Estimation in Wireless Sensor Networks with Security Constraints

Xiaoxi Guo

Submitted in total fulfilment of the requirements of the degree of

Doctor of Philosophy

Department of Electrical and Electronic Engineering
THE UNIVERSITY OF MELBOURNE

January 2016

Abstract

THIS thesis presents secure and energy-efficient power allocation algorithms for wireless sensor networks used in distributed estimation. We focus on power allocation policies that minimize the distortion level or distortion outage probability at a remote receiver when the network is under eavesdropping or active attacking. Various power allocation strategies are investigated and analysed under different channel assumptions and wireless sensor network models.

We first look at power allocation for distributed estimation in a wireless sensor network in the presence of an eavesdropper, where sensors send their observations to the fusion center through orthogonal multi-access channels, which at the same time is overheard by the eavesdropper. Depending on the available channel state information (CSI) at the fusion center and number of transmit antennas at sensors, we obtain optimal or suboptimal power allocation schemes that minimize estimation distortion errors at the fusion center subject to power constraint and secrecy constraints. Asymptotic expressions are obtained for the long-term distortion at the fusion center as the number of sensors or number of antennas becomes large.

We then focus on problems that require minimizing the distortion outage probability at the fusion center for distributed estimation in a wireless sensor network when an eavesdropper is present. Applying a rigorous probabilistic power allocation technique, we obtain power allocation schemes for the full CSI case. We study suboptimal power control policies to reduce the high computational cost for the case of a large number of sensors or receive antennas. Artificial noise techniques are also considered to assist with reducing the distortion outage probability at the fusion center. Simulation results show significant performance improvements when artificial noise techniques are employed.

Finally we look at the power transmission strategy of distributed estimation under the denial-of-service (DoS) attack in a single sensor network, where an active attacker jamming the communication channel attempts to reduce estimation quality at the fusion center. We study a game theoretic approach to capture the conflicting nature of both parties in DoS attacks for fading and non-fading scenarios, where the game is played simultaneously at both the sensor and attacker. Backwards induction and Nash Q-learning techniques are investigated to derive a strategy pair at a Nash equilibrium. When fading is present, apart from the full CSI, we also look at the partial CSI where type-contingent power transmission strategies at a Bayesian Nash equilibrium is obtained.

Declaration

This is to certify that

1. the thesis comprises only my original work towards the PhD,
2. due acknowledgement has been made in the text to all other material used,
3. the thesis is less than 100,000 words in length, exclusive of tables, maps, bibliographies and appendices.

Xiaoxi Guo, January 2016

Acknowledgements

Foremost, I would like to express my greatest gratitude to my supervisors Dr. Alex Leong and Prof. Subhrakanti Dey for their continuous support and endless patience of my Ph.D study and research. I would like to offer my sincere appreciation to Dr. Alex Leong for his motivation and technical insights. His guidance have helped me in all the time of research and writing of the thesis. I would also like to thank him for offering me financial support. Without him, I could not imagine how far I could go on my research journey. Additionally, I truly appreciate Dr. Yuanyuan He for her collaboration, invaluable advice and great help.

I would also like to extend my sincere appreciation to my advisory committee members Prof. Girish Nair and Prof. Erik Weyer, for their advice and support during my Ph.D candidature. Furthermore, I would like to thank Prof. Margreta Kuijper, Prof. Girish Nair and Prof. William Shieh for providing teaching and tutoring opportunities throughout the years. Many thanks to the staff at the University of Melbourne and CEET for providing a pleasant and friendly research environment and administrative tasks.

Looking back of all the joys during the past four years, my research life wouldn't have been this colorful without a number of friendship. I would like to thank all my friends who were there with me: Ehsan Nekouei, Chih-Yu Hsu, Athipat Limmanee, Meng Wang, Yuanyuan He, Adel Ahmadi Tabatabaei, Vijay Vijayalayan, Phee-Lep Yeoh, Hamid Khodakarami, Chrispin Gray, Dinuka Kudavithana, Rajitha Senanayake, Feng Li, and Gitansh Khirbat. I wish them all the best for their future endeavors.

Last but not least, I would like to express my deepest graduate to my family Qi Guo, Shuyu Wang and Yue Li, for their unconditional love and support. A special mention goes to my late grandpa Kerang Guo, who always believed in me. This thesis is dedicated

to him and my family.

To my family and my late grandpa, for their unconditional love and endless support

Contents

List of Figures

List of Tables

Chapter 1
Introduction

1.1 Motivation of the Thesis

SENSORS, which can sense, measure and gather information from the environment, provide human beings different ways to 'see' the world. Benefiting from Micro-Electro-Mechanical-System technology, wireless communications and digital electronics, sensors are now smaller in size and produced in a more economically friendly manner [18, 22]. A wireless sensor network (WSN) is formed when a few sensors are deployed in a geographical area. A typical WSN, shown in Figure 1.1, consists of some small, inexpensive and low-cost sensors, which can be used to monitor temperature, humidity, pressure, noise level, etc, and may communicate with a remote processor over wireless links. Because of the wide range of applications, WSNs have gained world wide attention and attracted much research interest.

Typically, sensors are equipped with a power source of limited power budget. If sensor nodes are deployed in a hostile or inhabitable environment to monitor the physical phenomenon, recharging or changing the power source can be very difficult. Yet, sensors need to have a lifetime long enough to fulfill the application requirements; in fact, someday, we expect sensors to be cheap enough that they are discarded rather than recharged. In general, the energy consumption in a sensor node mainly includes three components: (1) a sensing subsystem or sensor transducer for local data acquisition from the physical surrounding environment; (2) a radio communication subsystem for data transmission; (3) a processing subsystem for data storage and processing. Among the three parts, sensor transducer consumes the least amount of power. Although the en-

ergy consumption in data processing is large, it is considerably less compared to data transmission [37, 82, 84, 109]. In some cases, it is possible for sensors to scavenge energy from the environment, such as by using a solar panel; however, in order to perform a continuous and stable service a power supply, like batteries, is required [3].

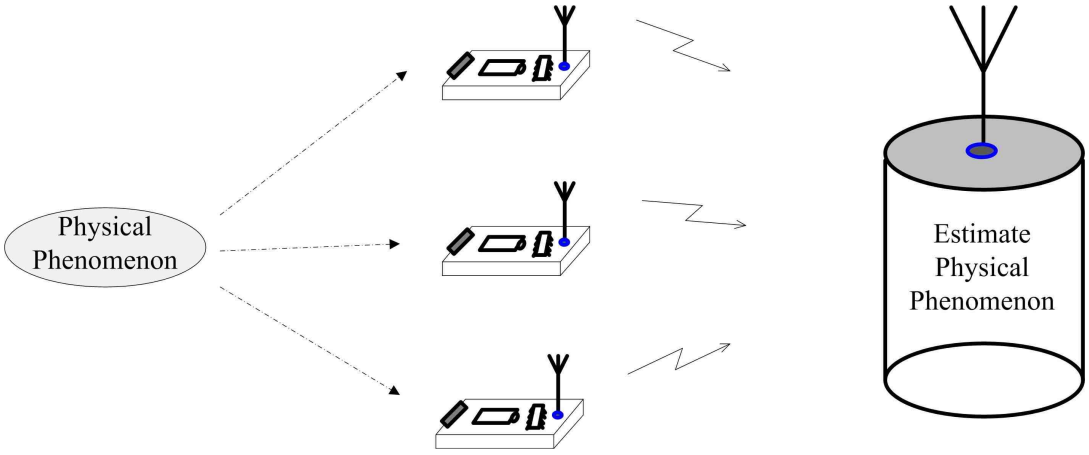

Figure 1.1: A WSN topology with a central controller.

A WSN normally has little or no infrastructure and sensor nodes are randomly placed into a field (in an ad hoc manner) to monitor a region to obtain data about the environment. As the performance of a sensor network application largely depends on the lifetime of the network, how to efficiently manage the power consumption of each sensor is crucial for a wireless sensor network. Power control is a way to set transmission powers for the sensors in a WSN. Generally, there are two ways to conduct power control in a wireless sensor network: centralized and distributed. In centralized power control, a sensor node does not generate its own power schedules; instead, it executes schedules generated and sent from a central controller. The node simply collects information and then forwards them to the central controller using the power allocation schedule it receives. Although centralized power management may suffer some vulnerabilities such as heavy traffic loads as the remote processor collects channel information from all the active sensors, it is more likely to achieve the global optimal power scheduling if all collected information is properly taken into consideration. In a distributed power control scheduling, each sensor node is autonomous. The power transmission scheme is calculated by sensors based on the local information it collects from the surrounding

environment[1].

WSNs are a special type of network. It shares some commonalities with a typical computer network, but also exhibits many unique characteristics. Unlike wired networks, wireless sensors broadcast their message to the medium, which leaves WSNs open to threats and risks ranging from passive eavesdropping to active interfering [107]. In passive eavesdropping, an adversary can easily retrieve valuable or private data from the transmitted packets. For example, by 'listening' to the water consumption and temperature reading of sensor networks inside a house, the tenants' personal daily activities are exposed. Traditional data encryption techniques can partially solve eavesdropping problems but suffer many vulnerabilities and are at a significant energy cost [12, 87]. First of all, if the adversary has sufficiently large computational power, cryptographic schemes with small key sizes may provide little secrecy. Apart from this, experiments in [37, 82] have shown that the energy cost of each bit transmitted by sensors in WSNs consumes about as much power as executing 800 to 1000 instructions. A high security level WSN deploying cryptography techniques suffers huge energy cost for the message expansion in the security mechanisms. Symmetric key cryptographic mechanisms are proposed to reduce the computational intensity, in which it uses a single shared key between the sensor nodes and the remote controller for both encryption and decryption. However, apart from a big challenge in key management, in quantum computing, Grover's algorithm allows one to break a symmetric key of complexity $O(N)$ in $O\left(\sqrt{N}\right)$ time [31]. One could foresee that in near future, cryptography techniques with small key size poses a security threat to WSN.

Active external attacks disrupt network functionality by deploying denial-of-service (DoS) attacks that diminish or attempt to reduce a network's capacity to perform its expected functions [102]. DoS attacks can happen at the physical layer via radio jamming. Even with less powerful jamming sources, an adversary is capable of disrupting the communications in the WSN [13, 81, 87, 102]. Some attackers specially target sensors' power supplies by keeping their radio communication on at all times until the battery of sensors

[1] In a distributed sensor network, the sensors would reach an agreement regarding a certain quantity of interest via consensus algorithm that specifies the information exchange between an agent and all of its neighbour on the network [77].

runs out or imposes an unnecessary working load. For example, a deceptive jammer may simply transmit random signals at the same operation frequency as the sensor to increase the packet dropping rate at the receiver side, the node hence needs to resend the same packet to the central control until its power drains out. A common method to defend against physical layer jamming is to deploy Spread-Spectrum communication; however, the power limited characteristic of sensors prevents us from using such techniques in WSNs. If sensors could identify a jamming attack, an ideal solution would be to simply put the node into sleep mode to save energy. Nevertheless, either way, the adversary causes degradation of performance in real-time applications.

Physical layer security explores the characteristic of wireless channel, such as fading and noise, providing a new security paradigm for the improvement of the communication security. The aforementioned observations and facts motivate us to conduct a series of studies on the joint issue of energy consumption and physical layer secrecy posed by wireless sensor networks in distributed estimation.

1.2 Literature Review

Many works have considered the estimation problem in an energy-constrained wireless sensor network. In a WSN, the sensors take measurements of a source independently and then transmit the measurements to a fusion center (FC) over a wireless link. After receiving the measurements from the sensors, the remote processor then attempts to reconstruct an estimate of the physical quantity the sensors observed. Many cross-layer optimization schemes have been proposed to maximize the lifetime of WSNs or enhance the estimation accuracy. In particular, the authors in [15] studied the optimal power allocation strategies for minimizing estimation error under a total power constraint, and minimizing total power subject to a distortion constraint, with orthogonal multiple access channels (MACs) adopted between the sensors and the FC. With the assumption of perfect synchronization between sensors and the FC, the authors in [105] adopted coherent MAC and derived the optimal power scheduling. However, these optimal power schemes were obtained based on static channel environments which may not always

meet a strict distortion constraint in the varying channel conditions. In [6], the block fading channel was considered in WSNs where the authors solved the optimal power allocation schemes that minimize the total transmission power while subject to estimation distortion constraints. Under fading channel scenarios, due to the randomness of the fading channels, the distortion becomes a random variable as a function of the channel gains, and one may not be able to satisfy the distortion constraint at all times. This leads to the notion of estimation outage or distortion outage, which is defined as the probability that the estimation distortion exceeds a certain threshold [15]. Applying similar techniques as developed in [64], the authors in [99] investigated the distortion outage performance for wireless sensor networks, and derived optimal power allocation schemes that minimize the distortion outage probability. They also showed that in Rayleigh fading and with more than one sensor the outage probability could be made to zero with finite total powers, and obtained an approximation for the minimum number of sensors needed.

However, because of the broadcast nature of wireless communication, security of information transfer via wireless networks remains a challenging issue. Eavesdropping and jamming are two primary attacks at the physical layer of a wireless network. Generally, channel and code design, secure capacity, and the power and signal design approaches are the five major categories in the existing physical layer security techniques [92]. The channel design approaches involve explore the channel characteristics including radio frequency fingerprinting, randomization of multiple-input multiple-output (MIMO) transmission coefficient and algebraic channel decomposition multiplexing precoding. Error correction coding and spread spectrum coding are common techniques used to secure the transmission when code design approaches are considered. The notion of perfect secrecy, introduced by Shannon [90], provides a different perspective on the data confidentiality. Under perfect secrecy, the signal received by an eavesdropper does not provide any additional information about the transmitted message. Later, in the 1970s, Wyner suggested that perfect secrecy is achievable as long as the channels that are unknown to unauthorized users are more noisy than that of the authorized users [103]. In other words, if the adversary's channel is a degraded version of the legitimate receiver's, reliable information can be received at the legitimate user

without the eavesdropper being aware of almost any useful information. The authors in [30] considered the secure transmission over an ergodic fading channel in the presence of an eavesdropper. In their setup, the transmitter is assumed to have access to the channel gains of both the legitimate receiver and the eavesdropper. The secrecy capacity under this full channel state information (CSI) assumption was derived and served as an upper bound for the case when the transmitter only knows the CSI of the legitimate receiver. In [59], the authors investigated the secrecy capacity region for the parallel Gaussian broadcast channel, where a sensor node transmits common message to two receivers with one of which receiving confidential message from the sensor as well. The secret information needs be kept confidential from the other receiver in the network. Assuming that all parties are aware of the CSI, the authors obtained the secrecy capacity region and the corresponding optimal power allocations achieving the boundary of the capacity region. The secrecy capacity of the MIMO wiretap channel is considered in [76], where the number of antennas is arbitrary for both the transmitter and two receivers. The authors proved that the perfect secrecy capacity is the difference of the capacity of the legitimate user and the one of the eavesdropper.

The other two categories of physical layer security techniques are power and signal design approaches which involve the employment of artificial noise injection and directional antennas. The authors in [27, 110] proposed artificial noise schemes to ensure perfectly secure communications. Depending on the number of antennas at the transmitter, artificial noise is generated on the null space of the legitimate receiver's channel using either multiple antennas or collaboration among multiple sensor nodes. This approach discriminates against the non-legitimate receiver's channel while keeping the channel of the authorized receiver unimpaired. It was also shown that secret communication can be established even if the adversary has better channel conditions than the intended receiver.

It is known that the mutual information between the input and the output of a channel is at the core of information theory; given an input signal it measures the amount of coded information that can be reliably transmitted through a channel. Some literature above considered network security and privacy issues from an information-theoretic per-

spective. Its counterpart, minimum mean square error (MMSE), is a fundamental quantity in estimation theory, which indicates how accurately the input signal can be retrieved from the channel output.

A fundamental connection between information theory and estimation theory was discovered in [32]. The authors found that regardless of the input distribution, the derivative of the mutual information in nats w.r.t. SNR is equal to half the MMSE, as long as the input signals are observed through an additive Gaussian noise channel. This fundamental relationship and its generalizations [80] have been shown to be useful to provide insightful and simple proofs for deriving the capacity region of Gaussian multi-receiver wiretap channel. The authors in [10] used such techniques to derive a closed-form expression for the secrecy capacity of the MIMO Gaussian wiretap channel under a power constraint. The proof provides the missing intuition regarding the existence and construction of an enhanced degraded channel that does not increase the secrecy capacity. In [71], the Chief Executive Officer (CEO) problem is investigated in which a center controller (or CEO) attempts to minimize the estimate distortion from the noisy observations of a random source it receives. The authors related the equivocation rate[2] to the normalized distortion at the eavesdropper in the CEO problem with additional secrecy constraints, where they showed that the estimation error at the eavesdropper is an upper bound of the equivocation rate.

These works motivate us to study estimation problems in an energy-constrained wireless sensor network, while looking at the physical layer secrecy from the signal estimation viewpoint by employing MMSE as the security performance metric. To secure the transmission in scenarios of passive eavesdropping, the signal processing side of transmission schemes are widely explored in [4, 46, 52, 61, 67]. In particular, distributed detection with censoring sensors was investigated in [67] where the authors used divergence as a security metric and set it to zero at the adversary to guarantee that no information is leaked. The authors in [61] focused on secret transmit beamforming approaches. To secure the confidential transmission between the legitimate receiver and transmitter, the maximum

[2]The equivocation rate is defined as the conditional entropy of the confidential information given the signal seen by the eavesdropper. It indicates the eavesdropper's uncertainty about the confidential message [76].

allowable Signal-to-Interference-and-Noise Ratios (SINR) at the eavesdroppers was proposed. In [52], the performance of a link adaptation and untrusted relay assignment framework was investigated, where the authors used bit error rate to measure a reliable transmission against untrusted relays.

In the additional to passive eavesdropping, an adversary could also actively interferes the quality of message received by the remote processor. In the framework of DoS attacks, both the attacker and defender behave selfishly, in which the defender intends to protect against DoS attack and establish a reliable link for data delivery; whereas the attacker has completely contradictory objectives. An effective defense strategy for the sensor should not only depend on its own behaviors but also take into consideration the actions of the opponent. Game theory provides a framework to capture and analyze the conflicting nature of DoS attacks in WSNs. Instead of a static analysis focusing on only one side transmitter or attacker, the game-theoretic approach is able to model the interactions between two conflicting parties, perform tactical analysis of potential threats and provide strategic suggestions against such threats.

In [57], jamming and anti-jamming in multi-channel wireless communication systems were studied and modeled as a zero-sum stochastic game. The Nash equilibrium of the game was analyzed in which a linear quadratic function is used as the payoff function to be maximized/minimized by the sensor/jammer. From an information theoretic perspective, the authors in [47] considered a zero-sum mutual information game on MIMO Gaussian Rayleigh fading channels. With the assumption that the jammer has access to the channel input, the authors showed that at the Nash equilibrium the amount of damage caused by the jammer to the communication is only as much as the one without the extra channel information. A similar setup was investigated in [88] where the authors studied a muli-user system under correlated jamming. Depending on the user's channel knowledge at the jammer, the Nash equilibrium may or may not be exist. More recently, the authors in [58] investigated remote state estimation of cyber-physical systems under SINR-based DoS attack. A Markov game framework is built to model the online interactive decision-making process. The authors then applied a modified Nash Q-learning algorithm to solve the associated optimality (Bellman) equations. The aforementioned

work is based on either complete channel information at the jammer or static channels which do not take into account the characteristic of incomplete information for DoS attacks.

1.3 Contributions and Overview of the Thesis

The focus of this thesis is to investigate energy-efficiency transmission power algorithms for passive eavesdropping issues as well as active interfering issues in distributed estimation of a wireless sensor network. In particular, we introduce the performance metric, minimum mean square error, to measure how accurately an input signal can be retrieved at the legitimate receiver and how distorted the signal would be seen at the passive eavesdropper. When the message is severely distorted at the receiver it is almost unlikely can be recovered. Therefore, the minimum mean square error not only tells us the security level of a wireless network by observing it at the eavesdropper side but also indicates the estimation accuracy when it is considered at the FC side. Optimal and suboptimal power allocation policies and transmitting strategies are studied and obtained for different scenarios for a wireless sensor network. In the thesis, we start by discussing some background and fundamentals of estimation in signal processing in Chapter 2. The concept of game theory is also presented, which is used in the framework of DoS attacks. Brief summaries of each remaining chapters in the thesis are presented below.

Chapter 3 In this chapter, we focus on the performance of distributed estimation in a wireless sensor network with the presence of an eavesdropper. Under amplify-and-forward transmission and orthogonal multi-access protocol, we obtain power transmission schemes that minimize estimation error at the remote processor while keeping the information acquired by the adversary at an acceptable level. Depending on the sensors' awareness about eavesdropper's CSI or/and the number of antennas equipped on each sensor, we investigate various power transmit policies that satisfy secrecy constraints in every transmission time slot or over a few fading blocks. When a sensor has multiple communication antennas we show that zero information leakage can be achieved if the sensor knows the adversary's channel; otherwise, the sensor could broadcast additional

noise on its channel null space to distract the eavesdropper while keeping the legitimate receiver unaffected. A similar concept is introduced in the multiple sensors scenario, where some sensors serve as friendly relays to produce random signals that can be cancelled off at the intended receiver. For a given power budget, we obtain the asymptotic expression of the expected distortion at the FC and show that the distortion decreases to a constant at the rate $1/N_t$ for the multiple antennas scenario and at a rate of $1/K$ for the multiple sensors scenario, where N_t and K stand for the number of antennas and sensors respectively. Numerical results show that given the same total number of transmitting antennas, the multiple-antenna sensor network is superior to the performance of the multiple-sensor single antenna network.

Chapter 4 In this chapter, we formulate a distortion outage minimization problem for a wireless sensor network with multiple receive antennas at both the eavesdropper and the legitimate receiver. We investigate power allocation algorithms that minimize distortion outage probability at the FC, where sensors apply amplify-and-forward techniques to transmit their signals to the receiver via orthogonal multi-access channels. After considering the full CSI case we extend the problem and study a more practical scenario where only the adversary's statistical channel information is available. To compensate for the high computational complexity caused by large numbers of sensors or a large number of antennas at the FC, we propose a sub-optimal scheme for implementation. In the additional to single-antenna sensors, we also look at a multiple-antenna scenario. In this case, the distortion outage at the FC can be dramatically reduced and in some cases eliminated at the FC in both the full CSI and partial CSI cases.

