the presence of byzantine attacks," *IEEE Transactions on Signal Processing*, vol. 57, no. 1, pp. 16–29, 2009.

[36] S. Marano, V. Matta, and P. K. Willett, "Distributed detection with censoring sensors under physical layer secrecy," *IEEE Transactions on Signal Processing*, vol. 57, no. 5, pp. 1976–1986, May 2009.

[37] T. Melzer, "Svd and its application to generalized eigenvalue problems," Vienna University of Technology, Tech. Rep., 2004.

[38] L. B. Milstein, "Interference rejection techniques in spread spectrum communications," *Proceedings of the IEEE*, vol. 76, no. 6, pp. 657–671, Jun 1988.

[39] A. Mpitziopoulos, D. Gavalas, C. Konstantopoulos, and G. Pantziou, "A survey on jamming attacks and countermeasures in wsns," *IEEE Communications Surveys & Tutorials*, vol. 11, no. 4, pp. 42–56, 2009.

[40] A. Mukherjee, S. A. A. Fakoorian, J. Huang, and A. L. Swindlehurst, "Principles of physical layer security in multiuser wireless networks: A survey," *IEEE Communications Surveys & Tutorials*, vol. 16, no. 3, pp. 1550–1573, 2014.

[41] V. S. S. Nadendla, "Secure distributed detection in wireless sensor networks via encryption of sensor decisions," Master's thesis, Louisiana State University, 2009.

[42] V. S. S. Nadendla, H. Chen, and P. Varshney, "Minimax games for cooperative spectrum sensing in a centralized cognitive radio network in the presence of interferers," in *Proc. MILCOM-2011*, Baltimore, MD.

[43] V. S. S. Nadendla, H. Chen, and P. K. Varshney, "On jamming models against collaborative spectrum sensing in a simple cognitive radio network," in *Proc. ALLERTON-2010*, Pacific Grove, CA.

[44] G. N. Nayak and S. Samaddar, "Different flavours of man-in-the-middle attack, consequences and feasible solutions," in *Computer Science and Information Technology (ICCSIT), 2010 3rd IEEE International Conference on*, vol. 5, 2010, pp. 491–495.

[45] R. Negi and S. Goel, "Secret communication using artificial noise," in *Proc. VTC-2005-Fall Vehicular Technology Conf. 2005 IEEE 62nd*, vol. 3, 2005, pp. 1906–1910.

[46] J. Newsome, E. Shi, D. Song, and A. Perrig, "The sybil attack in sensor networks: Analysis & defenses," in *Proceedings of the 3rd International Symposium on Information Processing in Sensor Networks*, ser. IPSN '04. New York, NY, USA: ACM, 2004, pp. 259–268. [Online]. Available: http://doi.acm.org/10.1145/984622.984660

[47] N. Patwari, J. Ash, S. Kyperountas, A. Hero, R. Moses, and N. Correal, "Locating the nodes: cooperative localization in wireless sensor networks," *IEEE Signal Processing Magazine*, vol. 22, no. 4, pp. 54–69, 2005.

[48] A. Perrig, J. Stankovic, and D. Wagner, "Security in wireless sensor networks," *Communications of the ACM*, vol. 47, no. 6, pp. 53–57, 2004.

[49] A. Perrig, R. Szewczyk, J. D. Tygar, V. Wen, and D. E. Culler, "Spins: security protocols for sensor networks," *Wirel. Netw.*, vol. 8, no. 5, pp. 521–534, Sep. 2002. [Online]. Available: http://dx.doi.org/10.1023/A:1016598314198

[50] R. Pickholtz, D. Schilling, and L. Milstein, "Theory of spread-spectrum communications–a tutorial," *IEEE Transactions on Communications*, vol. 30, no. 5, pp. 855–884, 1982.

[51] J. G. Proakis, *Digital signal processing: principles algorithms and applications*. Pearson Education India, 2001.

[52] 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*, vol. 59, no. 2, pp. 774–786, 2011.

[53] A. Ribeiro and G. B. Giannakis, "Bandwidth-constrained distributed estimation for wireless sensor networks - part i: Gaussian case," *IEEE Transactions on Signal Processing*, vol. 54, no. 3, pp. 1131–1143, 2006.

[54] R. T. Rockafeller, *Convex Analysis*, ser. Princeton Landmarks in Mathematics and Physics. Princeton University Press, 1996.

[55] Y. E. Sagduyu, R. A. Berry, and A. Ephremides, "Jamming games in wireless networks with incomplete information," *IEEE Communications Magazine*, vol. 49, no. 8, pp. 112–118, August 2011.

[56] S. Shafiee and S. Ulukus, "Mutual information games in multiuser channels with correlated jamming," *IEEE Transactions on Information Theory*, vol. 55, no. 10, pp. 4598–4607, Oct 2009.

[57] E. Soltanmohammadi, M. Orooji, and M. Naraghi-Pour, "Decentralized hypothesis testing in wireless sensor networks in the presence of misbehaving nodes," *IEEE Transactions on Information Forensics and Security*, vol. 8, no. 1, pp. 205–215, 2013.