Chapter 5 In this chapter, we study a game theoretic approach for the distributed estimation in a wireless sensor network, where the single sensor is under DoS attacks. We first look at an additive white Gaussian noise (AWGN) environment in which players interact with each other only once. Next, a dynamic game is formulated. The idea is that both players will interact with each other many times. Although the players are assumed to be ignorant about the opponent's actions, they could learn each other's strategies from the interactions. When a finite number of games is played, a two-person feedback game is considered. In this scenario, we propose an algorithm to recursively derive behavioral

strategies under a limited power budget. Multiple agents Q-learning is also studied to derive a Nash equilibrium strategy for the infinite horizon scenario where players keep playing the same static game. In discrete fading, apart from complete information game, we study scenarios where players are unaware of the channel type of its opponent, i.e., incomplete information game. Such games are modeled as Bayesian games, and the power allocations are adjusted according to the player's own channel information and the belief it has on the statistical channel information of the other.

Chapter 6 This chapter provides concluding remarks of the thesis, and presents some future research ideas related to the thesis.

1.4 List of Publications

Journal Papers

1. X. Guo, A. S. Leong and S. Dey, "Distortion Outage Minimization in Distributed Estimation with Estimation Secrecy Outage Constraints," *Submitted to IEEE Transactions on Signal and Information Processing over Networks*, Oct. 2015.

2. X. Guo, A. S. Leong and S. Dey, "Estimation in Wireless Sensor Networks with Security Constraints," *submitted to IEEE Transactions on Aerospace and Electronic Systems*, Nov. 2015.

3. "Sensor Power Allocation for Distributed Detection with Limited Channel Feedback," *in preparation*.

4. "A Game-Theoretic Approach to DoS attack in Distributed Estimation," *in preparation*.

Conference Paper

1. X. Guo, A. S. Leong and S. Dey, "Power Allocation for Distortion Minimization in Distributed Estimation with Security Constraints," *in 2014 IEEE 15th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Jun. 2014, pp. 299–303.

2. X. Guo, A. S. Leong and S. Dey, "Power Allocation for Estimation Outage Minimization with Secrecy Outage Constraints," *in Proc. Australian Communications Theory Workshop*, Melbourne, Australia, pp. 71-76, Nov. 2015.

Chapter 2
Background and Fundamental Concepts

This chapter discusses the background and fundamental concepts that will be referred to in the thesis. We first introduce minimum mean square error (MMSE), which is often chosen to be as the error criterion in signal estimation. The linear MMSE estimator is obtained when it is a fixed linear function of the measured random variable. We then present the fundamental concepts of game theory and three game theoretic models that are applied later in our work.

2.1 Signal Estimation

The primary goal of estimation theory is to extract information from noise-corrupted observations or signal waveforms. One of the key procedures is the introduction of an error criterion that measures the error between the desired quantity and its estimates [78]. Often this estimate is chosen to be the one with minimum mean square error (MMSE).

2.1.1 Minimum Mean Square Error Estimators

Suppose at time k we have a set of L observations about a parameter θ, represented by the $L-$dimensional vector $\mathbf{X} = [x[1] \dots x[L]]^T$. Based on the knowledge of the PDF of θ, we wish to obtain an estimate of this parameter, denoted as $\hat{\theta}$, so as to minimize the mean square error between the actual value of θ and our estimate $\hat{\theta}$. Specifically, with additional information in the form of observation vector $\mathbf{X}$, we choose $\hat{\theta}$ to minimize

$$\mathbb{E}\left[\left\|\theta - \hat{\theta}\right\|^2 \middle| \mathbf{X}\right].$$

From probability theory it is well-known [91] that the mean squared error is minimized by choosing the estimator be the conditional expected value of the parameter given the observation, i.e.,

$$\hat{\theta} = \mathbb{E}\left[\theta \middle| \mathbf{X}\right].$$

The conditional expectation $\mathbb{E}\left[\theta \middle| \mathbf{X}\right]$ would be easy to compute if one could determine the conditional density $f_{\theta|\mathbf{X}}\left(\theta \middle| \mathbf{X}\right)$. However, this may not be the case. In general, finding $f_{\theta|\mathbf{X}}\left(\theta \middle| \mathbf{X}\right)$ is difficult. A useful and widely used compromise is to restrict the estimates $\hat{\theta}$ to be a fixed linear function of each observation element.

2.1.2 Linear Minimum Mean Square Error Estimators

When we restrict the estimator to be a fixed linear function of the measured random variable and choose the linear relationship so as to minimize the mean square error, we obtain the linear minimum mean square error (LMMSE) estimator. To be more specific, an estimator of the form [49]

$$\hat{\theta} = a_0 + \sum_{l=1}^{L} a_l x[l]$$

and with weighting coefficients a_l's to minimize the mean square error $\mathbb{E}\left[\|\theta - \hat{\theta}\|^2\right]$ is the LMMSE estimator, where the expectation is with respect to the PDF $f_{(\mathbf{X},\theta)}\left(\mathbf{X}, \theta\right)$.

The coefficient a_l of the LMMSE estimator can be found from the first two moments of PDF $f_{(\mathbf{X},\theta)}\left(\mathbf{X}, \theta\right)$. To minimize $\mathbb{E}\left[\|\theta - \hat{\theta}\|^2\right]$, we differentiate it with respect to a_l for $l = 0, 1, \ldots, L$, and set each of the derivatives to zero. Let $\mathbf{a} = \left[a_1, \ldots, a_L\right]^T$, we have

$$\begin{aligned} a_0 &= \mathbb{E}\left[\theta\right] - \sum_{l=1}^{L} a_l \mathbb{E}\left[x[l]\right], \\ \mathbf{a} &= \Sigma_{xx}^{-1} \Sigma_{x\theta}. \end{aligned}$$

Here Σ_{xx} is the $L \times L$ covariance matrix of $\mathbf{X}$, and $\Sigma_{x\theta}$ is the $1 \times L$ cross-covariance vector. Assuming the means of θ and $\mathbf{X}$ are zero, the LMMSE estimator is then $\Sigma_{\theta x} \Sigma_{xx}^{-1} \mathbf{X}$ and the

associated mean squared error is

$$\mathbb{E}\left[\|\theta-\hat{\theta}\|^2\right]=\Sigma_{\theta\theta}-\Sigma_{\theta x}\Sigma_{xx}^{-1}\Sigma_{x\theta}.$$

The theory of signal estimation has been widely used in many areas. In wireless communications such as data transmission to a remote processor, estimation theory provides a guide to the design of effective communication receivers and/or efficient transmission strategy at transmitters. Throughout the thesis, a remote fusion center is assumed to apply the linear MMSE estimator to reconstruct an estimate of the physical quantity observed. We then investigate the best power allocation schemes for a network under different security issues.

2.2 Game Theory

Game theory is a collection of analytical tools helping decision makers to unlock the insights of the 'games' they play. In social science, it entails to understand human behaviors such as bidders competing in an auction, firms fighting for business, and the candidates competing for votes [79]. In engineering and computer science, it assists with infrastructure planing such as limited resources allocation, traffic congestion control, and pricing of the Internet service [65]. In economics and finance, it is used to deal with trade and production, such as to analyze stock markets and exchange rates [20]. Game theory provides comprehensive ways to capture and analyze the interactive decisions among multiple decision makers, thus suggesting reasonable strategies.

A game describes strategic interaction of decision makers in a certain environment; hence the basic entities in a game theoretic model are decision makers, called *players* (or *agents* in some context). If the players behave individually in a non-cooperative manner or players have their own interests which are conflicting with others', we are in the realm of *noncooperative game theory* [8]; if the individuals work as a group or in a collective move, we then call it *cooperative games*. In a game, each player has its own interests or objectives (called *payoffs* or *costs*), such as to reduce the bidding price in an auction, or to gain the best interests in limited resource allocation. In order to achieve such objectives in a game,

each player has a series of moves, called *strategy*, and based on which a player decides its *actions* at a given situation.

The concept of optimality in multi-person decision making is not well-defined; however, the Nash equilibrium solution is generally considered as 'optimal', which defines the best strategies for the players so that no players could do strictly better than currently achievements by unilaterally adopting another strategy [8]. Below, we will only discuss the following three game theoretic models for *two-person zero-sum games*[1] that are applied in Chapter 5. Details can be found in [8,21,79,95] for a more comprehensive discussion.

2.2.1 Static Games of Complete Information

Depending on common knowledge among the players, a game can be classified as complete information or incomplete information. A game of complete information [95] requires players are perfectly informed of:

- All the possible actions of all the players.
- All the possible outcomes, including the effect on the outcomes from each combination of actions of all players.
- The preferences of each and every player over all outcomes.

Static games with complete information is the simplest game model describing situations when players simultaneously and independently choose a decision. In other words, all players choose their actions at the same moment without any knowledge of the decisions made by their counterparts[2].

In general, there are two different representations for two-person games with finite strategy sets: *matrix form* (also known as *normal form*) and *extensive form*. In matrix games, e.g. Table 2.1 [95], each entry of the matrix is an outcome of the game corresponding to a certain pair of strategies used by the two players, where M, F stand for the possible

[1] As the name suggests, a two-person zero-sum game involves only two players in a game with each having the completely opposite interests of another. In such a game, the sum of the objective functions of the two players is zero, or can be made to zero by positive scaling and/or translation that is independent of the players' actions [8].

[2] As players have complete information and move simultaneously, this type of game is also known as a game of complete but imperfect information.

actions at **P1** and m, f represent actions at **P2**. Equivalently, applying a tree structure, extensive form uses nodes and branches to provide a more explicit representation about the interactive behaviors of the players, including the playing orders and available information to each player in the decision process [8]. Figure 2.1 is the extensive form of the same game as shown in Table 2.1, where the dotted circle indicates the available information at players (known as the *information set*) at the time of his play. In Figure 2.1 it says **P2** is not clear about which branches he is in, and this implies both players act simultaneously.

		P2	
		m	f
P1	M	4, 4	-1, 5
	F	5, -1	1, 1

Table 2.1: Prisoner's Dilemma Game in matrix form.

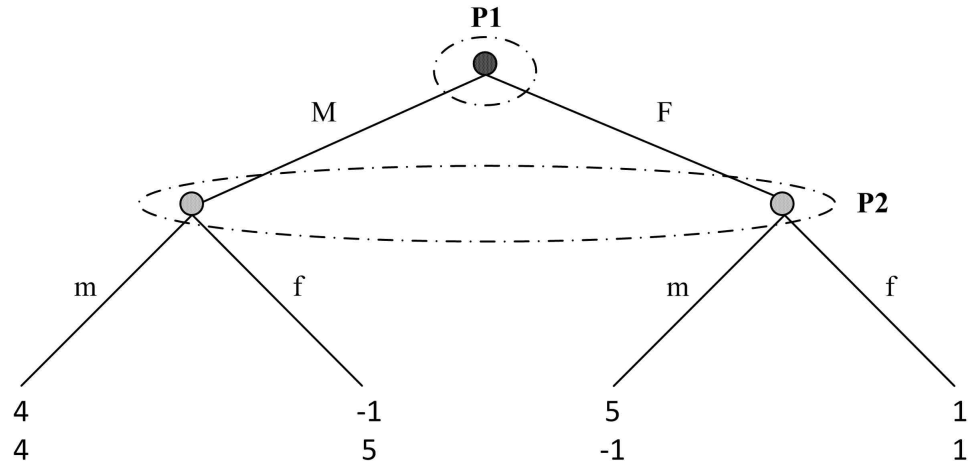

Figure 2.1: Prisoner's Dilemma Game in extensive form.

2.2.2 Multi-Stage Games

A multi-stage game is defined as a finite sequence of normal-form stage-games, in which players are allowed to act more than once [95]. At each stage, the game is an independent and well-defined complete but imperfect game that is played by the same group of play-

ers. The total payoff of the game is evaluated by the sum or discounted sum of outcomes at each stage. In a multi-stage game, it is generally assumed that all players are able to observe the outcome of the game played at each stage; in other words, at the time of action, each player has perfect information regarding the payoffs of previous stage. This is also known as multi-act feedback games in [8], where the two-person zero-sum game is discussed.

		P2	
		l	r
P1	L	0, 0	-4, -1
	R	-1, -4	-3, -3

Table 2.2: Revenge Game in matrix form.

An example is provided. Suppose that after completing the Prisoner's Dilemma Game, the same two players in Section 2.2.1 play a different game with possible actions L, R at **P1** and l, r at **P2**, as shown in Table 2.2 [95]. The extensive form of this multi-stage game is given in Figure 2.2, which explicitly captures the information set to each player at the time of his decision. After completing the Prisoner's Dilemma Game at stage-one, players receive feedback on the outcomes of the game they played. Thus, at stage-two, the information sets of **P1** are singletons; and the information sets of **P2** do not include nodes corresponding to branches coming from two or more different information sets of **P1**.

A special case of multistage games is *repeated games*, where the same game is played at every stage [95]. For instance, instead of playing the Revenge Game game at the second stage, the same two players in Section 2.2.1 keep playing the Prisoner's Dilemma Game. Repeated games capture the idea of continuous interaction between parties in a rarely changing environment. This game model can be applied to many realistic settings. For example, bidders compete over a few rounds in an auction, firms fight over time in the same market, etc.

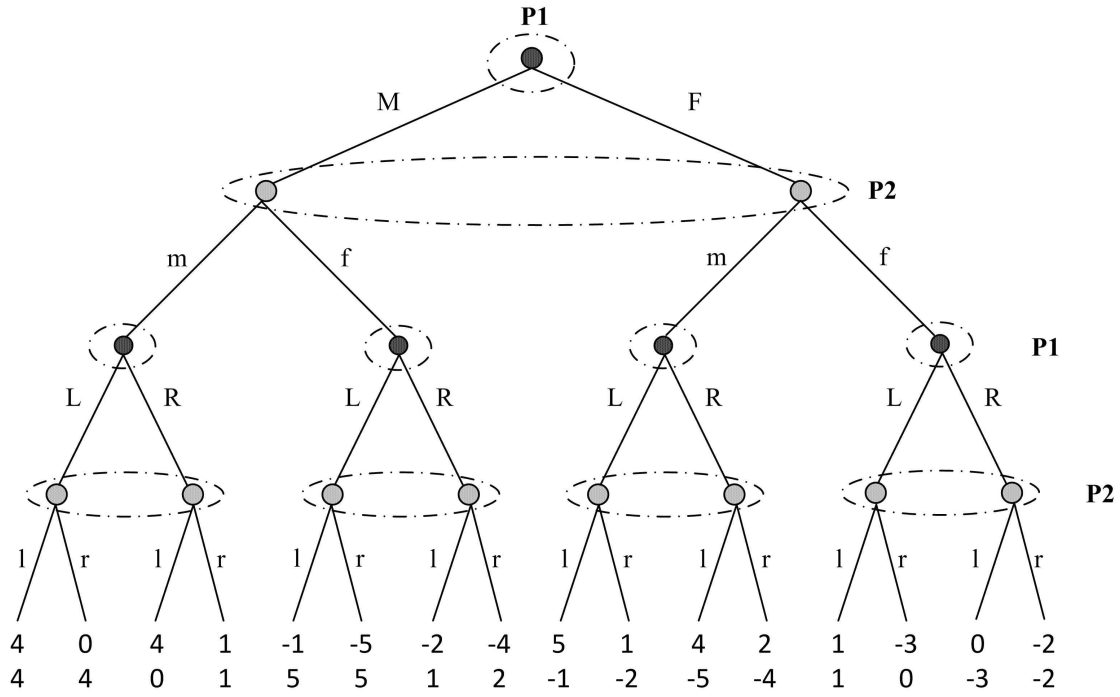

Figure 2.2: Prisoner-Revenge Game in extensive form.

2.2.3 Bayesian Games of Incomplete Information

In a game model, especially for a noncooperative game, it may be more reasonable to assume that players have some ideas about their opponents' characteristics, but not the complete knowledge. For instance, in a two-person game, one player may find out all the possible actions of his opponent's, but he may not be able to be exactly aware of the other's objective (or payoff functions). *Incomplete information* captures the players' uncertainty about some important characteristics of the game situation [35]. The concept of *type* is then introduced as a player's private information that is not common knowledge to others but is relevant to the player's decision making. The framework of incomplete information game suggested by Harsanyi is as follows:

- Nature[3] draws a type for each player from a well-defined probability distribution which is common knowledge.

[3]Nature can be thought as someone who is able to choose a game from a probability distribution over players' types [95]. It can also be viewed as a game with an additional player named 'nature' who has a fixed mixed strategy [8].

- Nature reveals the type to each corresponding player but not to any other players.
- The players simultaneously choose their type-contingent moves from their action set.

Bayesian games are widely used to model a game of incomplete information. In economics and finance, a player's payoff not only depends on his own actions and the actions of others, but also some unknown economic fundamentals. In engineering and computer science, a player may be uncertain about network parameters such as number of nodes in the network, Quality-of-Service (QoS) requirement, operating bandwidth, etc. In Chapter 5, where the fusion center is under DoS attack, the system is modeled as a Bayesian game when the sensor and attacker only have partial channel information regarding the other. In this scenario, a type is the player's instantaneous channel information, as it is a private information only known to the player itself.

Chapter 3

Estimation in Wireless Sensor Networks with Security Constraints

In this chapter, we investigate the performance of distributed estimation schemes in a wireless sensor network (WSN) in the presence of an eavesdropper. The sensor(s) transmit observations to the fusion center (FC), which at the same time is overheard by the eavesdropper. Both the FC and the eavesdropper reconstruct a minimum mean square error (MMSE) estimate of the physical quantity observed. We address the problem of transmit power allocation for system performance optimization subject to a total average power constraint on the sensor(s), and a security/secrecy constraint on the eavesdropper for two scenarios : 1) a single sensor with multiple transmit antennas and 2) multiple sensors with each sensor having a single transmit antenna. Asymptotic expressions are derived for the long-term distortion at the FC as the number of sensors or the number of antennas becomes large.

3.1 Introduction

WIRELESS Sensor Networks (WSNs) are networks consisting of some small, inexpensive, and low-power sensors, which are deployed over a region and may communicate with a remote processor over wireless links. Due to their low cost, robustness, and high flexibility, WSNs are widely employed in many military and civilian applications such as environmental monitoring, traffic control, battlefield surveillance etc. [2]. In distributed estimation, sensors independently collect data about some phenomenons and send the measurements to a fusion center (FC) which then attempts to reconstruct the phenomenon.

One crucial issue in WSNs is the limited battery life of the sensors. As sensors are normally geographically widespread, replacing batteries can be costly. The problem of

power allocation for distributed estimation has been well-studied. In [14, 106], a digital approach was considered where the analog observations are digitised into bits and then modulated and transmitted. In [15], the authors considered the problems of minimizing transmit power under distortion (or mean squared error) constraints and minimizing distortion under power constraints for an orthogonal multiple access channel (MAC). Employing a universal decentralized quantization/estimation scheme and an uncoded quadrature amplitude modulated transmission strategy, the authors in [104] studied the optimal power scheduling problem in an inhomogeneous sensor network; while the power allocation policies for a vector source were investigated in [105]. In [6], the authors investigated the energy-efficient distributed estimation problem for spatially correlated observations in WSNs. The diversity order of decentralized estimation in terms of increasing numbers of sensors has also been explored in [15, 56].

Due to the broadcast nature of wireless communications, security and privacy issues have become one of the biggest challenges in WSNs. The traditional encryption schemes or cryptography might be vulnerable because of problems such as secret key distribution and management. In addition, if an eavesdropper has sufficiently large computational power, cryptographic schemes with small key size may provide little secrecy. As an alternative, the notion of perfect secrecy[1], introduced by Shannon [90], provides a different perspective on the data confidentiality. Later, in 1970s, Wyner introduced the concept of wiretap channel [103], and showed that if the adversary's channel is a degraded version of the legitimate receiver's, reliable information can be received at the legitimate receiver without the eavesdropper being able to extract almost any useful information. From an information theoretic perspective, the authors in [30, 51, 60] studied the secrecy capacity in the case of full channel state information (CSI) or partial CSI, and investigated MIMO channels in [10, 50, 62]. Multiterminal source coding or CEO problems with secrecy constraints were also considered in [5, 17, 71, 96]. In particular, in [96], the authors investigated secure lossy source coding in the presence of an eavesdropper who is

[1]Perfect secrecy was first introduced in 1949 by Shannon. In this model it is assumed that the confidential message W is encrypted and then transmitted over a noiseless channel [90]. In information theory, perfect secrecy requires that $I(W;Z) = 0$; it indicates that the signal Z received by the eavesdropper does not provide any additional information about the transmitted message W. A weaker definition was given in [103], which requires the mutual information rate $\frac{1}{n} I(W;Z)$ goes to zero, as n, the number of bits in Z goes to infinity.

able to observe the coded information bits and has access to correlated side information. Under these assumptions, the authors derived inner and outer bounds on the achievable rate region. The authors in [48] considered a different scenario where the eavesdropper can obtain the size of the packets, thus parsing the bit stream into separate encrypted messages. Bounds on coding rate and key rate are derived for perfect zero-delay secrecy. However, although such secure source coding techniques enable one to gain information-theoretic insights, it does not provide a closed form expression for distortion achievable via multi-sensor estimation over fading channels. Thus motivated, we investigate the secure estimation problem from a signal processing viewpoint where sensors employ simple uncoded analog-forwarding techniques [23] to transmit their observations to the FC. In this way, a direct expression for the distortion over fading channels can be obtained, which is more desirable for deriving analytical results. In fact, various secrecy schemes from a 'signal processing' rather than information theoretic point of view have also been studied in [52,61,67,93], where different performance metrics, such as bit-error-rate, signal-interference-to-noise ratio, Ali-Silvey distances or error probability were used to measure secrecy in a system. Related techniques based on cooperating relays, artificial noise generation or beamforming were also implemented in [16,26,27,61] to secure a system.

Moreover, it is known that the mutual information between the input and the output of a channel is at the core of information theory; given an input signal it measures the amount of coded information that can be reliably transmitted through a channel. Its counterpart, minimum mean-square error (MMSE), is a fundamental quantity in estimation theory, which indicates how accurately the input signal can be retrieved from the channel output. In [32], the authors discovered that regardless of the input distribution the derivative of the mutual information in nats w.r.t. SNR is equal to half the MMSE, as long as the input signals are observed through an additive Gaussian noise channel. In [71], the authors related the equivocation rate to the normalized distortion at the eavesdropper in the CEO problem with additional secrecy constraints, where they showed that the estimation error at the eavesdropper is an upper bound of the equivocation rate.

Therefore, in favour of a closed form distortion expression for multi-sensor estimation

over fading channels and close relationship between MMSE and mutual information, we consider analog uncoded transmission at the sensors and introduce the MMSE as security metric to secure the system at the physical layer.

In this chapter, we consider the estimation of a single point Gaussian source by a sensor network in the presence of an eavesdropper, where the analog amplify and forward technique over a slow-fading orthogonal MAC[2] is used. We assume the same observed signal passes through another orthogonal MAC before reaching the eavesdropper, and both the FC and the eavesdropper attempt to reconstruct a MMSE estimation of the observations. The main contributions of the chapter are:

- We consider power allocation problems that minimize the distortion at the FC subject to a total transmit power constraint at the sensor(s) and a security/secrecy constraint at the eavesdropper.

- In the multiple-antenna single sensor system, we can achieve zero information leakage in the full CSI case by transmitting the signal onto the eavesdropper's channel null space, and also enhance the system performance dramatically by employing the technique of artificial noise for the partial CSI case. We give theoretical analysis on the long-term distortion for a power allocation scheme where a beamforming vector is aligned with the FC's channel direction. We also study the asymptotic distortion at the FC under the secrecy constraints when the number of antennas grows large.

- In the multiple-sensor scenario, we consider a short-term power allocation problem in the full CSI case, and long-term power allocation problems in both the full CSI and partial CSI cases. The asymptotic behaviour of the long-term distortion at the FC is also studied under the equal power allocation scheme as the number of sensors increases.

This chapter is organized as follows. In Section 3.2 we give the general problem formulation of the decentralized estimation for a system with a multiple-antenna single sen-

[2]When orthogonal MAC, such as TDMA and FDMA, is employed, only pairwise synchronization between each sensor and the FC is sufficient; whereas in the case of coherent MAC, synchronization between all sensors and the FC are required [15].

sor, and study the optimal power scheduling. We also explore other techniques that can be utilized in the multiple-antenna scenario. In Section 3.3, we explore a multiple-sensor single-antenna network and solve the power allocation problems for different scenarios. In Section 3.4, we consider a multiple-sensor-multiple-antenna network. Simulation results are given in Section 3.5, followed by concluding remarks in Section 3.6.

3.2 Multiple Antennas Scenario

Consider a wireless network with one sensor equipped with N_t transmit antennas observing a single point independent and identically distributed (i.i.d.) Gaussian source, denoted by $\theta[t]$, $t = 0, 1, 2, \ldots$, which has zero mean and variance σ_θ^2. The measurement received by the sensor at time t is given as,

$$x[t] = \theta[t] + \omega[t], \tag{3.1}$$

where we assume $\omega[t]$ is i.i.d. Gaussian noise over time t, with zero mean and variance σ_ω^2.

The analog amplify and forward techniques [24, 25] are employed, where the sensor transmits over fading channels a scaled version of the analog measurements to the fusion center (FC). It has been shown in [7, 25] that this technique is asymptotically optimal, and exactly optimal in [24] under certain situations for Gaussian source estimation in the coherent MAC. In our model, the sensor amplifies the signal with a beamforming vector $\boldsymbol{\beta}[t] \in \mathbb{C}^{N_t \times 1}$ before transmitting it to the FC in the presence of an eavesdropper, as illustrated in Figure 3.1. We assume both channels experience block fading, where the channels remain constant during each coherence time interval, and are i.i.d. over different time intervals [11]. We further assume that full channel state information (CSI) of the FC is available, while the eavesdropper's CSI may or may not be available to the FC. The FC designs the optimal power allocation strategy based on the available CSI, and then sends $\boldsymbol{\beta}[t]$ back to the sensor via a public feedback link[3]. Note that CSI at the FC can

[3]In this case, it can be seen later in (3.3b) that the minimum distortion level at the eavesdropper is achieved by implementing the linear MMSE estimator. When the feedback link is secure, the estimation

be obtained by employing pilot training signals transmitted from the sensor.

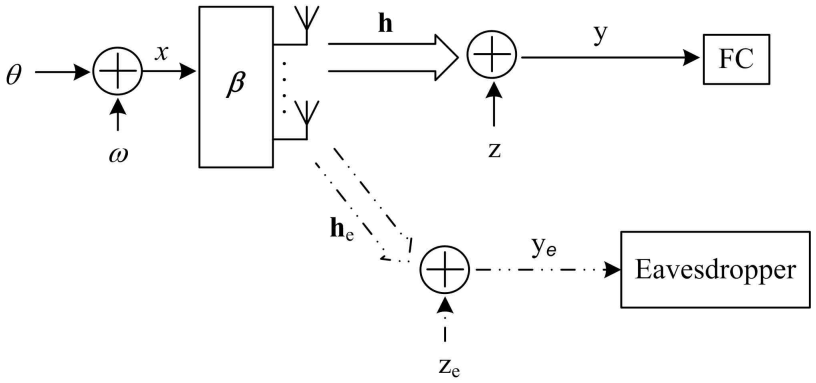

Figure 3.1: Diagram of a multiple-antenna single sensor system with the presence of an eavesdropper.

The signals received by the FC and the eavesdropper are given by, respectively

$$y[t] = \mathbf{h}[t]\boldsymbol{\beta}[t]\theta[t] + \mathbf{h}[t]\boldsymbol{\beta}[t]\omega[t] + z[t], \tag{3.2a}$$

$$y_e[t] = \mathbf{h}_e[t]\boldsymbol{\beta}[t]\theta[t] + \mathbf{h}_e[t]\boldsymbol{\beta}[t]\omega[t] + z_e[t], \tag{3.2b}$$

where both $z[t]$ and $z_e[t]$ are i.i.d. zero mean complex Gaussian channel noise at the FC and the eavesdropper with variance σ_n^2 and σ_e^2 respectively, and $\mathbf{h}[t] = [h_1[t], \ldots, h_n[t], \ldots, h_{N_t}[t]]$ and $\mathbf{h}_e[t] = [h_{e1}[t], \ldots, h_{en}[t], \ldots, h_{eN_t}[t]]$ are respectively the channels from the sensor to the FC and to the eavesdropper. We assume that $\{h_n[t]\}$ are i.i.d. complex Gaussian with zero mean and variance σ_h^2, and the elements in $\mathbf{h}_e[t]$ are also i.i.d. complex Gaussian, with zero mean and variance $\sigma_{h_e}^2$.

The linear minimum mean square error (MMSE) estimator is well known to be the optimal[4] estimator for θ under the model (3.2) [49]. At time t the mean squared error (MSE) or *distortion* at the FC and the eavesdropper can be shown to be, respectively[5],

$$D[t] = \sigma_\theta^2 - \frac{\sigma_\theta^4 \left(\mathbf{h}[t]\boldsymbol{\beta}[t]\right)^{\mathrm{H}} \mathbf{h}[t]\boldsymbol{\beta}[t]}{\sigma_n^2 + \left(\sigma_\theta^2 + \sigma_\omega^2\right)\left(\mathbf{h}[t]\boldsymbol{\beta}[t]\right)^{\mathrm{H}} \mathbf{h}[t]\boldsymbol{\beta}[t]}, \tag{3.3a}$$

distortion seen by the eavesdropper will be even larger than $D_e[t]$ given in (3.3b), due to the lack of $\boldsymbol{\beta}[t]$.

[4]It is also the best linear estimator for non-Gaussian distributions for the source and noise.

[5]The notation $\mathbf{x}^{\mathrm{H}}$ refers to the conjugate transpose of $\mathbf{x}$.

$$D_e[t] = \sigma_\theta^2 - \frac{\sigma_\theta^4 \left(\mathbf{h}_e[t]\boldsymbol{\beta}[t]\right)^{\mathrm{H}} \mathbf{h}_e[t]\boldsymbol{\beta}[t]}{\sigma_e^2 + \left(\sigma_\theta^2 + \sigma_\omega^2\right) \left(\mathbf{h}_e[t]\boldsymbol{\beta}[t]\right)^{\mathrm{H}} \mathbf{h}_e[t]\boldsymbol{\beta}[t]}. \tag{3.3b}$$

For a limited transmission power budget $\mathcal{P}_{\text{tot}}$, we would like to minimize the distortion at the FC by adapting the sensor's transmit power $\boldsymbol{\beta}[t]^{\mathrm{H}}\boldsymbol{\beta}[t]$, while maintaining a certain level of security of the transmission. In information theoretic security, the secrecy capacity is defined as the maximum transmission rate at which the mutual information between the confidential message and the signal received by the eavesdropper is less than a threshold [30]. Motivated by this idea, plus a close relation between MMSE and the mutual information of the channel input and output [32, 71], we consider a notion of *secrecy in estimation* from a non-information theoretic viewpoint by requiring the distortion at the eavesdropper to be greater than a threshold $\mathbb{D}_e$. In this way, some level of confidentiality can be achieved at the FC. We will refer to the minimum distortion threshold $\mathbb{D}_e$ as the *secrecy threshold* in the following.

Due to the assumption of system independence over time t, we will drop the time index t for the rest of the chapter.

3.2.1 Full CSI

In the case of full CSI, where we assume the FC can also acquire the channel information between the sensor and the eavesdropper, the power control policies can be derived such that the sensor is able to adjust the antenna transmission power depending on both the FC's and the eavesdropper's channel information. Clearly, the requirement of full CSI of the eavesdropper channels is infeasible in practice. However, the optimal distortion performance with this assumption is instructive as well as useful as a benchmark for the distortion performance with partial CSI of the eavesdropper channels, to be analysed subsequently.

Long-Term Optimal Power Allocation

In long-term power allocation, we assume that the crucial information lies in the long-term behaviour of the estimates, such as long-term trends in the physical process ob-

served, hence the FC would be more interested in the estimation over multiple fading blocks. We would like to minimize the long-term average distortion at the FC by adapting $\boldsymbol{\beta}$, where the average is across coherence time intervals, while keeping the long-term average sum of sensor transmission powers, defined as

$$\mathbb{E}\left[\boldsymbol{\beta}^{\mathrm{H}}\boldsymbol{\beta}\mathbb{E}\left[x_k^2\right]\right] = \mathbb{E}\left[\boldsymbol{\beta}^{\mathrm{H}}\boldsymbol{\beta}(\sigma_\theta^2+\sigma_{\omega k}^2)\right]$$

to be less than the power budget $\mathcal{P}_{\mathrm{tot}}$. We also seek to maintain the average distortion at the eavesdropper to be greater than the threshold $\mathbb{D}_e$, i.e. $\mathbb{E}[D_e] \geq \mathbb{D}_e$, to ensure some level of confidentiality can be achieved at the FC over the long-term.

Furthermore, an additional constraint ensuring the average estimation quality being better at the FC is considered for more meaningful solutions. Therefore, the power control problem can be formulated as

$$\begin{aligned}
&\min_{\boldsymbol{\beta}} \ \mathbb{E}\left[D\right] \\
&s.t. \ \mathbb{E}\left[\left(\sigma_\theta^2+\sigma_\omega^2\right)\boldsymbol{\beta}^{\mathrm{H}}\boldsymbol{\beta}\right] \leq \mathcal{P}_{\mathrm{tot}}, \\
&\quad \mathbb{E}\left[D_e\right] \geq \mathbb{D}_e, \\
&\quad \mathbb{E}\left[D_e\right] \geq \mathbb{E}\left[D\right].
\end{aligned} \tag{3.4}$$

Remark: A necessary existence condition for above problem is $0 \leq \mathbb{D}_e \leq \mathbb{E}\left[D_e\left(\hat{\boldsymbol{\beta}}^*\right)\right]$, where $\hat{\boldsymbol{\beta}}^*$ is the optimal solution of $\min_{\hat{\boldsymbol{\beta}}} \mathbb{E}\left[D_e\left(\hat{\boldsymbol{\beta}}\right)\right]$, $s.t.$ $\mathbb{E}\left[\left(\sigma_\theta^2+\sigma_\omega^2\right)\hat{\boldsymbol{\beta}}^{\mathrm{H}}\hat{\boldsymbol{\beta}}\right] \leq \mathcal{P}_{\mathrm{tot}}$. First of all, we know that the constraint $\mathbb{E}\left[\mathbb{D}_e\right] \geq \mathbb{E}\left[\mathbb{D}\right]$ is always feasible for the given power budget unless all the eavesdropper's channel realizations are all better than that of the FC, which has zero probability. Next, this choice of $\mathbb{D}_e$ guarantees $\mathbb{E}\left[D_e\left(\boldsymbol{\beta}\right)\right] \geq \mathbb{D}_e$ for any $\boldsymbol{\beta}$ satisfies $\mathbb{E}\left[\left(\sigma_\theta^2+\sigma_\omega^2\right)\boldsymbol{\beta}^{\mathrm{H}}\boldsymbol{\beta}\right] \leq \mathcal{P}_{\mathrm{tot}}$.

Given the distortion expressions in (3.3), we can simplify problem (3.4) and rewrite it as

$$\begin{aligned}
&\min_{\boldsymbol{\beta}} \ \mathbb{E}\left[\left(\alpha+\left(\mathbf{h}\boldsymbol{\beta}\right)^{\mathrm{H}}\mathbf{h}\boldsymbol{\beta}\right)^{-1}\right] \\
&s.t. \ \mathbb{E}\left[\boldsymbol{\beta}^{\mathrm{H}}\boldsymbol{\beta}\right] \leq \frac{\mathcal{P}_{\mathrm{tot}}}{\sigma_\theta^2+\sigma_\omega^2},
\end{aligned} \tag{3.5a}$$

$$\mathbb{E}\left[\left(\alpha_e + (\mathbf{h}_e\boldsymbol{\beta})^{\mathrm{H}}\mathbf{h}_e\boldsymbol{\beta}\right)^{-1}\right] \geq D_{\mathrm{ma\_L}}, \tag{3.5b}$$

$$\mathbb{E}\left[\frac{\alpha_e}{\alpha_e+(\mathbf{h}_e\boldsymbol{\beta})^{\mathrm{H}}\mathbf{h}_e\boldsymbol{\beta}}\right] \geq \mathbb{E}\left[\frac{\alpha}{\alpha+(\mathbf{h}\boldsymbol{\beta})^{\mathrm{H}}\mathbf{h}\boldsymbol{\beta}}\right], \tag{3.5c}$$

where $D_{\mathrm{ma\_L}} = \left(\frac{\mathbb{D}_e}{\frac{\sigma_\theta^2\sigma_\omega^2}{\sigma_\theta^2+\sigma_\omega^2}} - 1\right) \Big/ \frac{\sigma_e^2\sigma_\theta^2}{\left(\sigma_\theta^2+\sigma_\omega^2\right)\sigma_\omega^2}$, $\alpha = \frac{\sigma_n^2}{\sigma_\theta^2+\sigma_\omega^2}$ and $\alpha_e = \frac{\sigma_e^2}{\sigma_\theta^2+\sigma_\omega^2}$.

To solve problem (3.5), we apply the technique of Lagrange multipliers. The dual problem of (3.5) is defined as

$$\max_{\lambda,\,\nu,\,\tau} g(\lambda,\nu,\tau), \tag{3.6}$$

where λ, ν and τ are nonnegative Lagrange multipliers, and the dual function $g(\lambda,\nu,\tau)$ associated with problem (3.5) is

$$\begin{aligned} &g(\lambda,\nu,\tau) \\ &= \min_{\beta_n(\mathbf{h},\mathbf{h}_e),\ \forall n} \int_{\mathbf{h}}\int_{\mathbf{h}_e} l\left(\{\beta_n(\mathbf{h},\mathbf{h}_e)\},\lambda,\nu,\tau\right) f_{\mathbf{h}} f_{\mathbf{h}_e}\, d\mathbf{h}\, d\mathbf{h}_e + \nu D_{\mathrm{ma\_L}} - \lambda\frac{\mathcal{P}_{\mathrm{tot}}}{\sigma_\theta^2+\sigma_\omega^2}, \end{aligned} \tag{3.7}$$

where $[\beta_1(\mathbf{h},\mathbf{h}_e),\ldots,\beta_{N_t}(\mathbf{h},\mathbf{h}_e)]^{\mathrm{T}} = \boldsymbol{\beta}(\mathbf{h},\mathbf{h}_e)$ are complex gains allocated on each antenna, $f_{\mathbf{h}} = \prod_{n=1}^{N_t} f(h_n)$ and $f_{\mathbf{h}_e} = \prod_{n=1}^{N_t} f(h_{en})$, with $f(\cdot)$ denoting the probability density function. Also $l\left(\{\beta_n(\mathbf{h},\mathbf{h}_e)\},\lambda,\nu,\tau\right) = \frac{1+\tau\sigma_n^2/\sigma_e^2}{\alpha+(\mathbf{h}\boldsymbol{\beta})^{\mathrm{H}}\mathbf{h}\boldsymbol{\beta}} + \lambda\boldsymbol{\beta}^{\mathrm{H}}\boldsymbol{\beta} - \frac{\tau+\nu}{\alpha_e+(\mathbf{h}_e\boldsymbol{\beta})^{\mathrm{H}}\mathbf{h}_e\boldsymbol{\beta}}$.

It is not difficult to show that problem (3.5) is non-convex. We can obtain a locally optimal solution from the following necessary Karush-Kuhn-Tucker (KKT) conditions [9] from the Lagrangian formulation for the optimal point:

$$\frac{-h_n^{\mathrm{H}}(\mathbf{h}\boldsymbol{\beta})^{\mathrm{H}}}{\left[\alpha+(\mathbf{h}\boldsymbol{\beta})^{\mathrm{H}}\mathbf{h}\boldsymbol{\beta}\right]^2} + \frac{h_{e_n}^{\mathrm{H}}(\mathbf{h}_e\boldsymbol{\beta})^{\mathrm{H}}(\nu+\tau)}{\left(1+\tau\sigma_n^2/\sigma_e^2\right)\left[\alpha_e+(\mathbf{h}_e\boldsymbol{\beta})^{\mathrm{H}}\mathbf{h}_e\boldsymbol{\beta}\right]^2} + \frac{\lambda}{1+\tau\sigma_n^2/\sigma_e^2}\beta_n^{\mathrm{H}} = 0,\ \forall n \tag{3.8a}$$

$$\lambda\left(\mathbb{E}\left[\boldsymbol{\beta}^{\mathrm{H}}\boldsymbol{\beta}\right] - \frac{\mathcal{P}_{\mathrm{tot}}}{{\sigma_\theta}^2+{\sigma_\omega}^2}\right) = 0, \tag{3.8b}$$

$$\nu\left(D_{\mathrm{ma\_L}} - \mathbb{E}\left[\left(\alpha_e+(\mathbf{h}_e\boldsymbol{\beta})^{\mathrm{H}}\mathbf{h}_e\boldsymbol{\beta}\right)^{-1}\right]\right) = 0, \tag{3.8c}$$

$$\tau\left(\frac{\sigma_n^2}{\sigma_e^2}\mathbb{E}\left[\left(\alpha+(\mathbf{h}\boldsymbol{\beta})^{\mathrm{H}}\mathbf{h}\boldsymbol{\beta}\right)^{-1}\right] - \mathbb{E}\left[\left(\alpha_e+(\mathbf{h}_e\boldsymbol{\beta})^{\mathrm{H}}\mathbf{h}_e\boldsymbol{\beta}\right)^{-1}\right]\right) = 0. \tag{3.8d}$$

To be more specific, we first assign arbitrary initial values to λ, ν and τ, then iteratively apply the following **Step 1** and **Step 2** until (3.8b), (3.8c) and (3.8d) are satisfied.

Step 1: With fixed $\tau^{(i)}$, $\lambda^{(i)}$ and $\nu^{(i)}$, find the optimal solution $\boldsymbol{\beta}^*$ of the Lagrange dual function (3.7), which can be obtained by solving the equations in (3.8a).

Step 2: With the resulting allocated power, apply the subgradient method to update the Lagrange multipliers, i.e.,

$$\begin{aligned}
\lambda^{(i+1)} &= \left[\lambda^{(i)} + \epsilon\left(\mathbb{E}\left[\boldsymbol{\beta}^{*\mathrm{H}}\boldsymbol{\beta}^*\right] - \frac{\mathcal{P}_{\mathrm{tot}}}{{\sigma_\theta}^2 + {\sigma_\omega}^2}\right)\right]^+, \\
\nu^{(i+1)} &= \left[\nu^{(i)} + \kappa\left(D_{\mathrm{ma_L}} - \mathbb{E}\left[\left(\alpha_e + \left(\mathbf{h}_e\boldsymbol{\beta}^*\right)^{\mathrm{H}}\mathbf{h}_e\boldsymbol{\beta}^*\right)^{-1}\right]\right)\right]^+, \\
\tau^{(i+1)} &= \left[\tau^{(i)} + \upsilon\left(\frac{\sigma_n^2}{\sigma_e^2}\mathbb{E}\left[\left(\alpha + \left(\mathbf{h}\boldsymbol{\beta}^*\right)^{\mathrm{H}}\mathbf{h}\boldsymbol{\beta}^*\right)^{-1}\right] - \mathbb{E}\left[\left(\alpha_e + \left(\mathbf{h}_e\boldsymbol{\beta}^*\right)^{\mathrm{H}}\mathbf{h}_e\boldsymbol{\beta}^*\right)^{-1}\right]\right)\right]^+, \quad (3.9)
\end{aligned}$$

where υ, κ and ϵ are sufficiently small step-sizes for updating τ, ν and λ respectively.

Zero Information Leakage

Other than diversity gain, another advantage with multiple transmit antennas is that we can employ techniques to hide the observation data from the eavesdropper by transmitting it onto the null space of the eavesdropper's channel. As a result, the eavesdropper is unable to detect any information about x.

Let the singular value decomposition of $\mathbf{h}_e$ be $\mathbf{h}_e = \mathbf{U}\mathbf{S}\mathbf{V}^{\mathrm{H}}$. The null space of the eavesdropper's channel can be represented by the span of the orthonormal column vectors of $\tilde{\mathbf{V}}$, where $\tilde{\mathbf{V}}$ is the last $N_t - 1$ columns of $\mathbf{V}$. Then we can express the eavesdropper's channel null space as $\tilde{\mathbf{V}}\tilde{\mathbf{V}}^{\mathrm{H}}$ [108].