[58] Z. B. Tang, K. R. Pattipati, and D. L. Kleinman, "Optimization of detection networks. ii. tree structures," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 23, no. 1, pp. 211–221, Jan 1993.

[59] W. P. Tay, J. N. Tsitsiklis, and M. Z. Win, "Data fusion trees for detection: Does architecture matter?" *IEEE Transactions on Information Theory*, vol. 54, no. 9, pp. 4155–4168, Sept 2008.

[60] J. Tsitsiklis, "Decentralized detection by a large number of sensors," *Math. Control, Signals, Systems*, vol. 1, no. 2, pp. 167–182, 1988.

[61] J. N. Tsitsiklis, "Decentralized detection," in *Advances in Signal Processing*, H. V. Poor and J. B. Thomas, Eds. JAI Press, 1993, vol. 2, pp. 297–344.

[62] J. Tsitsiklis, "Extremal properties of likelihood-ratio quantizers," *Communications, IEEE Transactions on*, vol. 41, no. 4, pp. 550–558, Apr 1993.

[63] J. Tsitsiklis and M. Athans, "On the complexity of decentralized decision making and detection problems," *IEEE Transactions on Automatic Control*, vol. 30, no. 5, pp. 440–446, May 1985.

[64] P. K. Varshney, *Distributed Detection and Data Fusion*. Springer, New York, 1997.

[65] V. V. Veeravalli and P. K. Varshney, "Distributed inference in wireless sensor networks," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 370, no. 1958, pp. 100–117, 2012.

[66] A. Vempaty, K. Agrawal, H. Chen, and P. Varshney, "Adaptive learning of byzantines' behavior in cooperative spectrum sensing," in *Proc. IEEE Wireless Communications and Networking Conf. (WCNC)*, 2011, pp. 1310–1315.

[67] A. Vempaty, O. Ozdemir, K. Agrawal, H. Chen, and P. Varshney, "Localization in wireless sensor networks: Byzantines and mitigation techniques," *IEEE Transactions on Signal Processing*, vol. 61, no. 6, pp. 1495–1508, 2013.

[68] A. Vempaty, L. Tong, and P. K. Varshney, "Distributed inference with byzantine data: State-of-the-art review on data falsification attacks," *IEEE Signal Processing Magazine*, vol. 30, no. 5, pp. 65–75, 2013.

[69] R. Viswanathan and P. K. Varshney, "Distributed detection with multiple sensors i. fundamentals," *Proceedings of the IEEE*, vol. 85, no. 1, pp. 54–63, 1997.

[70] D. Vouyioukas, "A survey on beamforming techniques for wireless mimo relay networks," *International Journal of Antennas and Propagation*, vol. 2013, 2013.

[71] W. Weil, Ed., *Spatial Point Processes and their Applications*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007, pp. 1–75. [Online]. Available: http://dx.doi.org/10.1007/978-3-540-38175-4_1

[72] B. Widrow, J. R. Glover, J. M. McCool, J. Kaunitz, C. S. Williams, R. H. Hearn, J. R. Zeidler, J. E. Dong, and R. C. Goodlin, "Adaptive noise cancelling: Principles and applications," *Proceedings of the IEEE*, vol. 63, no. 12, pp. 1692–1716, Dec 1975.

[73] M. Wiese, J. Nötzel, and H. Boche, "A channel under simultaneous jamming and eavesdropping attack: Correlated random coding capacities under strong secrecy criteria," *IEEE Transactions on Information Theory*, vol. 62, no. 7, pp. 3844–3862, July 2016.

[74] A. Wood and J. A. Stankovic, "Denial of service in sensor networks," *IEEE Computer*, vol. 35, no. 10, pp. 54–62, October 2002.

[75] W. Xu, K. Ma, W. Trappe, and Y. Zhang, "Jamming sensor networks: attack and defense strategies," *IEEE Network*, vol. 20, no. 3, pp. 41–47, 2006.

[76] Q. Zhang, P. K. Varshney, and R. D. Wesel, "Optimal bi-level quantization of i.i.d. sensor observations for binary hypothesis testing," *IEEE Transactions on Information Theory*, vol. 48, no. 7, pp. 2105–2111, 2002.

[77] Q. Zhao and B. M. Sadler, "A survey of dynamic spectrum access," *IEEE Signal Processing Magazine*, vol. 24, no. 3, pp. 79–89, 2007.

VITA

NAME: Venkata Sriram Siddhardh Nadendla

PLACE OF BIRTH: Rajahmundry, Andhra Pradesh, India

DATE OF BIRTH: April 18, 1986

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

Sri Chandrasekharendra Saraswathi Viswa Mahavidyalaya, Kanchipuram, TN, India.

Louisiana State University, Baton Rouge, LA, USA.

DEGREES AWARDED:

B.E.(ECE), 2007, Sri Chandrasekharendra Saraswathi Viswa Mahavidyalaya, Kanchipuram, TN, India.

M.S.(EE), 2009, Louisiana State University, Baton Rouge, LA, USA.