Next, we define a precoding matrix

$$\mathbf{W} = \tilde{\mathbf{V}}\tilde{\mathbf{V}}^{\mathrm{H}},$$

where $\mathbf{W} \in \mathbb{C}^{N_t}$. The sensor sends $\mathbf{W}\boldsymbol{\beta}x$. The signal received by the FC and the eaves-

dropper are given by, respectively

$$y = \mathbf{h}\mathbf{W}\boldsymbol{\beta}x + z = \mathbf{h}\mathbf{W}\boldsymbol{\beta}\theta + \mathbf{h}\mathbf{W}\boldsymbol{\beta}\omega + z, \tag{3.10a}$$

$$y_e = \mathbf{h}_e\mathbf{W}\boldsymbol{\beta}x + z_e = z_e, \tag{3.10b}$$

and the transmission power can be computed as $\left((\mathbf{W}\boldsymbol{\beta})^{\mathrm{H}}\mathbf{W}\boldsymbol{\beta}\right)(\sigma_\theta^2 + \sigma_\omega^2)$. Since the eavesdropper receives only noise, the distortion at the eavesdropper reaches its highest level of σ_θ^2, and hence we can remove constraints (3.5b) and (3.5c) in problem (3.5). In addition, we know that the beamforming vector should line-up with $\mathbf{hW}$ to minimize the distortion at the FC; thus, $\boldsymbol{\beta} = \sqrt{p_0}\frac{(\mathbf{hW})^{\mathrm{H}}}{\|\mathbf{hW}\|}$ with p_0 being real-valued[6]. Therefore, problem (3.5) can be then simplified and rewritten as

$$\begin{aligned} &\min_{p_0 \geq 0} \ \mathbb{E}\left[\left(\alpha + p_0\mathbf{h}\mathbf{W}\mathbf{h}^{\mathrm{H}}\right)^{-1}\right] \\ &s.t. \ \ p_0 \leq \frac{\mathcal{P}_{\mathrm{tot}}}{\sigma_\theta^2 + \sigma_\omega^2}. \end{aligned} \tag{3.11}$$

It can be seen that when $p_0 = \frac{\mathcal{P}_{\mathrm{tot}}}{\sigma_\theta^2+\sigma_\omega^2}$ the long-term distortion at the FC reaches its minimum.

Remark: The signal is transmitted on the eavesdropper's null space via the precoding matrix $\mathbf{W}$. Therefore, we have the effective FC channel $\mathbf{hW}$, which is the projection of $\mathbf{h}$ on the null space of $\mathbf{h}_e$.

Short-Term Optimal Power Allocation

We can formulate a power allocation problem that minimizes the distortion at the FC, while satisfying a secrecy constraint at the eavesdropper and meeting the total power budget *in every transmission time slot*. We refer this as short-term power allocation. Note that for short-term optimal power allocation we cannot guarantee that the estimation quality is better at the FC at all times. Keeping this in mind, the optimal power allocation

[6]The notation $\|\mathbf{x}\|$ refers the Euclidean norm of the vector $\mathbf{x}$.

problem for a given set of channels can be cast as

$$\begin{aligned}\min_{\boldsymbol{\beta}} \; D = \min_{\boldsymbol{\beta}} \;\; & \sigma_\theta^2 - \frac{\alpha\sigma_\theta^4}{\sigma_n^2}\left[1 - \frac{\alpha}{\alpha + (\mathbf{h}\boldsymbol{\beta})^{\mathrm{H}}\mathbf{h}\boldsymbol{\beta}}\right] \\ s.t. \;\; & \boldsymbol{\beta}^{\mathrm{H}}\boldsymbol{\beta} \leq \frac{\mathcal{P}_{\mathrm{tot}}}{\sigma_\theta^2 + \sigma_\omega^2}, \\ & \sigma_\theta^2 - \frac{\alpha_e\sigma_\theta^4}{\sigma_e^2}\left[1 - \frac{\alpha_e}{\alpha_e + (\mathbf{h}_e\boldsymbol{\beta})^{\mathrm{H}}\mathbf{h}_e\boldsymbol{\beta}}\right] \geq \mathbb{D}_e, \end{aligned} \tag{3.12}$$

where $\mathbb{D}_e$ and $\mathcal{P}_{\mathrm{tot}}$ are respectively the distortion threshold at the eavesdropper and the total transmission power budget.

We aim to find an optimal beamforming vector $\boldsymbol{\beta}$ which meets the constraints in every fading block. The Lagrange multiplier technique is applied to solve this non-convex optimization problem, and the details are omitted to avoid repetition.

3.2.2 Partial CSI

Due to the difficulties of perfectly acquiring the eavesdropper's CSI in practical setups, in this subsection we assume that the FC only has statistical knowledge of the eavesdropper's channel information. As it is not practical to consider short-term constraints that need to be satisfied at every time instance, we only look at the long-term scenario for the partial CSI case. We first explore the power allocation problem that minimizes the long-term distortion at the FC via the Lagrange multiplier technique. Next, we study the technique of artificial noise, where the artificial interference is transmitted to confuse the eavesdropper. We also analyse the asymptotic behaviour of the distortion at the FC when equal power allocation is employed.

The power allocation problem is formulated similar to (3.4) but now with $\boldsymbol{\beta}$ being a function of $\mathbf{h}$, rather than a function of $\mathbf{h}$ and $\mathbf{h}_e$ as in the full CSI case. As the problem is again non-convex, a locally optimal solution can be obtained as follows. Similar to (3.7), we define the Lagrange dual function as

$$g(\lambda,\nu,\tau) = \min_{\beta_n(\mathbf{h}),\,\forall n} \int_{\mathbf{h}} l\left(\{\beta_n(\mathbf{h})\},\lambda,\nu,\tau\right) f_{\mathbf{h}}\, d\mathbf{h} + \nu D_{\mathrm{ma\_L}} - \lambda\frac{\mathcal{P}_{\mathrm{tot}}}{\sigma_\theta^2 + \sigma_\omega^2},$$

with $l\left(\{\beta_n(\mathbf{h})\}, \lambda, \nu, \tau\right)$ expressed as

$$l\left(\{\beta_n(\mathbf{h})\}, \lambda, \nu, \tau\right) = \frac{1+\tau\sigma_n^2/\sigma_e^2}{\alpha+\left(\mathbf{h}_e\boldsymbol{\beta}\right)^{\mathrm{H}}\mathbf{h}_e\boldsymbol{\beta}}+\lambda\boldsymbol{\beta}^{\mathrm{H}}\boldsymbol{\beta}-\int_{\mathbf{h}_e}\frac{\tau+\nu}{\alpha_e+\left(\mathbf{h}_e\boldsymbol{\beta}\right)^{\mathrm{H}}\mathbf{h}_e\boldsymbol{\beta}}f_{\mathbf{h}_e}d\mathbf{h}_e. \tag{3.13}$$

For any set of channels $\mathbf{h}$, the optimal transmission power of the sensor is determined by the stationary points (or KKT points). We can then adapt similar methods as described in Section 3.2.1. In **Step 1** the power policies $\boldsymbol{\beta}^*(\mathbf{h})$ can be derived by applying **Algorithm 1** below. For fixed $\tau^{(i)}$, $\lambda^{(i)}$ and $\nu^{(i)}$, **Algorithm 1** sequentially updates the transmit power on each antenna by minimizing the function given in (3.13), until a locally optimal solution is found. In **Step 2** we update the Lagrange multipliers via the subgradient method.

Algorithm 1

1: Initialize the iteration index $q=0$, choose an arbitrary initial value for $\left\{\beta_n(\mathbf{h})^{(q)}\right\}_{n=1}^{N_t}$, and obtain $l^{(q)} = l\left(\left\{\beta_n(\mathbf{h})^{(q)}\right\}, \lambda, \nu, \tau\right)$ from (3.13).

2: **repeat**

3: for $j = 1 : N_t$

1. Find the complex gain $\beta_j'(\mathbf{h})$ on antenna j such that $l\left(\left\{\left\{\beta_n(\mathbf{h})^{(q)}\right\}_{n\neq j}, \beta_j'(\mathbf{h})\right\}, \lambda, \nu, \tau\right)$ is minimized.
2. Update the transmission power of antenna j by $\left[\beta_1(\mathbf{h})^{(q)}, \ldots, \beta_j'(\mathbf{h})^{(q)}, \ldots, \beta_{N_t}(\mathbf{h})^{(q)}\right]$.

4: update $l^{(q+1)} = l\left(\left\{\beta_n'(\mathbf{h})^{(q)}\right\}, \lambda, \nu, \tau\right)$, and $q = q+1$.

5: **until** convergence: $\left(l^{(q+1)} - l^{(q)}\right) / \left(l^{(q+1)}\right) < \zeta$; set $\{\beta_n^*(\mathbf{h})\} = \left\{\beta_n'(\mathbf{h})^{(q)}\right\}$.

Remark: In **Step 1**, $\tau^{(i)}$, $\lambda^{(i)}$ and $\nu^{(i)}$ are fixed, hence we drop the iteration number i in **Algorithm 1**; and ζ is a pre-specified convergence criterion. Additionally, **Algorithm 1** only gives a locally optimal solution, as the different initial values of $\boldsymbol{\beta}(\mathbf{h})^{(0)}$ may lead l in (3.13) to converge to a different minimum. Thus, in practice, the FC begins with several different initial points, and chooses the best resulting powers and forwards them to the sensor.

Artificial Noise

To enhance the system performance, we can use the technique of artificial noise to degrade the eavesdropper's channel. The artificial noise is generated by the transmitter (the sensor) in a way that the additional noise lies in the null space of the intended receiver's (the FC's) channel; as a result, the noise would not cause any damage towards the message received by the FC but would degrade the eavesdropper's channel [27,74].

To be more specific, let the column vectors of $\hat{\mathbf{W}}^{\mathrm{H}} = [\mathbf{w}_1 \mathbf{W}_2]$ be an orthonormal basis of $\mathbb{C}^{N_t}$, with $\mathbf{w}_1^{\mathrm{T}} \in \mathbb{C}^{1\times N_t}$ representing the signal space of $\mathbf{h}$. The sensor then transmits

$$\mathbf{w}_1 \sqrt{p_s} x + \mathbf{W}_2 \mathbf{v}, \tag{3.14}$$

where $\mathbf{W}_2\mathbf{v}$ is the artificial noise, which is chosen to be a random vector in the null space of $\mathbf{h}$ to reduce the possibility of small 'noise' seen by the eavesdropper. Here $\mathbf{v} \in \mathbb{C}^{(N_t-1)\times 1}$ has $N_t - 1$ i.i.d. complex Gaussian entries with each having zero mean and variance p_a. Hence the transmit power in each fading block is given as $p_s\left(\sigma_\theta^2 + \sigma_\omega^2\right) + p_a\left(N_t - 1\right)$. The signal received by the FC and the eavesdropper are respectively

$$\begin{aligned} y &= \mathbf{h}\hat{\mathbf{W}}^{\mathrm{H}} \left[\sqrt{p_s}x,\ \mathbf{v}^{\mathrm{T}}\right]^{\mathrm{T}} + z = \mathbf{h}\mathbf{w}_1\sqrt{p_s}x + \mathbf{h}\mathbf{W}_2\mathbf{v} + z \\ &= \mathbf{h}\mathbf{w}_1\sqrt{p_s}x + z, \\ y_e &= \mathbf{h}_e\hat{\mathbf{W}}^{\mathrm{H}} \left[\sqrt{p_s}x, \mathbf{v}^{\mathrm{T}}\right]^{\mathrm{T}} + z_e = \mathbf{h}_e\mathbf{w}_1\sqrt{p_s}x + \mathbf{h}_e\mathbf{W}_2\mathbf{v} + z_e. \end{aligned}$$

Remark: As $\mathbf{h}_e$ has i.i.d. entries and $\hat{\mathbf{W}}$ is a unitary matrix, we know that $\mathbf{h}_e\hat{\mathbf{W}}^{\mathrm{H}}$ also has i.i.d. elements. This indicates that $\mathbf{h}_e\mathbf{w}_1$ is independent of $\mathbf{h}_e\mathbf{W}_2$. As a result, the effective noise at the eavesdropper becomes $\mathbf{h}_e\mathbf{W}_2\mathbf{v} + z_e$.

Our objective is to derive the power used to produce artificial noise and to forward the observation signal so that the long-term distortion at the FC is minimized, while satisfying the three long-term constraints as described in Section 3.2.1. Assuming both the FC and the eavesdropper use the optimal MMSE estimator, the functional optimization

problem can be written as (3.15) below.

$$\begin{aligned}
\min_{p_s(\mathbf{h}),\, p_a(\mathbf{h})} \quad & \mathbb{E}\left[\left(\frac{1}{\sigma_\theta^2}+\frac{p_s(\mathbf{h})\mathbf{h}\mathbf{h}^{\mathrm{H}}}{\sigma_n^2+p_s(\mathbf{h})\mathbf{h}\mathbf{h}^{\mathrm{H}}\sigma_\omega^2}\right)^{-1}\right] \\
s.t. \quad & (\sigma_\theta^2+\sigma_\omega^2)\mathbb{E}\left[p_s(\mathbf{h})\right]+(N_t-1)\,\mathbb{E}\left[p_a(\mathbf{h})\right] \leq \mathcal{P}_{\mathrm{tot}} \\
& \mathbb{E}\left[\left(\frac{1}{\sigma_\theta^2}+\frac{p_s(\mathbf{h})|\mathbf{h}_e\mathbf{w}_1|^2}{\sigma_e^2+p_s(\mathbf{h})|\mathbf{h}_e\mathbf{w}_1|^2\sigma_\omega^2+\mathbf{h}_e\mathbf{W}_2\mathbf{W}_2^{\mathrm{H}}\mathbf{h}_e^{\mathrm{H}}p_a(\mathbf{h})}\right)^{-1}\right] \geq \mathbb{D}_e \\
& \mathbb{E}\left[\left(\frac{1}{\sigma_\theta^2}+\frac{p_s(\mathbf{h})|\mathbf{h}_e\mathbf{w}_1|^2}{\sigma_e^2+p_s(\mathbf{h})|\mathbf{h}_e\mathbf{w}_1|^2\sigma_\omega^2+\mathbf{h}_e\mathbf{W}_2\mathbf{W}_2^{\mathrm{H}}\mathbf{h}_e^{\mathrm{H}}p_a(\mathbf{h})}\right)^{-1}\right] \geq \\
& \qquad\qquad \mathbb{E}\left[\left(\frac{1}{\sigma_\theta^2}+\frac{p_s(\mathbf{h})\mathbf{h}\mathbf{h}^{\mathrm{H}}}{\sigma_n^2+p_s(\mathbf{h})\mathbf{h}\mathbf{h}^{\mathrm{H}}\sigma_\omega^2}\right)^{-1}\right]
\end{aligned} \tag{3.15}$$

To solve problem (3.15), we apply the technique of Lagrange multipliers and use two steps similar to those described in Section 3.2.1, where in **Step 1**, with fixed Lagrange multipliers, we need to sequentially find the $p_s(\mathbf{h})$ and $p_a(\mathbf{h})$.

Asymptotic Analysis

In this subsection, we are interested in seeing how the long-term distortion decays at the FC as the number of antennas N_t increases, under both the power constraint and the secrecy constraints.

Notation: For two functions $f_1(\cdot)$ and $f_2(\cdot)$, we use the standard asymptotic notation and say that $f_1 \sim f_2$ as $t \to t_0$, if $f_1(t)/f_2(t) \to 1$ as $t \to t_0$ [101].

We consider the case where the beamforming vector is chosen to be lined up with the FC's channel in order to minimize the distortion at the FC, i.e.,

$$\boldsymbol{\beta} = \frac{\sqrt{p_0}\mathbf{h}^{\mathrm{H}}}{\|\mathbf{h}\|}, \tag{3.16}$$

where $p_0 = \min\left[\frac{1-\alpha_e D_{\mathrm{ma\_L}}}{\sigma_{h_e}^2 D_{\mathrm{ma\_L}}},\ \frac{\mathcal{P}_{\mathrm{tot}}}{\sigma_\theta^2+\sigma_\omega^2}\right]$. This choice of p_0 guarantees that the three long-term constraints are satisfied. To see this, we first rewrite $\mathbb{E}\left[\left(\alpha_e+\left(\mathbf{h}_e\boldsymbol{\beta}\right)^{\mathrm{H}}\mathbf{h}_e\boldsymbol{\beta}\right)^{-1}\right]$ as

$$\mathbb{E}\left[\left(\alpha_e+\left(\mathbf{h}_e\boldsymbol{\beta}\right)^{\mathrm{H}}\mathbf{h}_e\boldsymbol{\beta}\right)^{-1}\right]$$

$$
\begin{aligned}
&= \mathbb{E}\left[\left(\alpha_e + p_0 \frac{\mathbf{h}\mathbf{h}_e^{\mathrm{H}}\mathbf{h}_e\mathbf{h}^{\mathrm{H}}}{\|\mathbf{h}\|^2}\right)^{-1}\right] \\
&= \mathbb{E}\left[\left(\alpha_e + p_0 \left(\|\mathbf{h}_e\| \left|\tilde{\mathbf{h}}\tilde{\mathbf{h}}_e^{\mathrm{H}}\right|\right)^2\right)^{-1}\right], \qquad (3.17)
\end{aligned}
$$

where $\tilde{\mathbf{h}} = \frac{\mathbf{h}}{\|\mathbf{h}\|}$ and $\tilde{\mathbf{h}}_e = \frac{\mathbf{h}_e}{\|\mathbf{h}_e\|}$, which are two independent isotropic vectors on the N_t-dimensional unit sphere.

The first thing to be noticed from (3.17) is that $\|\mathbf{h}_e\| \cdot \left|\tilde{\mathbf{h}}\tilde{\mathbf{h}}_e^{\mathrm{H}}\right|$ can be thought of as the magnitude of the vector $\mathbf{h}_e$ projected onto the vector space of $\mathbf{h}$, as the second term can be written as $\left|\tilde{\mathbf{h}}\tilde{\mathbf{h}}_e^{\mathrm{H}}\right| = \left|\cos\left(\angle\left(\tilde{\mathbf{h}}, \tilde{\mathbf{h}}_e\right)\right)\right|$[7]. This also indicates that $\left|\tilde{\mathbf{h}}\tilde{\mathbf{h}}_e^{\mathrm{H}}\right|$ is only related to the difference in the two channel directions. Therefore, by exploiting the independence of channel norm and channel direction [45], we can simplify (3.17) as

$$
\mathbb{E}\left[\left(\alpha_e + p_0 XY\right)^{-1}\right], \qquad (3.18)
$$

where $X = \|\mathbf{h}_e\|^2$ and $Y = \left|\tilde{\mathbf{h}}\tilde{\mathbf{h}}_e^{\mathrm{H}}\right|^2$ are two independent random variables, with X being Gamma distributed as

$$
X \sim \Gamma\left(N_t, \sigma_{h_e}^2\right), \qquad f_X(x) = \frac{x^{N_t-1} e^{-x/\sigma_{h_e}^2}}{\sigma_{h_e}^{2N_t}\,(N_t-1)!}.
$$

In addition, as $\tilde{\mathbf{h}}$ and $\tilde{\mathbf{h}}_e$ are independent isotropic vectors, we have Y being Beta distributed with parameters 1 and $N_t - 1$:

$$
Y \sim Beta\left(1, N_t - 1\right), \qquad f_Y(y) = (N_t - 1)\,(1-y)^{N_t-2}.
$$

Since (3.18) is convex with respect to XY, applying Jensen's inequality and using the fact that $\mathbb{E}[X] = \sigma_{h_e}^2 N_t$, $\mathbb{E}[Y] = \frac{1}{N_t}$, we obtain a lower bound of (3.18) as:

$$
\begin{aligned}
&\mathbb{E}\left[\left(\alpha_e + p_0 XY\right)^{-1}\right] \\
&\geq \left(\alpha_e + p_0 \mathbb{E}[X]\,\mathbb{E}[Y]\right)^{-1}
\end{aligned}
$$

[7] $\angle(\mathbf{x}, \mathbf{y})$ is the angle between two vectors $\mathbf{x}$ and $\mathbf{y}$, and $|\cos(\angle(\mathbf{x}, \mathbf{y}))| = \frac{|\mathbf{x}^{\mathrm{H}}\mathbf{y}|}{\|\mathbf{x}\| \cdot \|\mathbf{y}\|}$.

$$= \left(\alpha_e + p_0 \sigma_{h_e}^2\right)^{-1}, \tag{3.19}$$

from which we can obtain a lower bound of the long-term distortion at the eavesdropper given as

$$\mathbb{E}\left[D_e\right] \geq \frac{\sigma_\theta^2 \sigma_\omega^2}{\sigma_\theta^2 + \sigma_\omega^2} + \frac{\sigma_\theta^4 / \left(\sigma_\theta^2 + \sigma_\omega^2\right)}{1 + \sigma_{h_e}^2 p_0 \, / \alpha_e}, \tag{3.20}$$

which is independent of the number of transmit antennas, and decreases to $\frac{\sigma_\theta^2 \sigma_\omega^2}{\sigma_\theta^2 + \sigma_\omega^2}$ when the total transmit power is increased to infinity (as the long-term transmit power is $\left(\sigma_\theta^2 + \sigma_\omega^2\right) \mathbb{E}\left(\boldsymbol{\beta}^{\mathrm{H}} \boldsymbol{\beta}\right) = \left(\sigma_\theta^2 + \sigma_\omega^2\right) p_0$). Hence, we can set $\frac{\sigma_\theta^2 \sigma_\omega^2}{\sigma_\theta^2 + \sigma_\omega^2} + \frac{\sigma_\theta^4 / \left(\sigma_\theta^2 + \sigma_\omega^2\right)}{1 + \sigma_{h_e}^2 p_0 / \alpha_e} \geq \mathbb{D}_e$ to guarantee that the secrecy constraint at the eavesdropper is satisfied, i.e., $p_0 \leq \frac{1 - \alpha_e D_{\mathrm{ma\_L}}}{\sigma_{h_e}^2 D_{\mathrm{ma\_L}}}$. Therefore, given a total transmit power budget $\mathcal{P}_{\mathrm{tot}}$, we see that the long-term power constraint as well as the secrecy constraint are met when $p_0 = \min\left[\frac{1 - \alpha_e D_{\mathrm{ma\_L}}}{\sigma_{h_e}^2 D_{\mathrm{ma\_L}}}, \ \frac{\mathcal{P}_{\mathrm{tot}}}{\sigma_\theta^2 + \sigma_\omega^2}\right]$.

Furthermore, using the beamforming vector (3.16) gives us the long-term distortion at the FC:

$$\begin{aligned}
\mathbb{E}\left[D\right] =& \frac{\sigma_\theta^2 \sigma_\omega^2}{\sigma_\theta^2 + \sigma_\omega^2} + \frac{\sigma_\theta^4 \alpha^2}{\sigma_n^2} \mathbb{E}\left[\left(\alpha + p_0 \|\mathbf{h}\|^2\right)^{-1}\right] && \text{(3.21a)} \\
=& \frac{\sigma_\theta^2 \sigma_\omega^2}{\sigma_\theta^2 + \sigma_\omega^2} + \frac{\sigma_\theta^4 \alpha^2}{\sigma_n^2} \mathbb{E}\left[\left(\alpha + p_0 \sum_{n=1}^{N_t} |h_n|^2\right)^{-1}\right] && \text{(3.21b)} \\
\overset{(a)}{\sim}& \frac{\sigma_\theta^2 \sigma_\omega^2}{\sigma_\theta^2 + \sigma_\omega^2} + \frac{\sigma_\theta^4 \alpha^2}{\sigma_n^2} \left(\alpha + p_0 N_t \sigma_h^2\right)^{-1} && \text{(3.21c)} \\
\sim& \frac{\sigma_\theta^2 \sigma_\omega^2}{\sigma_\theta^2 + \sigma_\omega^2} + \frac{\sigma_\theta^4 \alpha^2}{\sigma_n^2 \sigma_h^2 p_0} \frac{1}{N_t}, && \text{(3.21d)}
\end{aligned}$$

which is asymptotically equal to the constant $\frac{\sigma_\theta^2 \sigma_\omega^2}{\sigma_\theta^2 + \sigma_\omega^2}$ plus a term that decays to zero at the rate $1/N_t$, where (a) holds providing the expectation $\mathbb{E}\left[|h_1|^2\right]$ exists (which in this case is σ_h^2) and applying the strong law of large numbers.

From (3.21d), we notice that if the beamforming vector has the form $\boldsymbol{\beta} = \frac{\sqrt{p_0} \mathbf{h}^{\mathrm{H}}}{\|\mathbf{h}\|}$, the long-term distortion at the FC decreases as we increase N_t, whereas the lower bound of the distortion at the eavesdropper, as shown in (3.20), is dependent on the transmission power. Therefore, we conclude that, given a limited transmit power budget, the

long-term distortion at the FC is always smaller than the distortion at the eavesdropper when the number of transmission antennas is large; in other words, all three long-term constraints can be satisfied when $\boldsymbol{\beta} = \frac{\sqrt{p_0}\mathbf{h}^{\mathrm{H}}}{\|\mathbf{h}\|}$ with $p_0 = \min\left[\frac{1-\alpha_e D_{\mathrm{ma\_L}}}{\sigma_{h_e}^2 D_{\mathrm{ma\_L}}},\ \frac{\mathcal{P}_{\mathrm{tot}}}{\sigma_\theta^2+\sigma_\omega^2}\right]$.

3.3 Multiple Sensors Scenario

For the single point source estimation, if applying multiple antennas is not an option, an alternative way to improve the estimation accuracy at the FC is to employ multiple sensors. Therefore, in this section, we investigate the behaviour of a multiple-sensor signal antenna system, followed by multi-antenna multi-sensor systems in Section 3.4. In both cases we assume that the FC and eavesdropper have a single receive antenna.

A schematic diagram of the wireless system model is shown in Figure 3.2. We assume that the same single point Gaussian source θ as defined in Section 3.2 is observed by K sensors. The measurement received by the kth sensor is corrupted with noise ω_k and given as,

$$x_k = \theta + \omega_k, \tag{3.22}$$

where we assume ω_k is i.i.d. Gaussian noise over time, with zero mean and variance $\sigma_{\omega k}^2$. We assume pairwise synchronization between each sensor and the FC. The sensors employ the analog amplify and forward technique [24,25] to scale the signal with $\beta_k \in \mathbb{C}$ before sending it to the FC via a set of orthogonal channels $[h_1, \ldots, h_K]$. The observation $\{x_k\}$ is also listened to by the eavesdropper via another set of orthogonal channels $[h_{e1}, \ldots, h_{eK}]$.

The signals received by the FC and the eavesdropper from the kth sensor are given by, respectively,

$$y_k = h_k\beta_k\theta + h_k\beta_k\omega_k + z_k, \tag{3.23a}$$

$$y_{e_k} = h_{ek}\beta_k\theta + h_{ek}\beta_k\omega_k + z_{ek}, \tag{3.23b}$$

where both h_k and h_{ek} are zero mean i.i.d. complex Gaussian channels (Rayleigh fading)

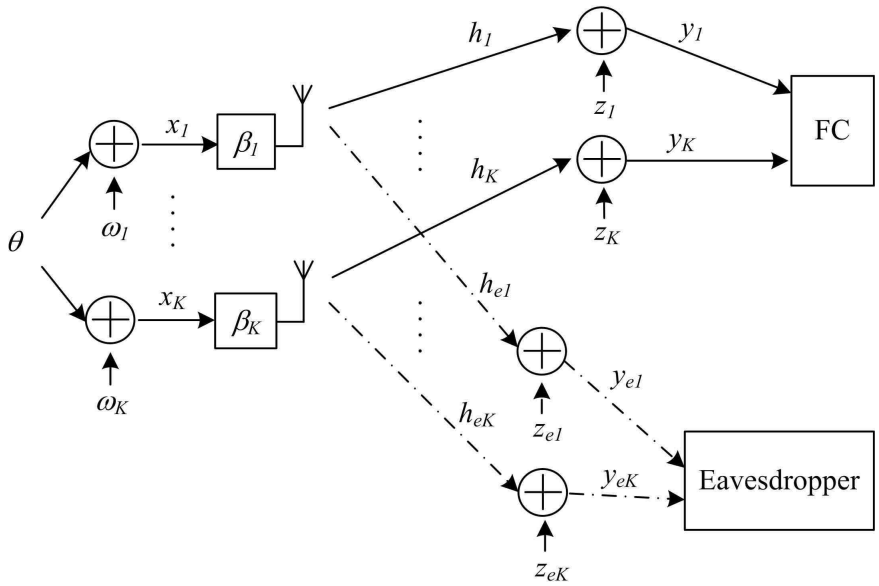

Figure 3.2: Diagram of a wireless sensor network using orthogonal MAC scheme with the presence of an eavesdropper.

from sensor k to the FC and the eavesdropper with variances $\sigma^2_{h_k}$ and $\sigma^2_{h_e k}$ respectively, and z_k and z_{ek} represent i.i.d. complex Gaussian noise with zero mean and variances σ^2_{nk} at the FC and $\sigma^2_{e\,k}$ at the eavesdropper respectively.

The optimal MMSE estimator is used at both the FC and the eavesdropper to measure θ. At each channel instance, the mean squared error or *distortion* at the FC and the eavesdropper can be shown to be, respectively,

$$\begin{aligned} D &= \left(\frac{1}{\sigma_\theta^2} + \sum_{k=1}^{K} \frac{(h_k\beta_k)^{\mathrm{H}} h_k\beta_k}{(h_k\beta_k)^{\mathrm{H}} h_k\beta_k\sigma^2_{\omega k} + \sigma^2_{nk}} \right)^{-1} \\ &= \left(\frac{1}{\sigma_\theta^2} + \sum_{k=1}^{K} \frac{g_k p_k}{g_k p_k \sigma^2_{\omega k} + \sigma^2_{nk}} \right)^{-1}, \end{aligned} \tag{3.24a}$$

$$\begin{aligned} D_e &= \left(\frac{1}{\sigma_\theta^2} + \sum_{k=1}^{K} \frac{(h_{ek}\beta_k)^{\mathrm{H}} h_{ek}\beta_k}{(h_{ek}\beta_k)^{\mathrm{H}} h_{ek}\beta_k\sigma^2_{\omega k} + \sigma^2_{e\,k}} \right)^{-1} \\ &= \left(\frac{1}{\sigma_\theta^2} + \sum_{k=1}^{K} \frac{g_{ek} p_k}{g_{ek} p_k \sigma^2_{\omega k} + \sigma^2_{e\,k}} \right)^{-1}, \end{aligned} \tag{3.24b}$$

where $g_k = h_k^{\mathrm{H}} h_k \in \mathbb{R}$ and $g_{ek} = h_{ek}^{\mathrm{H}} h_{ek} \in \mathbb{R}$ are respectively the channel power gains from sensor k to the FC and the eavesdropper, and $p_k = \beta_k^{\mathrm{H}}\beta_k \in \mathbb{R}$ is the power allocated

on the kth sensor. This means that for a given set of $\{p_k\}$, any $\{\beta_k\}$ satisfying $\beta_k^{\mathrm{H}}\beta_k = p_k, \forall k$ would result in the same distortion, which implies β_k does not necessarily need to line-up with sensor k's channel direction; hence we mainly focus on $\{p_k\}$ in multiple-sensor scenarios.

In the following, we first look at the optimal power allocation, where the optimal power policies are designed by the FC based on the available CSI, and then sends $\{p_k\}$ back to the sensors via a public channel. Applying a similar idea as in Section 3.2.2 of increasing the interference seen by the adversary in such a way that the channel is degraded while the channel of the legitimate is not, we then consider a scenario where some of the sensors are employed to broadcast artificial interference which can be canceled off at the FC, but will in general degrade the eavesdropper's channel. The asymptotic behaviour is also studied for the partial CSI case at the end of this section.

3.3.1 Full CSI - Optimal Power Allocation

In order to extend sensors' lifespan meanwhile maintaining a certain level of security for the network, we would like to minimize the distortion at the FC by adapting the sensors' transmit powers while satisfying the same three constraints as considered in multiple-antenna scenarios. With full knowledge of the eavesdropper's channel information, the power control problem can be formulated as:

$$\begin{aligned}
\min_{p_k(g_k,g_{e_k})\geq 0,\ \forall k}\ & \mathbb{E}\left[\left(\frac{1}{\sigma_\theta^2}+\sum_{k=1}^{K}\frac{g_k p_k}{g_k p_k \sigma_{\omega k}^2+\sigma_{nk}^2}\right)^{-1}\right] \\
s.t.\ & \mathbb{E}\left[\sum_{k=1}^{K}\left(\sigma_\theta^2+\sigma_{\omega k}^2\right)p_k\right] \leq \mathcal{P}_{\mathrm{tot}} \\
& \mathbb{E}\left[\left(\frac{1}{\sigma_\theta^2}+\sum_{k=1}^{K}\frac{g_{e_k} p_k}{g_{e_k} p_k \sigma_{\omega k}^2+\sigma_{e k}^2}\right)^{-1}\right] \geq \mathbb{D}_e \\
& \mathbb{E}\left[\left(\frac{1}{\sigma_\theta^2}+\sum_{k=1}^{K}\frac{g_{e_k} p_k}{g_{e_k} p_k \sigma_{\omega k}^2+\sigma_{e k}^2}\right)^{-1}\right] \geq \mathbb{E}\left[\left(\frac{1}{\sigma_\theta^2}+\sum_{k=1}^{K}\frac{g_k p_k}{g_k p_k \sigma_{\omega k}^2+\sigma_{nk}^2}\right)^{-1}\right].
\end{aligned} \tag{3.25}$$

Similar setups have been considered in [33], where a minimum distortion threshold is set at the eavesdropper to ensure that the estimation error at the eavesdropper is no smaller than the requirement. In (3.25), an additional constraint guaranteeing a larger error always occurs at the eavesdropper is considered, hence one would expect no better performance being achieved at the FC compared with the results in [33] because of a smaller feasible region. Despite this, one can use the same methods as described in Section 3.2.1 by applying KKT condition and then numerically obtain locally optimal solutions for problem (3.25). Simulation results are given in Section 3.5.

Next, we explore the short-term distortion performance at the FC while satisfying a secrecy constraint at the eavesdropper and a total power constraint at the sensors *in every transmission instant*. As for the short-term optimal power allocation we cannot guarantee the distortion to be smaller at the legitimate receiver than the distortion at the eavesdropper for every fading block. For example, if the instantaneous channel SNR of the eavesdropper is greater than the channel SNR of the FC ($\frac{g_{e_k}}{\sigma_{\tilde{e}k}^2} > \frac{g_k}{\sigma_{nk}^2}$, $\forall k$), all sensors will have to stop transmitting, which is not so interesting. Therefore, the power allocation problem in short-term scenario is only considered in the case of full CSI. We can formulate the optimization problem and rewrite it as

$$\begin{aligned} \min_{p_k \geq 0,\, \forall k} \quad & \sum_{k=1}^{K} \frac{-g_k p_k}{g_k p_k \sigma_{\omega k}^2 + \sigma_{nk}^2} \\ s.t. \quad & \sum_{k=1}^{K} \left(\sigma_\theta^2 + \sigma_{\omega k}^2\right) p_k \leq \mathcal{P}_{\text{tot}}, \qquad (3.26a) \\ & \sum_{k=1}^{K} \frac{g_{e_k} p_k}{g_{e_k} p_k \sigma_{\omega k}^2 + \sigma_{\tilde{e}k}^2} \leq I_{\text{ms}}, \qquad (3.26b) \end{aligned}$$

where $I_{\text{ms}} = \frac{1}{\mathbb{D}_e} - \frac{1}{\sigma_\theta^2}$. As similar techniques depicted in Section 3.2.1 can be used to find a locally optimal solution, we omit details to avoid repetition.

Remark: Once the Lagrange dual functions are written for problem (3.25) and problem (3.26), one could notice that in the short-term scenario the power on the kth sensor depends only on its own channel conditions; whereas in the long-term scenario the transmission power of sensor k is a function of all sensors' channel information.

3.3.2 Partial CSI

Optimal Power Allocation

The optimal power allocation in partial CSI case is considered when the FC knows its channel full CSI but only has statistical knowledge of the eavesdropper. In this scenario, the problem is formulated similarly as problem (3.25) with the power scheme $\{p_k\}$ only being a function of the FC's channel information. A locally optimal solution can be then derived by applying similar techniques as used in [33] and Section 3.2.2; thus details regarding the optimal power allocation in partial CSI case are omitted. The simulation results are given in Section 3.5 for comparison.

Partial CSI - Artificial Noise with Relays

In a multiple-sensor network with only the FC's channel information, artificial noise can be produced when the observation information is crucial or there is a high security requirement. Different from Section 3.2.2, the concept of artificial noise in the multiple-sensor scenario is to transmit some interfering signals from a few sensors which can be cancelled off at the intended receiver (the FC), but would significantly degrade the eavesdropper's channel [27]. Instead of forwarding the observation signal to the FC, some sensors broadcast artificial noise to confuse the eavesdropper in the network. In this subsection, we assume that a total number of M sensors estimate the source θ and then transmit the observation information $\{x_m\}_{m=1}^{M}$ to the FC, while the remaining $K-M$ sensors work as relays aiming to boost the secret transmission of the information $\{x_m\}$. This extends the setup of [27] in which there is one transmitter and $K-1$ relays.

The transmission is completed in two stages. Let h_{s_mF}, h_{s_me}, and $h_{s_mr_k}$ be the channels from sensor m to the FC, the eavesdropper and relay k respectively. Denote h_{Fe} and h_{Fr_k} as the channels from the FC to the eavesdropper and relay k respectively. At the first stage, as shown in Figure 3.3, sensor m and the FC transmit $n_s h_{s_mF}$ and n_F respectively,

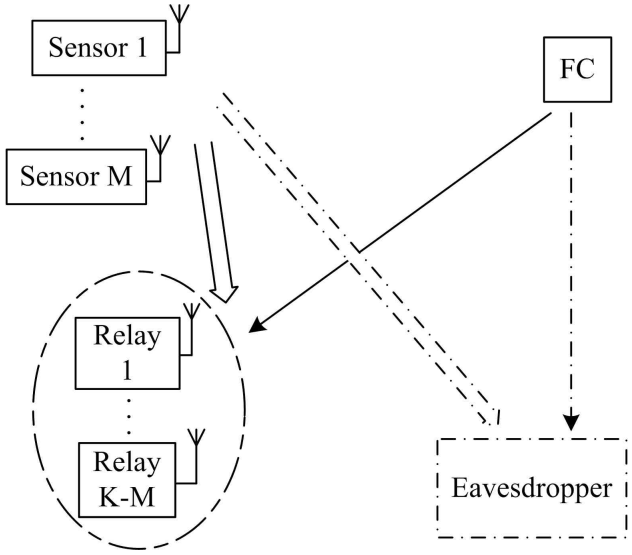

Figure 3.3: Diagram of stage one transmission in the artificial noise with relays.

and the eavesdropper and relay k receive, respectively,

$$y_{e,1} = n_s \sum_{m=1}^{M} h_{s_m F} h_{s_m e} + h_{Fe} n_F + z_{e,1},$$
$$y_{r_k,1} = n_s \sum_{m=1}^{M} h_{s_m F} h_{s_m r_k} + h_{F r_k} n_F + z_{k,1},$$

where $z_{e,1}$ and $z_{k,1}$ are zero mean i.i.d. complex Gaussian channel noises at the eavesdropper and at the kth relay with variances σ_e^2 and σ_r^2 respectively. n_s and n_F are artificial noises with variances σ_{ns}^2 and σ_{nF}^2 respectively.

At the second stage, sensor m forwards to the FC the amplified observation signal $x_m \beta_m$, and it also utilizes the public weight sequences $\{\bar{\gamma}_k\}$, which is a publicly available sequence of weights that are known to every participant (may also be seen by the eavesdropper) in the network, to transmit $-n_s \sum_{k=1}^{K-M} \bar{\gamma}_k h_{s_m r_k} h_{r_k F}$. We assume all the public sequences $\{\bar{\gamma}_k\}$ are i.i.d. zero mean complex Gaussian random variables with variance $\sigma_{\bar{\gamma}}^2$, thus $\{\bar{\gamma}_k\}$ varies at each transmission to reduce the probability of the artificial noise being nulled at the eavesdropper. On the other hand, relay k transmits $\bar{\gamma}_k y_{r_k,1}$. Therefore,

at the second stage the eavesdropper and the FC receive respectively,

$$\begin{aligned} y_{e,2} = &\sum_{m=1}^{M} x_m h_{s_m e}\beta_m + n_s \sum_{k=1}^{K-M} \bar{\gamma}_k \sum_{m=1}^{M} h_{s_m r_k}\left(h_{s_m F} h_{r_k e} - h_{r_k F} h_{s_m e}\right) \\ &+ n_F \sum_{k=1}^{K-M} \bar{\gamma}_k h_{F r_k} h_{r_k e} + \sum_{k=1}^{K-M} \bar{\gamma}_k h_{r_k e} z_{k,1} + z_{e,2}, \end{aligned} \tag{3.27}$$

$$\begin{aligned} y = &\sum_{m=1}^{M} x_m h_{sF}\beta_m + \sum_{k=1}^{K-M} \bar{\gamma}_k h_{r_k F}\left(h_{F r_k} n_F + z_{k,1}\right) + z \\ \overset{(b)}{\Longrightarrow} &\sum_{m=1}^{M} x_m h_{s_m F}\beta_m + \sum_{k=1}^{K-M} \bar{\gamma}_k z_{k,1} h_{r_k F} + z, \end{aligned} \tag{3.28}$$

where $z_{e,2}$ and z are zero mean i.i.d. complex Gaussian channel noises at the eavesdropper and the FC respectively with variances σ_e^2 and σ_n^2. In (3.28), (b) holds as n_F is known to the FC which can be cancelled off. Note that we assume synchronization between all sensors and the FC is available in this part of work, and the two-stage transmission can be completed in one fading block, as a result all the channels remain the same at the second stage transmission.

Remark: It is clear that the second term of (3.27) corresponds to the artificial noise generated from the M sensors at the first stage of transmission, which vanishes at the second stage as it reaches the FC (as can be seen in (3.28)). In channel conditions where $h_{s_m F} h_{r_k e}$ is close to $h_{r_k F} h_{s_m e}$, instead of increasing the transmit power at both the first and the second stage transmissions to boost noise level, we expect to use the third term of (3.27) to increase the noise level at the eavesdropper with little power consumption.

Combining the two-stage transmission, we have that the signal received by the eavesdropper given as

$$\mathbf{y}_e = \left[0, \sum_{m=1}^{M} h_{s_m e}\beta_m\left(\theta + \omega_m\right)\right]^{\mathrm{T}} + \mathbf{H}_{re}\left[n_s, n_F\right]^{\mathrm{T}} + \mathbf{z}_e, \tag{3.29}$$

where $\mathbf{z}_e = \left[z_{e,1},\ \sum_{k=1}^{K-M} \bar{\gamma}_k h_{r_k e} z_{k,1} + z_{e,2}\right]^{\mathrm{T}}$ and $\mathbf{H}_{re}$ is expressed in (3.30).

$$\mathbf{H}_{re} = \begin{bmatrix} \sum_{m=1}^{M} h_{s_m F} h_{s_m e} & h_{Fe} \\ \sum_{k=1}^{K-M} \bar{\gamma}_k \sum_{m=1}^{M} h_{s_m r_k} \left(h_{s_m F} h_{r_k e} - h_{r_k F} h_{s_m e}\right) & \sum_{k=1}^{K-M} \bar{\gamma}_k h_{F r_k} h_{r_k e} \end{bmatrix} \tag{3.30}$$

Hence, the total power consumption P_{stages} for the two-stage transmission can be also derived as

$$\begin{aligned} P_{\text{stages}} =& \sigma_{ns}^2 \left(\sum_{m=1}^{M} |h_{s_m F}|^2 + \sigma_{\bar{\gamma}}^2 \sum_{m=1}^{M} |h_{s_m F}|^2 \sum_{k=1}^{K-M} |h_{s_m r_k}|^2 + \sigma_{\bar{\gamma}}^2 \sum_{m=1}^{M} \sum_{k=1}^{K-M} |h_{s_m r_k}|^2 |h_{r_k F}|^2 \right) \\ &+ \sigma_{\bar{\gamma}}^2 (K-M) \sigma_r^2 + \sigma_{nF}^2 \sigma_{\bar{\gamma}}^2 \sum_{k=1}^{K-M} |h_{F r_k}|^2 + \sum_{m=1}^{M} |\beta_m|^2 \left(\sigma_{\omega m}^2 + \sigma_{\omega}^2\right). \end{aligned} \tag{3.31}$$

Let $\mathbf{K}_e$ be the covariance matrix of $\left[0,\ \sum_{m=1}^{M} h_{s_m e} \beta_m \omega_m\right]^{\mathrm{T}} + \mathbf{H}_{re} \left[n_s, n_F\right]^{\mathrm{T}} + \mathbf{z}_e$. As n_s, n_F, $z_{e,1}$, $z_{e,2}$, $\{z_{k,1}\}$ and $\{\omega_m\}$ are all independent random noises, $\mathbf{K}_e$ can be computed as

$$\mathbf{K}_e = \begin{bmatrix} \sigma_{ns}^2 \sum_{m=1}^{M} |h_{s_m F}|^2 |h_{s_m e}|^2 + \sigma_{nF}^2 |h_{Fe}|^2 + \sigma_e^2 & 0 \\ 0 & k_{e22} \end{bmatrix},$$

where k_{e22} is given as:

$$\begin{aligned} k_{e22} =& \sigma_{\bar{\gamma}}^2 \sigma_{ns}^2 \sum_{k=1}^{K-M} \sum_{m=1}^{M} |h_{s_m r_k}|^2 \left|h_{s_m F} h_{r_k e} - h_{r_k F} h_{s_m e}\right|^2 + \sigma_{\bar{\gamma}}^2 \sigma_{nF}^2 \sum_{k=1}^{K-M} |h_{F r_k}|^2 |h_{r_k e}|^2 \\ &+ \sigma_{\bar{\gamma}}^2 \sigma_r^2 \sum_{k=1}^{K-M} |h_{r_k e}|^2 + \sum_{m=1}^{M} |\beta_m h_{s_m e}|^2 \sigma_{\omega m}^2 + \sigma_e^2. \end{aligned} \tag{3.32}$$

Using the optimal MMSE estimator [49], from (3.28), (3.29) and (3.32) we can express the distortion D at the FC and the distortion D_e at the eavesdropper as

$$D = \left(\frac{1}{\sigma_\theta^2} + \frac{\left|\sum_{m=1}^{M} \beta_m h_{s_m F}\right|^2}{\sum_{m=1}^{M} |h_{s_m F} \beta_m|^2 \sigma_{\omega m}^2 + \sigma_{\bar{\gamma}}^2 \sigma_r^2 \sum_{k=1}^{K-M} |h_{r_k F}|^2 + \sigma_n^2} \right)^{-1} \tag{3.33}$$

$$D_e = \sigma_\theta^2 \left(1 - \frac{\sigma_\theta^2 \left| \sum_{m=1}^{M} h_{s_m e} \beta_m \right|^2}{k_{e22} + \sigma_\theta^2 \left| \sum_{m=1}^{M} h_{s_m e} \beta_m \right|^2} \right). \tag{3.34}$$

In the partial CSI scenario, the FC is able to obtain the channel information of $\{h_{s_m F}\}$, $\{h_{F r_k}\}$, $\{h_{s_m r_k}\}$ and $\{h_{r_k F}\}$ at each fading block, thus it can develop an intelligent transmission strategy such that the $\{\beta_m\}$, the variance of the public sequences $\sigma_{\tilde{\gamma}}^2$ and the artificial noise powers σ_{ns}^2, σ_{nF}^2 can be adapted in different fading blocks, while satisfying the long-term constraints as described in Section 3.3.1. Let $G = [\{h_{s_m F}\},\ \{h_{F r_k}\},\ \{h_{s_m r_k}\},\ \{h_{r_k F}\}]$. The functional optimization problem can be then formulated as

$$\begin{aligned} \min_{\{\beta_m(G)\},\, \sigma_{\tilde{\gamma}}^2(G),\, \sigma_{ns}^2(G),\, \sigma_{nF}^2(G)} \quad & \mathbb{E}\left[D\right] \\ s.t. \quad & \mathbb{E}\left[P_{\text{stages}}\right] \leq \mathcal{P}_{\text{tot}}, \\ & \mathbb{E}\left[D_e\right] \geq \mathbb{D}_e, \\ & \mathbb{E}\left[D_e\right] \geq \mathbb{E}\left[D\right], \end{aligned} \tag{3.35}$$

where P_{stages}, D and D_e are expressed in (3.31), (3.33) and (3.34), which are functions of $\{\beta_m\}$, $\sigma_{\tilde{\gamma}}^2$, σ_{ns}^2, and σ_{nF}^2. We can then employ the same Lagrange multiplier technique as described in Section 3.2.2 to solve problem (3.35), where in **Algorithm 1**, we need to sequentially find $\{\beta_m(G)\}$, $\sigma_{\tilde{\gamma}}^2(G)$, $\sigma_{ns}^2(G)$, and $\sigma_{nF}^2(G)$. The details are omitted for brevity.

Partial CSI - Asymptotic Analysis

In order to see how the system performs as the number of sensors increases, in this section, we explore the asymptotic long-term distortion at the FC in the case of partial CSI. For analytical tractability, we consider a homogeneous wireless sensor network where all the measurement noise and fading distributions are i.i.d.. As a consequence, we denote

$\sigma_{\omega k}^2 = \sigma_\omega^2$, $\mathbb{E}[g_k] = \sigma_h^2$, $\mathbb{E}[g_{e_k}] = \sigma_{h_e}^2$, $\sigma_{n k}^2 = \sigma_n^2$ and $\sigma_{e\,k}^2 = \sigma_e^2$, $\forall k$[8]. We also assume that the channel conditions of the FC and the eavesdropper satisfy $\frac{\sigma_h^2}{\sigma_n^2} \geq \frac{\sigma_{h_e}^2}{\sigma_e^2}$, as a result the FC always has better estimation quality than that of the eavesdropper when the number of sensors is sufficiently large. In addition, if the secrecy constraint and the transmit power constraint are satisfied at every transmission, the long-term power constraint as well as the long-term secrecy constraint can also be met.

With equal power allocation, i.e. $p_k = p\ \forall k$, we can rewrite the short-term secrecy constraint (3.26b) as

$$\frac{K}{\sigma_\omega^2 p}\frac{1}{K}\sum_{k=1}^{K}\frac{1}{g_{e_k}+\frac{\sigma_e^2}{\sigma_\omega^2 p}} \geq \frac{K - I_{\text{ms}}\sigma_\omega^2}{\sigma_e^2}. \tag{3.36}$$

It is straightforward to show that $\frac{1}{g_{e_k}+\frac{\sigma_e^2}{\sigma_\omega^2 p}}$ is convex in $g_{e_k}\ \forall k$. For large K, applying Jensen's inequality we have

$$\begin{aligned}
&\frac{K}{\sigma_\omega^2 p}\frac{1}{K}\sum_{k=1}^{K}\frac{1}{g_{e_k}+\frac{\sigma_e^2}{\sigma_\omega^2 p}} \\
\geq &\frac{K}{\sigma_\omega^2 p}\frac{1}{\frac{\sum_{k=1}^{K} g_{e_k}}{K}+\frac{\sigma_e^2}{\sigma_\omega^2 p}} \\
\sim &\frac{K}{\sigma_\omega^2 p}\frac{1}{\sigma_{h_e}^2+\frac{\sigma_e^2}{\sigma_\omega^2 p}}.
\end{aligned} \tag{3.37}$$

Therefore, we can set $\frac{K}{\sigma_\omega^2 p}\frac{1}{\sigma_{h_e}^2+\sigma_e^2/(\sigma_\omega^2 p)} \geq \frac{K-I_{\text{ms}}\sigma_\omega^2}{\sigma_e^2}$ to guarantee that the secrecy constraint is met (for large K). Let $r_e = \sigma_{h_e}^2 \Big/ \sigma_e^2$. Together with the short-term transmit power constraint (3.26a), the transmission power is given as

$$p = \min\left[\frac{\mathcal{P}_{\text{tot}}}{K\left(\sigma_\omega^2+\sigma_\theta^2\right)},\ \frac{I_{\text{ms}}}{r_e\left(K - I_{\text{ms}}\sigma_\omega^2\right)}\right]. \tag{3.38}$$

[8] As both channels of the FC and the eavesdropper are distributed as i.i.d. zero mean complex Gaussian (Rayleigh fading) with variances σ_h^2 and $\sigma_{h_e}^2$ respectively, we know that the channel power gains g_k and g_{e_k} are exponentially distributed with means σ_h^2 and $\sigma_{h_e}^2$ respectively.

When $p = \frac{I_{\text{ms}}}{r_e(K - I_{\text{ms}}\sigma_\omega^2)}$, from (3.24a) we have

$$\begin{aligned} D &= \left(\frac{1}{\sigma_\theta^2} + \frac{1}{\sigma_\omega^2} \sum_{k=1}^{K} \frac{g_k}{g_k + \frac{\sigma_n^2}{\sigma_\omega^2 p}} \right)^{-1} \\ &= \left(\frac{1}{\sigma_\theta^2} + \frac{1}{\sigma_\omega^2} \sum_{k=1}^{K} \frac{g_k}{g_k + \frac{\sigma_n^2 r_e}{\sigma_\omega^2 I_{\text{ms}}} K - \sigma_n^2 r_e} \right)^{-1} \\ &\overset{(c)}{\sim} \left(\frac{1}{\sigma_\theta^2} + \frac{K}{\sigma_\omega^2} \mathbb{E}\left[\frac{g_1}{g_1 + \frac{\sigma_n^2 r_e}{\sigma_\omega^2 I_{\text{ms}}} K - \sigma_n^2 r_e} \right] \right)^{-1}, \end{aligned} \tag{3.39}$$

provided the expectation $\mathbb{E}\left[\frac{g_k}{g_k + \sigma_e^2/(\sigma_\omega^2 p)}\right]$ exists. (c) is the result of applying a strong law of large numbers for triangular arrays [40]. Hence, the long-term distortion at the FC is given as

$$\mathbb{E}[D] \sim \left(\frac{1}{\sigma_\theta^2} + \frac{K}{\sigma_\omega^2} \mathbb{E}\left[\frac{g_1}{g_1 + \frac{\sigma_n^2 r_e}{\sigma_\omega^2 I_{\text{ms}}} K - \sigma_n^2 r_e} \right] \right)^{-1}. \tag{3.40}$$

As $g_k, \forall k$ is exponentially distributed with mean σ_h^2, we have

$$\begin{aligned} &\mathbb{E}\left[\frac{g_1}{g_1 + \frac{\sigma_n^2 r_e}{\sigma_\omega^2 I_{\text{ms}}} K - \sigma_n^2 r_e} \right] \\ =&1 + \frac{-\sigma_n^2 r_e \left(K - \sigma_\omega^2 I_{\text{ms}}\right)}{\sigma_h^2 \sigma_\omega^2 I_{\text{ms}}} e^{\frac{\sigma_n^2 r_e \left(K - \sigma_\omega^2 I_{\text{ms}}\right)}{\sigma_h^2 \sigma_\omega^2 I_{\text{ms}}}} \mathrm{E}_1\left[\frac{\sigma_n^2 r_e \left(K - \sigma_\omega^2 I_{\text{ms}}\right)}{\sigma_h^2 \sigma_\omega^2 I_{\text{ms}}} \right] \\ \sim&\frac{\sigma_h^2 \sigma_\omega^2 I_{\text{ms}}}{\sigma_n^2 r_e \left(K - \sigma_\omega^2 I_{\text{ms}}\right)} - \frac{2\sigma_h^4 \sigma_\omega^4 I_{\text{ms}}^2}{\sigma_n^4 r_e^2 \left(K - \sigma_\omega^2 I_{\text{ms}}^2\right)^2}, \end{aligned} \tag{3.41}$$

where function $\mathrm{E}_1[z]$ is related to the exponential integral $\mathrm{Ei}[z]$ through the expression $\mathrm{E}_1[z] = -\mathrm{Ei}[-z] = \int_z^\infty e^{-t} t^{-1} dt$ [42].

The case when $p = \frac{\mathcal{P}_{\text{tot}}}{K\left(\sigma_\omega^2 + \sigma_\theta^2\right)}$ has been explored in [56]. Therefore, combining the results of (3.38), (3.40) and (3.41), we have the long-term distortion at the FC being written as

$$\mathbb{E}[D] \sim \frac{\sigma_\theta^2 \sigma_n^4}{\sigma_n^4 + \sigma_h^2 \sigma_\theta^2 \sigma_n^2 \psi - \frac{2\sigma_h^4 \sigma_\omega^2 \sigma_\theta^2 \psi^2}{K}}$$

$$\sim \frac{\sigma_\theta^2 \sigma_n^2}{\sigma_n^2 + \sigma_h^2 \sigma_\theta^2 \psi} + \frac{2\sigma_h^4 \sigma_\omega^2 \psi^2 \sigma_\theta^4}{\left(\sigma_n^2 + \sigma_h^2 \psi \sigma_\theta^2\right)^2 K}, \tag{3.42}$$

where

$$\psi = \begin{cases} \frac{I_{\text{ms}}}{r_e(1-I_{\text{ms}})}, & I_{\text{ms}} < \frac{\mathcal{P}_{\text{tot}} r_e K}{K\left(\sigma_\theta^2+\sigma_\omega^2\right)+\mathcal{P}_{\text{tot}} r_e \sigma_\omega^2} \\ \frac{\mathcal{P}_{\text{tot}}}{\sigma_\theta^2+\sigma_\omega^2}, & \text{otherwise.} \end{cases} \tag{3.43}$$

Remark: It can be noticed from (3.42) and (3.43), that when the channel conditions of the FC and the eavesdropper satisfy $\frac{\sigma_h^2}{\sigma_n^2} \geq \frac{\sigma_{h_e}^2}{\sigma_e^2}$, for any given total transmission power $\mathcal{P}_{\text{tot}}$ and secrecy threshold at the eavesdropper $\mathbb{D}_e$ ($I_{\text{ms}} = \frac{1}{\mathbb{D}_e} - \frac{1}{\sigma_\theta^2}$ as defined in (3.26)), ψ is fixed for all fading blocks. In addition, the distortion at the FC decays to $\frac{\sigma_\theta^2 \sigma_n^2}{\sigma_n^2+\sigma_h^2 \sigma_\theta^2 \psi}$ (as the number of sensors increases) at the rate $1/K$.

3.4 Multiple Sensors Multiple Antennas Scenario

In this section, we want to explore the distortion performance for multiple sensors each equipped with multiple transmit antennas but with a single receive antenna at the FC and the eavesdropper. Let $\mathbf{h}_k = \left[h_{k,1}, \ldots, h_{k,N_k}\right]$ and $\mathbf{h}_{ek} = \left[h_{ek,1}, \ldots, h_{ek,N_k}\right]$ be the channels from the kth sensor to the FC and the eavesdropper respectively. We assume the entries of both $\mathbf{h}_k$ and $\mathbf{h}_{ek}$ are i.i.d. distributed zero mean complex Gaussian with variances $\left\{\sigma_{h\,k}^2\right\}$ and $\left\{\sigma_{h_e k}^2\right\}$ respectively. At each transmission, sensor k adopts the analog amplify and forward techniques by scaling the measurement with a amplifying factor $\boldsymbol{\beta}_k \in \mathbb{C}^{N_k \times 1}$. The FC and the eavesdropper receive, respectively,

$$y_k = \mathbf{h}_k \boldsymbol{\beta}_k \theta + \mathbf{h}_k \boldsymbol{\beta}_k \omega_k + z_k, \tag{3.44a}$$

$$y_{e_k} = \mathbf{h}_{ek} \boldsymbol{\beta}_k \theta + \mathbf{h}_{ek} \boldsymbol{\beta}_k \omega_k + z_{ek}. \tag{3.44b}$$

As a result, by employing the MMSE estimator, the distortion D at the FC and the distortion D_e at the eavesdropper can be written as

$$D = \left(\frac{1}{\sigma_\theta^2} + \sum_{k=1}^{K} \frac{(\mathbf{h}_k \boldsymbol{\beta}_k)^{\mathrm{H}} \mathbf{h}_k \boldsymbol{\beta}_k}{(\mathbf{h}_k \boldsymbol{\beta}_k)^{\mathrm{H}} \mathbf{h}_k \boldsymbol{\beta}_k \sigma_{\omega k}^2 + \sigma_{n k}^2} \right)^{-1}, \tag{3.45a}$$

$$D_e = \left(\frac{1}{\sigma_\theta^2} + \sum_{k=1}^{K} \frac{(\mathbf{h}_{ek} \boldsymbol{\beta}_k)^{\mathrm{H}} \mathbf{h}_{ek} \boldsymbol{\beta}_k}{(\mathbf{h}_{ek} \boldsymbol{\beta}_k)^{\mathrm{H}} \mathbf{h}_{ek} \boldsymbol{\beta}_k \sigma_{\omega k}^2 + \sigma_{e k}^2} \right)^{-1}. \tag{3.45b}$$

For short-term optimal power allocation, with a transmit power constraint at the sensors and a secrecy constraint at the eavesdropper, the optimization problem can be formulated as

$$\begin{aligned} \min_{\boldsymbol{\beta}_k, \forall k} \quad & \sum_{k=1}^{K} \left(\frac{\sigma_{\omega k}^4}{\sigma_{n k}^2} (\mathbf{h}_k \boldsymbol{\beta}_k)^{\mathrm{H}} \mathbf{h}_k \boldsymbol{\beta}_k + \sigma_{\omega k}^2 \right)^{-1} \\ s.t. \quad & \sum_{k=1}^{K} \boldsymbol{\beta}_k^{\mathrm{H}} \boldsymbol{\beta}_k \left(\sigma_{\omega k}^2 + \sigma_\theta^2 \right) \leq \mathcal{P}_{\mathrm{tot}}, \\ & \sum_{k=1}^{K} \left(\frac{\sigma_{\omega k}^4}{\sigma_{e k}^2} (\mathbf{h}_{ek} \boldsymbol{\beta}_k)^{\mathrm{H}} \mathbf{h}_{ek} \boldsymbol{\beta}_k + \sigma_{\omega k}^2 \right)^{-1} \geq I_{\mathrm{msma}}, \end{aligned} \tag{3.46}$$

where $I_{\mathrm{msma}} = \sum_{k=1}^{K} 1/\sigma_{\omega k}^2 + 1/\sigma_\theta^2 - 1/\mathbb{D}_e$.

In the long-term optimal power allocation, we have an additional constraint to ensure that the FC has a better estimation quality than at the eavesdropper; thus, the functional optimization problem can be expressed in (3.47).

$$\begin{aligned} \min_{\boldsymbol{\beta}_k, \forall k} \quad & \mathbb{E}\left[\left(\frac{1}{\sigma_\theta^2} + \sum_{k=1}^{K} \frac{(\mathbf{h}_k \boldsymbol{\beta}_k)^{\mathrm{H}} \mathbf{h}_k \boldsymbol{\beta}_k}{(\mathbf{h}_k \boldsymbol{\beta}_k)^{\mathrm{H}} \mathbf{h}_k \boldsymbol{\beta}_k \sigma_{\omega k}^2 + \sigma_{n k}^2} \right)^{-1} \right] \\ s.t. \quad & \mathbb{E}\left[\sum_{k=1}^{K} \boldsymbol{\beta}_k^{\mathrm{H}} \boldsymbol{\beta}_k \left(\sigma_{\omega k}^2 + \sigma_\theta^2 \right) \right] \leq \mathcal{P}_{\mathrm{tot}}, \\ & \mathbb{E}\left[\left(\frac{1}{\sigma_\theta^2} + \sum_{k=1}^{K} \frac{(\mathbf{h}_{ek} \boldsymbol{\beta}_k)^{\mathrm{H}} \mathbf{h}_{ek} \boldsymbol{\beta}_k}{(\mathbf{h}_{ek} \boldsymbol{\beta}_k)^{\mathrm{H}} \mathbf{h}_{ek} \boldsymbol{\beta}_k \sigma_{\omega k}^2 + \sigma_{e k}^2} \right)^{-1} \right] \geq \mathbb{D}_e, \\ & \mathbb{E}\left[\left(\frac{1}{\sigma_\theta^2} + \sum_{k=1}^{K} \frac{(\mathbf{h}_{ek} \boldsymbol{\beta}_k)^{\mathrm{H}} \mathbf{h}_{ek} \boldsymbol{\beta}_k}{(\mathbf{h}_{ek} \boldsymbol{\beta}_k)^{\mathrm{H}} \mathbf{h}_{ek} \boldsymbol{\beta}_k \sigma_{\omega k}^2 + \sigma_{e k}^2} \right)^{-1} \right] \geq \end{aligned}$$

$$\mathbb{E}\left[\left(\frac{1}{\sigma_\theta^2}+\sum_{k=1}^{K}\frac{(\mathbf{h}_k\boldsymbol{\beta}_k)^{\mathrm{H}}\mathbf{h}_k\boldsymbol{\beta}_k}{(\mathbf{h}_k\boldsymbol{\beta}_k)^{\mathrm{H}}\mathbf{h}_k\boldsymbol{\beta}_k\sigma_{\omega k}^2+\sigma_{nk}^2}\right)^{-1}\right]. \tag{3.47}$$

We can apply the same techniques as previous sections to solve problems (3.46) and (3.47). We omit the details to avoid repetition.

3.5 Numerical Results

In this section, we first show the performance of a multiple-antenna single sensor system via numerical simulations. For simplicity, we consider the source θ to be Gaussian distributed with zero mean and variance $\sigma_\theta^2 = 1$ mW. The sensor measurement sensitivity is set to $\sigma_\omega^2 = 10^{-3}$ mW. We assume the same noise level for both the FC and the eavesdropper's channel, where $\sigma_n^2 = \sigma_e^2 = 10^{-8}$ mW. In the following simulation, the secrecy threshold is chosen from the range $0.05 \leq \mathbb{D}_e \leq 0.65$. Furthermore, we consider the pathloss of signal power at the FC and the eavesdropper following the free-space pathloss model [29]

$$PL = 20\log_{10}(Dist) + 20\log_{10}(f) - 27.55, \tag{3.48}$$

where $Dist \in \{d, d_e\}$ is the distance between the sensor and the FC or the eavesdropper in meters, and f is the signal frequency in megahertz (we assume the network uses operation frequency of 800MHz, and the sensor is closer to the FC than to the eavesdropper with the distance from the sensor to the FC and to the eavesdropper being set to 127m and 130m respectively). Thus, the channel power gain follows an exponential distribution with mean of $10^{-\frac{PL}{10}}$ mW.

Figure 3.4 illustrates the distortion performance at the FC when zero information leakage is achieved with the number of transmit antennas $N_t \in \{2, 3, 4\}$, for a wide range of transmission power budgets. With the eavesdropper's full CSI, we can rotate and transmit the information on the null space of the eavesdropper's channel by sacrificing only a proportion of the FC's channel gain, and hence no information is leaked to the eavesdropper. As $\mathcal{P}_{\text{tot}}$ increases, the distortion gradually approaches its lower bound $\sigma_\theta^2 - \frac{\sigma_\theta^2}{\sigma_\theta^2+\sigma_\omega^2}$,

i.e., 9.99×10^{-4}.

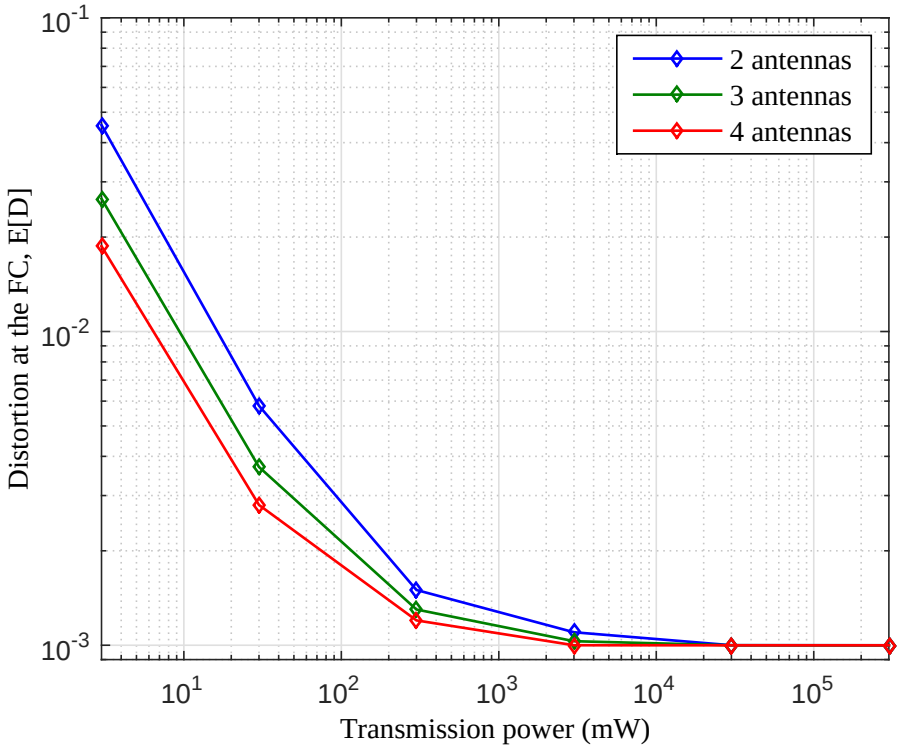

Figure 3.4: Performance comparison when zero information leakage is achieved.

Figure 3.5 depicts the distortion performance at the FC versus the secrecy threshold for a three-antenna single sensor system. For comparison, we plot the system performance under four scenarios: long-term full CSI, partial CSI, partial CSI with artificial noise and short-term full CSI. First, owing to the channel knowledge of both the FC and the eavesdropper, it is not surprising to see that the performance of the full CSI scenario is superior to the performance of partial CSI. Similar performance gains can be seen for the full CSI short-term distortion. We also notice the superior performance of artificial noise in the partial CSI case. This is because in the full CSI scenario, due to the full channel information of both FC and the eavesdropper, the direction of the beamformer can be designed to benefit the FC with little information being leaked to the eavesdropper; and in the case of artificial noise, a small amount of 'noise' is deliberately generated to degrade the eavesdropper's channel, which indicates that the secrecy threshold can be easily achieved without sacrificing much transmit power; whereas for the case of partial

CSI without artificial noise, some antennas need to be switched off to achieve the secrecy requirements, which is also the case for the short-term scenario[9].

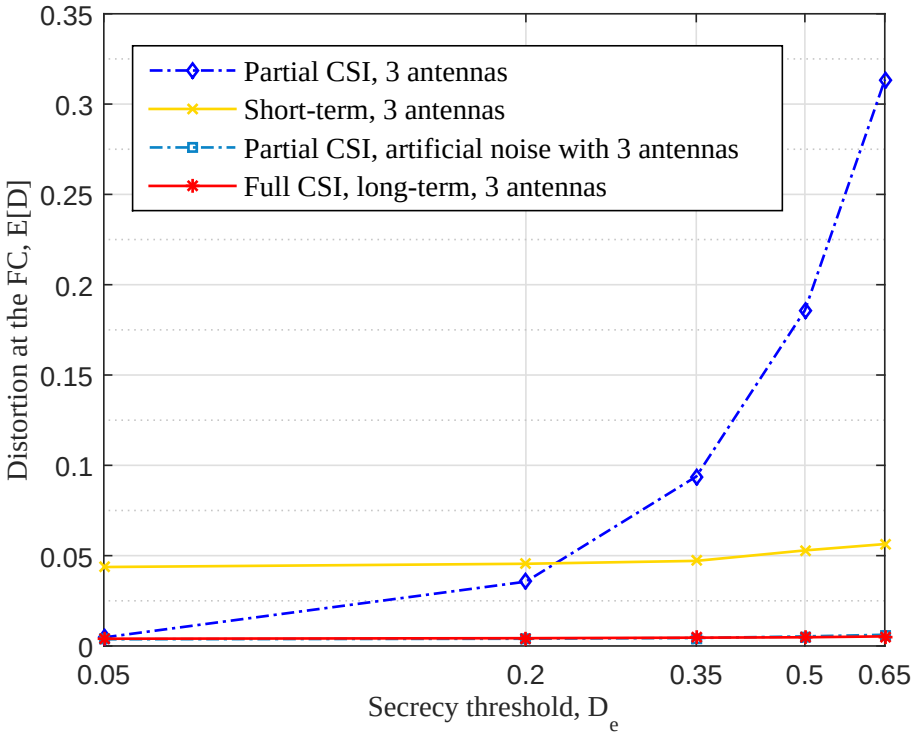

Figure 3.5: Performance comparison between full CSI, partial CSI and artificial noise in a multiple-antenna single sensor system.

We next present results for the asymptotic behaviour for the multiple-antenna single sensor scenario, where the beamforming vector is aligned with the FC's channel direction. In Figure 3.6, we can see that the asymptotic distortion performance of the results given in (3.21d) matches closely to the distortion at the FC obtained through simulations, and the gap gradually vanishes as N_t keeps increasing. Note that the asymptotic behaviour in a multiple-sensor network, obtained by applying (3.42), can be plotted similarly as Figure 3.6.

In the following, we study the distortion performance at the FC for a multiple-sensor network, where we assume the total transmit power budget is 30mW and all sensors share the same measurement sensitivity, i.e., $\sigma^2_{\omega_k} = \sigma^2_{\omega}$, $\forall k$. We apply the same pathloss

[9]Notice that the full CSI case is not completely overlapped with partial CSI artificial noise case, where the difference can be observed when $\mathbb{D}_e = 0.65$.

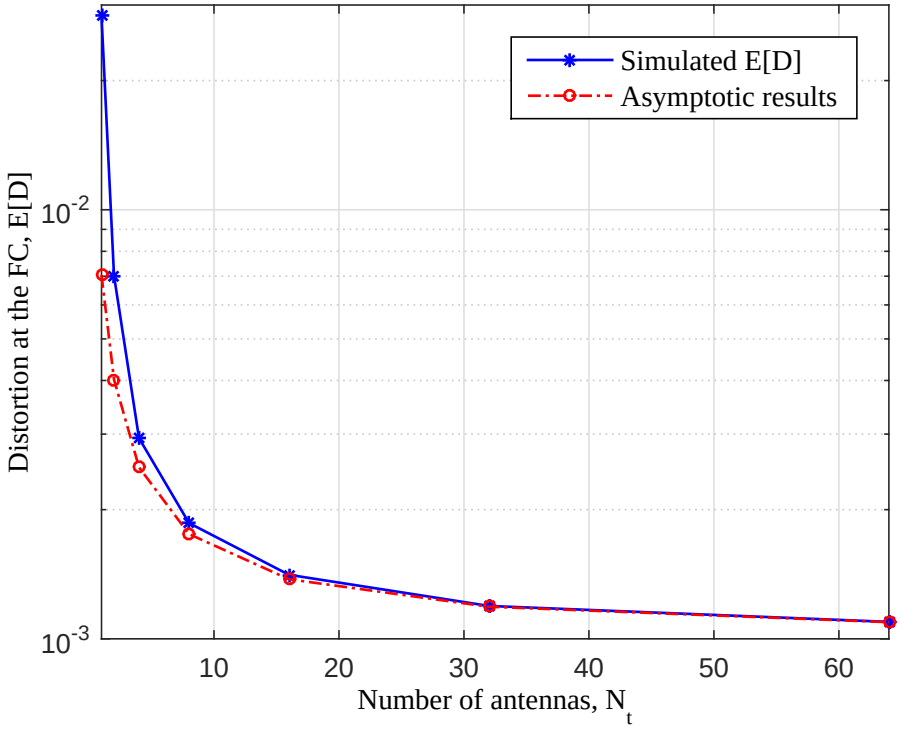

Figure 3.6: Asymptotic behaviour of $\mathbb{E}[D]$ in a multiple-antenna system.

model (3.48) and we consider the same noise level for both the FC and the eavesdropper's channel, where $\sigma_n^2 = \sigma_e^2 = 10^{-8}$ mW. From (3.24b) we notice that the distortion at the eavesdropper D_e drops to its minimum value $\frac{\sigma_\theta^2 \sigma_\omega^2}{K\sigma_\theta^2 + \sigma_\omega^2}$ as all the transmission powers approach infinity, and D_e would reach its maximum value σ_θ^2 when $\beta_k = 0$, $\forall k$.

In Figure 3.7, the secrecy threshold is chosen from the range $0.05 \leq \mathbb{D}_e \leq 0.25$. In the plot, the short-term distortion result is obtained by averaging over $10,000$ channel realizations. Not surprisingly, we can see that long-term distortion performances are superior to the performances of short-term power allocation problem due to a smaller feasibility region for the latter, where the sensors are required to ensure that the power constraint and the secrecy constraint are satisfied in every transmission slot.

In Figure 3.8, we study the system performance of a three-sensor network with two sensors working as relays to generate artificial noise. The secrecy threshold is set to $0.05 \leq \mathbb{D}_e \leq 0.8$. All the sensors (including two relays) are 127m away from the FC which is 3m closer than to the eavesdropper, and we also assume the distances from the

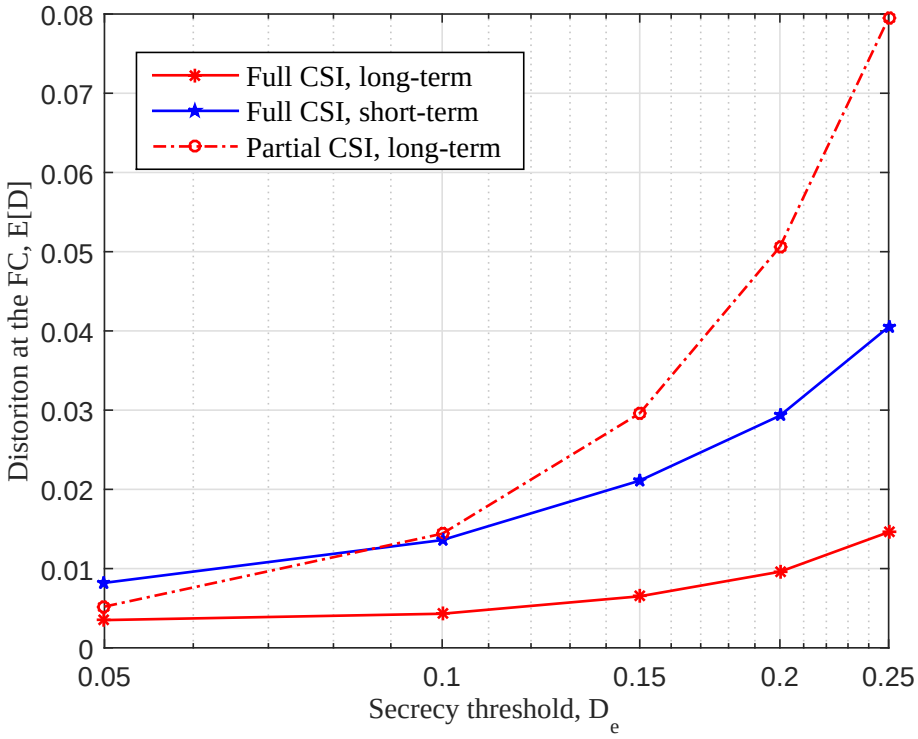

Figure 3.7: Performance comparison in an eight-sensor network, with $\sigma_\omega^2 = 10^{-3}$ mW, and the distance from each sensor to the FC and to the eavesdropper are 125m, 126m, 127m, 128m, 129m, 130m, 131m, 132m, and 139m, 138m, 137m, 136m, 135m, 131m, 130m, 129m respectively.

two relays to the sensor are 10m and 20m respectively. Due to diversity gains, it is clear to see the superior performance of the three-sensor network. As for the one-sensor two-relay network, it performs the same way as the one-sensor system when the distortion threshold is small; however, as the secrecy requirement increases at the eavesdropper, the performance gap grows. This is because the two relays are only activated when $\mathbb{D}_e$ is relatively large, where a small portion of the total transmit power is used to produce artificial noise to reach the secrecy threshold; whereas in the other two systems, without the eavesdropper's channel information, the sensor(s) may need to reduce the transmission power to achieve the high secrecy requirement.

In Figure 3.9, we compare the distortion performance of three different types of multiple-sensor network with a fixed total number of transmitting antennas of eight. It is seen that the distortion performance of the four-antenna two sensors network is followed by the

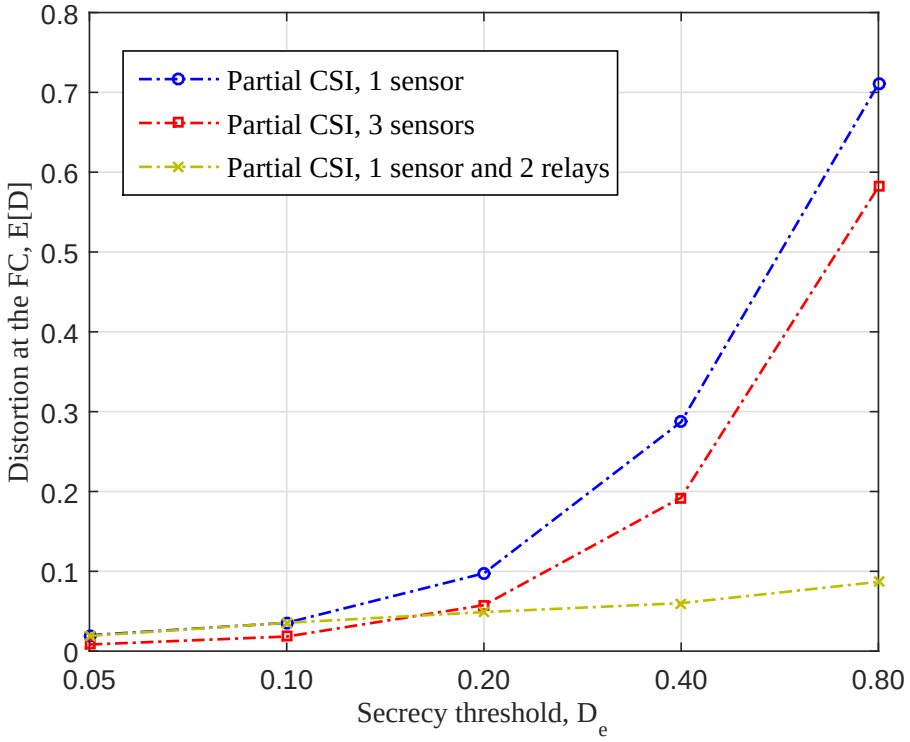

Figure 3.8: A multiple-sensor network with relays, with $\sigma_\omega^2 = 10^{-3}$ mW.

performance of the two-antenna four-sensor network, which are both superior to the single antenna eight-sensor scenario. This suggests that we can better utilize a multiple-antenna system for a point source estimation to achieve a better performance at the FC under the secrecy constraints.

3.6 Conclusion

In this chapter, we have considered the problem of transmit power allocation for distortion minimization in multisensor estimation in the presence of an eavesdropper, where the sensors can also have multiple transmit antennas. We studied the asymptotic behaviour for the long-term distortion at the FC under the equal power allocation for the multiple-sensor scenario, and also for the multiple-antenna-single-sensor scenario, where the transmit beamforming vector at the sensor is aligned with the direction of the FC. In addition, in a multiple-sensor network, when the secrecy requirement is high, some

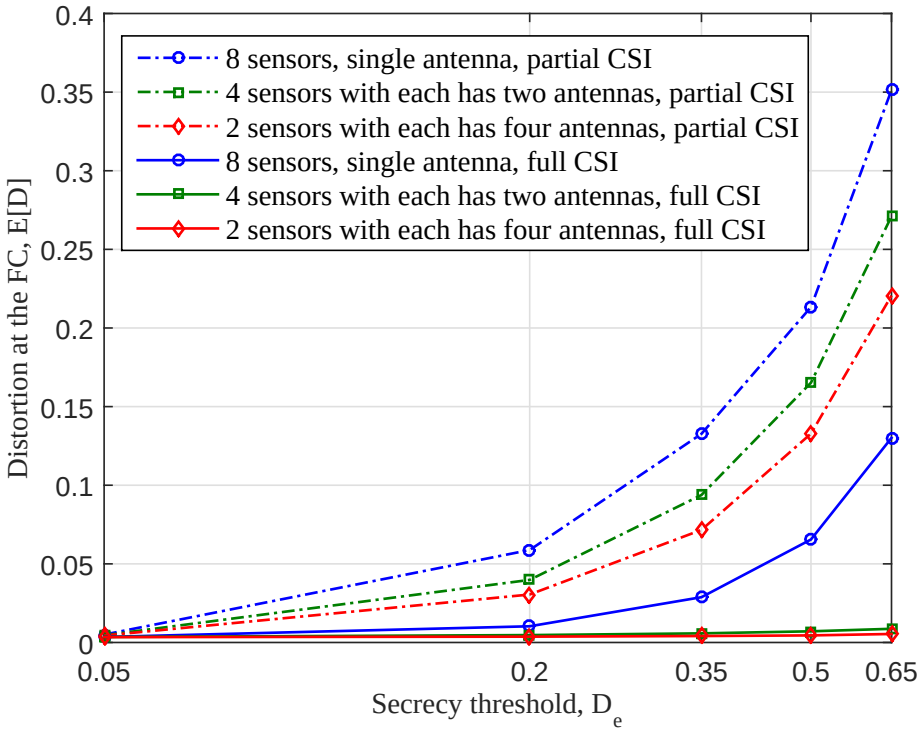

Figure 3.9: Performance comparison among multiple-sensor networks with the total number of transmitting antennas of eight.

sensors can be deployed to artificially produce noise to improve the transmission security. For the multiple-antenna-single-sensor system, depending on the availability of the eavesdropper's channel information, we can achieve zero information leakage or degrade the eavesdropper's channel and enhance the system performance by exploiting multiple-antenna techniques under the long-term power allocation scenario.

Chapter 4

Distortion Outage Minimization in Distributed Estimation with Estimation Secrecy Outage Constraints

Under fading channel assumptions, estimation error becomes a random variable as a function of the channel gains; hence it may not be always possible to satisfy the secrecy constraints considered in Chapter 3. In this chapter, we investigate the distortion outage minimization problem for a wireless sensor network (WSN) in the presence of an eavesdropper. Applying a rigorous probabilistic power allocation technique, we derive power policies for the full channel state information (CSI) case. Suboptimal power control policies are studied for the partial CSI case in order to reduce the high computational cost associated with large numbers of sensors or receive antennas. In the case of multiple transmit antennas, the distortion outage at the FC can be dramatically reduced and in some cases completely eliminated, by transmitting the observations on the null space of the eavesdropper's channel or deploying an artificial noise technique, in full CSI and partial CSI cases respectively.

4.1 Introduction

IN last chapter we looked at the optimal power allocation for a decentralized estimation problem in the presence of an eavesdropper. To secure the system, a minimum distortion threshold is set for the eavesdropper to ensure that the estimation error at the eavesdropper is no smaller than this threshold. However, due to the randomness of the fading channels, the quality of the estimate at the FC becomes a random variable. This might be detrimental to real-time applications when the distortion at the FC becomes large for a particular fading realisation, or the distortion at the eavesdropper becomes very small. Hence, for a delay constrained sensor network, instead of minimising a

long-term average estimation error at the FC as in Chapter 3, it is more appropriate to maintain a target distortion level throughout the fading process and minimise a *distortion outage probability*[1] at the FC and a secrecy outage constraint at the eavesdropper. This is the subject of our work in this chapter.

In the context of communications and information theory, the idea of *information outage* probability minimization was introduced in [64] for block-fading channels, and has been further extended in e.g. [64, 73]. A similar concept of *estimation outage* probability for distributed estimation was introduced by the authors in [15], which is defined as the probability that the estimation distortion exceeds a certain threshold. With full channel state information, the authors in [97] considered a clustered WSN and derived the optimal power allocation for estimation outage minimization problem; the results were extended to partial CSI with limited feedback in [98]. In [99], the authors explored the diversity order for distortion outage minimization over coherent MACs. Optimal power allocation for estimation outage probability minimization was also studied in [55] for state estimation of linear dynamical systems.

Apart from the limited battery life of the sensors, another crucial issue in a wireless sensor network is secrecy, as previously discussed in Chapter 3. Because of the open wireless media, maintaining a high level of secrecy in a wireless network is quite challenging. Various secrecy schemes have been investigated from a signal processing [52, 61, 67, 93] as well as from an information theoretic point of view [60, 62, 71, 96]. In favor of a closed form distortion expression for the sensor estimation over fading channels, we investigate the secure estimation problem from a signal processing viewpoint, which is more desirable for us to derive analytical results. Therefore, instead of applying secure source coding techniques we consider analog uncoded transmission at the sensors.

In this chapter, we look at a WSN where each sensor independently measures a single point Gaussian source, and then transmits the noisy measurements to the FC using an uncoded analog scheme over an orthogonal MAC in the presence of an eavesdropper or adversary. Both the FC and the adversary attempt to reconstruct a minimum mean

[1]This is analogous to the situation in wireless communication where the ergodic capacity describes the maximum achievable long term average rate without a delay constraint; however, in real-time applications because of the delay constraint it is more suitable to adopt the notion of the outage capacity, which determines the maximum achievable rate with an outage probability less than ϵ [11].

square error (MMSE) estimate of the observations. Under this setting, the main contributions of the chapter are: 1) We consider power allocation problems that minimise the distortion outage probability at the FC, subject to a long-term transmit power constraint and a secrecy outage constraint at the eavesdropper, where an *estimation secrecy outage* is defined as the event that the mean squared error (MSE) at the eavesdropper is below a minimum acceptable distortion level. In this way, the entire network is guaranteed to operate under a specified power constraint; while maintaining a certain level of confidentiality. 2) We study the distortion outage probability at the FC that can be achieved by adding multiple receive antennas in both the full CSI and partial CSI cases. In addition, we propose suboptimal power allocation policies to alleviate the high computational cost issues raised by computing for the locally optimal power policy in the partial CSI case. 3) As an alternative to having multiple sensors in a network, the scenario of a single sensor with multiple transmit antennas is investigated. Numerical studies illustrate that in both the full CSI and partial CSI cases, zero outage can be achieved at the FC with a sufficiently large power budget.

The rest of the chapter is organised as follows. In Section 4.2 we present the system model for a multiple-sensor network and solve the outage minimization problem. In Section 4.3 we investigate the secrecy outage problem for the multiple-antenna single sensor scenario and study optimal power control policies for both the full and partial CSI cases. In Section 4.4, alternative problems that can be solved by applying similar techniques are formulated. Illustrative numerical results are provided in Section 4.5, followed by concluding remarks in Section 4.6.

4.2 Multiple Sensors Scenario

A schematic diagram of the wireless sensor network model is shown in Figure 4.1, where we have K sensors observing a single point i.i.d. (independent and identically distributed) Gaussian source with zero mean and variance σ_θ^2, denoted by $\theta[t]$, $t = 0, 1, 2, \ldots$. The measurement $x_k[t]$ received by the kth sensor at time t is corrupted with noise and is

given by

$$x_k[t] = \theta[t] + \omega_k[t], \tag{4.1}$$

where $\omega_k[t]$ is the sensor measurement noise which is i.i.d. Gaussian with zero mean and variance $\sigma^2_{\omega k}$.

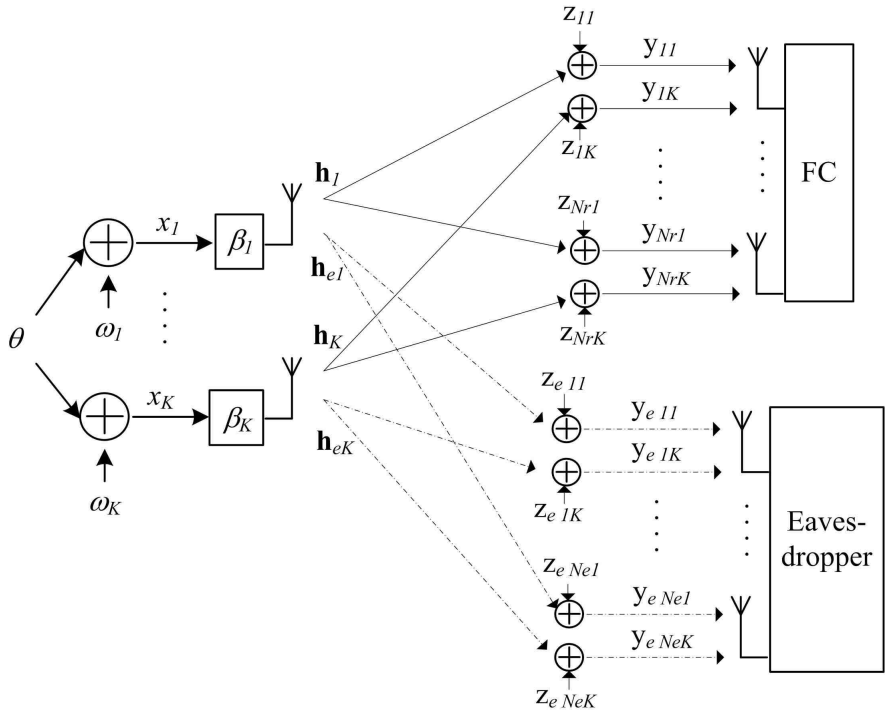

Figure 4.1: Diagram of a wireless sensor network using orthogonal MAC schemes with the presence of an eavesdropper.

The sensors are assumed to have a single transmit antenna, see Section 4.3 for the case of multiple transmit antennas. Each sensor amplifies and forwards its measurement to a N_r-antenna fusion center (FC) with amplification factor $\beta_k[t] \in \mathbb{C}$ via a slow-fading orthogonal multiple access channel (MAC), e.g. by using OFDMA or TDMA techniques. The transmissions are overheard by an eavesdropper who is equipped with N_e receive antennas. We assume that both the FC's and the eavesdropper's channels experience block fading, where the channels remain constant during each coherence time interval, and are i.i.d. over different time intervals [11]. The signals received by the FC and eaves-

dropper from the kth sensor are then given by, respectively,

$$\mathbf{y}_k[t] = \theta[t]\beta_k[t]\mathbf{h}_k[t] + \omega_k[t]\beta_k[t]\mathbf{h}_k[t] + \mathbf{z}_k[t], \tag{4.2a}$$

$$\mathbf{y}_{ek}[t] = \theta[t]\beta_k[t]\mathbf{h}_{ek}[t] + \omega_k[t]\beta_k[t]\mathbf{h}_{ek}[t] + \mathbf{z}_{ek}[t], \tag{4.2b}$$

where $\mathbf{y}_k[t] = [y_{1k}[t], \ldots, y_{N_r k}[t]]^{\mathrm{T}}$ and $\mathbf{y}_{ek}[t] = \left[y_{e1k}[t], \ldots, y_{eN_e k}[t]\right]^{\mathrm{T}}$, the entries of $\mathbf{h}_k[t]$ and $\mathbf{h}_{ek}[t]$ are the instantaneous zero mean i.i.d. complex Gaussian channels from sensor k to the FC and the eavesdropper with variances σ_{hk}^2 and $\sigma_{h_e k}^2$ respectively, and $\mathbf{z}_k[t] = [z_{1k}[t], \ldots, z_{N_r k}[t]]^{\mathrm{T}}$ and $\mathbf{z}_{ek}[t] = [z_{e1k}[t], \ldots, z_{eN_e k}[t]]^{\mathrm{T}}$ represent i.i.d. additive Gaussian noise with zero mean and covariances $\sigma_{nk}^2\mathbf{I}_{N_r}$ at the FC and $\sigma_{e\,k}^2\mathbf{I}_{N_e}$ at the eavesdropper respectively[2]. The set of received signals at the FC from all sensors can be written as

$$\begin{aligned}\mathbf{Y}[t] =& [\mathbf{y}_1[t], \ldots, \mathbf{y}_K[t]]^{\mathrm{T}} \\ =& \theta[t]\,[\beta_1[t]\mathbf{h}_1[t], \ldots, \beta_K[t]\mathbf{h}_K[t]]^{\mathrm{T}} \\ &+ [\mathbf{z}_1[t], \ldots, \mathbf{z}_K[t]]^{\mathrm{T}} + [\omega_1[t]\beta_1[t]\mathbf{h}_1[t], \ldots, \omega_K[t]\beta_K[t]\mathbf{h}_K[t]]^{\mathrm{T}}.\end{aligned} \tag{4.3}$$

Using the fact that each sensor transmits through an orthogonal MAC, the covariance of the noise factor $[\omega_1[t]\beta_1[t]\mathbf{h}_1[t], \ldots, \omega_K[t]\beta_K[t]\mathbf{h}_K[t]]^{\mathrm{T}} + [\mathbf{z}_1[t], \ldots, \mathbf{z}_K[t]]^{\mathrm{T}}$ can be derived as a $KN_r \times KN_r$ matrix:

$$C[t] = \begin{bmatrix} \sigma_{w1}^2\beta_1^2[t]\mathbf{h}_1[t]\mathbf{h}_1^{\mathrm{H}}[t] + \sigma_{n1}^2\mathbf{I}_{N_r} & & 0 \\ & \ddots & \\ 0 & & \sigma_{wK}^2\beta_K^2[t]\mathbf{h}_K[t]\mathbf{h}_K^{\mathrm{H}}[t] + \sigma_{nK}^2\mathbf{I}_{N_r} \end{bmatrix}. \tag{4.4}$$

The linear minimum mean square error (MMSE) estimator is well known to be the optimal estimator for θ under the model (4.2) [49]. At time t the mean squared error

[2]The notation $\mathbf{x}^{\mathrm{T}}$ and $\mathbf{x}^{\mathrm{H}}$ refers to the transpose of $\mathbf{x}$ and conjugate transpose of $\mathbf{x}$ respectively.

(MSE) or *distortion* at the FC using the MMSE estimator is

$$
\begin{aligned}
D[t] &= \left(\frac{1}{\sigma_\theta^2} + \begin{bmatrix} \beta_1[t]\mathbf{h}_1[t] \\ \vdots \\ \beta_K[t]\mathbf{h}_K[t] \end{bmatrix}^{\mathrm{H}} C[t]^{-1} \begin{bmatrix} \beta_1[t]\mathbf{h}_1[t] \\ \vdots \\ \beta_K[t]\mathbf{h}_K[t] \end{bmatrix} \right)^{-1} \\
&\overset{(a)}{=} \Bigg[\frac{1}{\sigma_\theta^2} + \sum_{k=1}^{K} \beta_k^{\mathrm{H}}[t]\beta_k[t] \Bigg(\sigma_{n_k}^{-2}\mathbf{h}_k^{\mathrm{H}}[t]\mathbf{h}_k[t] - \sigma_{n_k}^{-2}\mathbf{h}_k^{\mathrm{H}}[t]\mathbf{h}_k[t] \Big(\sigma_{w_k}^{-2}\beta_k^{-2}[t] \\
&\qquad + \sigma_{n_k}^{-2}\mathbf{h}_k^{\mathrm{H}}[t]\mathbf{h}_k[t] \Big)^{-1} \sigma_{n_k}^{-2}\mathbf{h}_k^{\mathrm{H}}[t]\mathbf{h}_k[t] \Bigg) \Bigg]^{-1} \\
&= \left(\frac{1}{\sigma_\theta^2} + \sum_{k=1}^{K} \frac{g_k[t]p_k[t]}{\sigma_{n_k}^2 + g_k[t]\sigma_{w_k}^2 p_k[t]} \right)^{-1}, \qquad (4.5)
\end{aligned}
$$

where (a) results from applying the Matrix Inversion Lemma [34], $p_k[t] \triangleq \beta_k^{\mathrm{H}}[t]\beta_k[t]$ is the power allocated on the kth sensor, and $g_k[t] \triangleq \mathbf{h}_k^{\mathrm{H}}[t]\mathbf{h}_k[t] = \sum_{m=1}^{N_r} h_{mk}^{\mathrm{H}}[t]h_{mk}[t]$ is the sum of channel power gains from the kth sensor to the FC with $h_{mk}[t]$ being the channel gain from sensor k to mth antenna at the FC. Note that for a given set of $\{p_k[t]\}$, any $\{\beta_k[t]\}$ satisfying $\beta_k[t]^{\mathrm{H}}\beta_k[t] = p_k[t], \forall k$ would result in the same distortion, hence our primary focus is $\{p_k[t]\}$. We assume the optimal power allocation strategy is designed by the FC, and then $\{p_k[t]\}$ are wirelessly transmitted to the sensors via a public channel[3]. The minimum distortion level at the eavesdropper is achieved by implementing the linear MMSE estimator, shown as

$$
D_e[t] = \left(\frac{1}{\sigma_\theta^2} + \sum_{k=1}^{K} \frac{g_{ek}[t]p_k[t]}{g_{ek}[t]p_k[t]\sigma_{\omega k}^2 + \sigma_{e\,k}^2} \right)^{-1}, \qquad (4.6)
$$

where $g_{ek}[t] \triangleq \mathbf{h}_{ek}^{\mathrm{H}}[t]\mathbf{h}_{ek}[t] = \sum_{n=1}^{N_e} h_{enk}^{\mathrm{H}}[t]h_{enk}[t]$ is the sum of channel power gains from the kth sensor to the eavesdropper and $h_{enk}[t]$ is the channel gain from sensor k to nth antenna at the eavesdropper. Due to the randomness of the fading channels, the instantaneous distortions at the FC and the eavesdropper, as shown in (4.5) and (4.6), change over time.

[3]When the feedback link is secure, the estimation distortion seen by the eavesdropper will be even larger than $D_e[t]$ given in (4.6), due to the lack of $\{p_k[t]\}$.

Different from our work in Chapter 3 where we studied optimal power allocation for an expected distortion minimization with security constraints at the eavesdropper, in this chapter we focus on the *distortion outage minimization* problem. For a given maximum acceptable distortion level $\mathbb{D}$ at the FC, we define a *distortion outage* to be the event that the instantaneous distortion $D[t]$ exceeds $\mathbb{D}$. The *distortion outage probability* at the FC is then given as $\Pr_{\text{outage\_FC}} \triangleq \Pr[D[t] > \mathbb{D}]$. At the eavesdropper, for a given minimum acceptable distortion level $\mathbb{D}_e$, a *secrecy outage* event is declared if the instantaneous distortion $D_e[t]$ is less than $\mathbb{D}_e$ (which means that the eavesdropper has a good quality estimate), and the *secrecy outage probability* is defined as $\Pr_{\text{outage\_EVE}} \triangleq \Pr[D_e[t] < \mathbb{D}_e]$. We assume that the full channel state information (CSI) of the sensor-to-FC channels are available at the FC, while eavesdropper's channel information may or may not be available at the FC.

In this chapter, we wish to minimise the distortion outage probability at the FC by adapting the transmit powers of the sensors at each channel instance, while keeping the secrecy outage probability under a certain threshold, i.e., $\Pr_{\text{outage\_EVE}} \leq \delta$, and the long-term average sum of sensor transmission powers, defined as $\mathbb{E}\left[\sum_{k=1}^{K} p_k \mathbb{E}\left[x_k^2[t]\right]\right] = \mathbb{E}\left[\sum_{k=1}^{K} p_k(\sigma_\theta^2 + \sigma_{\omega k}^2)\right]$, to be less than a power budget $\mathcal{P}_{\text{tot}}$.

Due to the assumption of system independence over time t, we will drop the time index t for the rest of the chapter.

4.2.1 Full CSI

In this section, we assume the FC can also acquire the channel information between the sensors and the eavesdropper. As a result, the power control policies can be derived such that sensors are able to adjust the transmission powers depending on both the FC's and the eavesdropper's channel information. Clearly, the requirement of full CSI of the eavesdropper channels is infeasible in practice. However, the optimal performance with this assumption is instructive as well as useful as a benchmark for the performance with partial CSI of the eavesdropper channels, to be analysed subsequently.

Let the channel states at the FC and the eavesdropper be denoted by $\mathbf{g} = [g_1, \ldots, g_K]$

and $\mathbf{g}_e = [g_{e1}, \ldots, g_{eK}]$ respectively. The outage minimization problem is

$$\begin{aligned} \min_{\mathbf{P}(\mathbf{G})} \quad & \Pr\left[D\left(\mathbf{G}, \mathbf{P}\left(\mathbf{G}\right)\right) > \mathbb{D}\right] \\ s.t. \quad & \Pr\left[D_e\left(\mathbf{G}, \mathbf{P}\left(\mathbf{G}\right)\right) < \mathbb{D}_e\right] \leq \delta, \qquad (4.7a) \\ & \mathbb{E}_{\mathbf{G},\mathbf{p}}\left[\langle \mathbf{P}\left(\mathbf{G}\right)\rangle\right] \leq \mathcal{P}_{\text{tot}}, \qquad (4.7b) \end{aligned}$$

where $\mathbf{G} = [\mathbf{g}; \mathbf{g}_e]$ and $\langle \mathbf{p}\left(\mathbf{G}\right)\rangle \triangleq \sum_{k=1}^{K}\left(\sigma_\theta^2 + \sigma_{\omega k}^2\right) p_k\left(\mathbf{G}\right)$ is the total power consumption. $\mathbf{P}\left(\mathbf{G}\right)$ is a vector of random variables with conditional probability density function $f_{\mathbf{P}|\mathbf{G}}\left(\mathbf{p}|\,\mathbf{G}\right)$, where $\mathbf{p}$ is one of the deterministic schemes and $\mathbf{p} = [p_1, \ldots, p_K]$ are the powers allocated across the sensors.

Notice that, from the expression of D_e in (4.6), when zero power is allocated to the sensors we obtain $D_e|_{\mathbf{p}=\mathbf{0}} = \sigma_\theta^2$, giving the largest possible distortion at the eavesdropper, while if the transmit power on each sensor approaches infinity we have the smallest possible distortion at the eavesdropper $D_e \to \left(\frac{1}{\sigma_\theta^2} + \sum_{k=1}^{K}\frac{1}{\sigma_{\omega k}^2}\right)^{-1}$. Therefore, in order to produce a meaningful solution to problem (4.7), $\mathbb{D}_e$ should satisfy $\left(\frac{1}{\sigma_\theta^2} + \sum_{k=1}^{K}\frac{1}{\sigma_{\omega k}^2}\right)^{-1} < \kappa < \mathbb{D}_e < \sigma_\theta^2$, where κ is a nonnegative threshold to ensure constraint (4.7a) is achievable for a given transmit power budget $\mathcal{P}_{\text{tot}}$ and a secrecy outage probability threshold δ.

In communications theory, it was shown in [11, 64] that for information outage minimization problems the optimal power allocation policy is in general a probabilistic policy, in particular this is often the case for discrete channel distributions. Motivated by these results, we start with a probabilistic power allocation $\mathbf{P}\left(\mathbf{G}\right)$.

Denote the indicator function by $1(x)$, where $1(x) = 1$ if x is true; otherwise $1(x) = 0$. With the assumption on the fading channels and perfect CSI at the FC, the distortion outage probability at the FC and the secrecy outage probability at the eavesdropper can be expressed as, respectively,

$$\begin{aligned} & \Pr\left[D\left(\mathbf{G}, \mathbf{P}\right) > \mathbb{D}\right] \\ = & \int\!\!\int 1\left\{D\left(\mathbf{G}, \mathbf{p}\right) > \mathbb{D}\right\} f_{\mathbf{P}|\mathbf{G}}\left(\mathbf{p}|\,\mathbf{G}\right) d\mathbf{p}\left(\mathbf{G}\right) dF\left(\mathbf{G}\right), \qquad (4.8) \\ & \Pr\left[D_e\left(\mathbf{G}, \mathbf{P}\right) < \mathbb{D}_e\right] \end{aligned}$$

$$= \iint 1\left\{D_e\left(\mathbf{G},\mathbf{p}\right) < \mathbb{D}_e\right\} f_{\mathbf{P}|\mathbf{G}}\left(\mathbf{p}|\,\mathbf{G}\right) d\mathbf{p}\left(\mathbf{G}\right) dF\left(\mathbf{G}\right). \tag{4.9}$$

We outline the strategy involved in solving problem (4.7), which are similar to techniques used in [64]. We first show that for an arbitrary feasible probabilistic power allocation $\mathbf{P}\left(\mathbf{G}\right)$, which can be divided into four non-overlapping power regions, we can always construct another feasible probabilistic power allocation $\hat{\mathbf{P}}\left(\mathbf{G}\right)$ that contains three power regions, with the powers in one of the regions all equal to zero, and such that $\hat{\mathbf{P}}\left(\mathbf{G}\right)$ gives no worse performance than $\mathbf{P}\left(\mathbf{G}\right)$. Next, based on $\hat{\mathbf{P}}\left(\mathbf{G}\right)$ we construct another feasible power scheme $\mathbf{P}'\left(\mathbf{G}\right)$ which is randomised among three deterministic power schemes $\left\{\mathbf{p}_i\left(\mathbf{G}\right)\right\}, i = 1,2,3$ with corresponding weighting factors $\left\{\omega_i\left(\mathbf{G}\right)\right\}$. Furthermore, we show that $\mathbf{P}'\left(\mathbf{G}\right)$ performs at least as well as $\hat{\mathbf{P}}\left(\mathbf{G}\right)$.

First, given a feasible probabilistic power scheme $\mathbf{P}\left(\mathbf{G}\right)$, we partition the powers into four non-overlapping power regions as given in (4.10).

$$\begin{aligned}
\mathcal{A}_1\left(\mathbb{D},\mathbb{D}_e,\mathbf{G}\right) &= \left\{\mathbf{p}\left(\mathbf{G}\right) : D\left(\mathbf{G},\mathbf{p}\left(\mathbf{G}\right)\right) \leq \mathbb{D}, D_e\left(\mathbf{G},\mathbf{p}\left(\mathbf{G}\right)\right) \geq \mathbb{D}_e|\mathbf{G}\right\} \\
\mathcal{A}_2\left(\mathbb{D},\mathbb{D}_e,\mathbf{G}\right) &= \left\{\mathbf{p}\left(\mathbf{G}\right) : D\left(\mathbf{G},\mathbf{p}\left(\mathbf{G}\right)\right) \leq \mathbb{D}, D_e\left(\mathbf{G},\mathbf{p}\left(\mathbf{G}\right)\right) < \mathbb{D}_e|\mathbf{G}\right\} \\
\mathcal{A}_3\left(\mathbb{D},\mathbb{D}_e,\mathbf{G}\right) &= \left\{\mathbf{p}\left(\mathbf{G}\right) : D\left(\mathbf{G},\mathbf{p}\left(\mathbf{G}\right)\right) > \mathbb{D}, D_e\left(\mathbf{G},\mathbf{p}\left(\mathbf{G}\right)\right) \geq \mathbb{D}_e|\mathbf{G}\right\} \\
\mathcal{A}_4\left(\mathbb{D},\mathbb{D}_e,\mathbf{G}\right) &= \left\{\mathbf{p}\left(\mathbf{G}\right) : D\left(\mathbf{G},\mathbf{p}\left(\mathbf{G}\right)\right) > \mathbb{D}, D_e\left(\mathbf{G},\mathbf{p}\left(\mathbf{G}\right)\right) < \mathbb{D}_e|\mathbf{G}\right\}
\end{aligned} \tag{4.10}$$

The objective is to minimise the distortion outage probability at the FC with the secrecy outage probability at the eavesdropper being less than δ. As $\mathcal{A}_3\left(\mathbb{D},\mathbb{D}_e,\mathbf{G}\right)$ and $\mathcal{A}_4\left(\mathbb{D},\mathbb{D}_e,\mathbf{G}\right)$ are power regions where outage occurs at the FC, and both $D\left(\mathbf{G},\mathbf{p}\left(\mathbf{G}\right)\right)$ and $D_e\left(\mathbf{G},\mathbf{p}\left(\mathbf{G}\right)\right)$ are convex functions over $\mathbf{p}\left(\mathbf{G}\right)$, we can replace the power regions $\mathcal{A}_3\left(\mathbb{D},\mathbb{D}_e,\mathbf{G}\right)$ and $\mathcal{A}_4\left(\mathbb{D},\mathbb{D}_e,\mathbf{G}\right)$ by a region where all the powers are set to $\mathbf{0}$, which saves transmit power and does not violate the constraints (4.7a) and (4.7b), We denote this new feasible probabilistic power scheme as $\hat{\mathbf{P}}\left(\mathbf{G}\right)$, which has three non-overlapping power regions for a given $\mathbf{G}$, namely,

$$\begin{aligned}
\mathcal{B}_1\left(\mathbb{D},\mathbb{D}_e,\mathbf{G}\right) &= \mathcal{A}_1\left(\mathbb{D},\mathbb{D}_e,\mathbf{G}\right), \\
\mathcal{B}_2\left(\mathbb{D},\mathbb{D}_e,\mathbf{G}\right) &= \mathcal{A}_2\left(\mathbb{D},\mathbb{D}_e,\mathbf{G}\right),
\end{aligned}$$

$$\mathcal{B}_3\left(\mathbb{D}, \mathbb{D}_e, \mathbf{G}\right) = \{\mathbf{0}\} \tag{4.11}$$

with all powers in $\mathcal{B}_3\left(\mathbb{D}, \mathbb{D}_e, \mathbf{G}\right)$ equal to zero.

Any optimal probabilistic power scheme can always be divided into the four non-overlapping regions as defined in (4.10). As $\mathcal{A}_3\left(\mathbb{D}, \mathbb{D}_e, \mathbf{G}\right)$ and $\mathcal{A}_4\left(\mathbb{D}, \mathbb{D}_e, \mathbf{G}\right)$ are two sets of powers that result in outage at the FC, replacing these two regions with $\mathcal{B}_3\left(\mathbb{D}, \mathbb{D}_e, \mathbf{G}\right)$ would not change the distortion outage probability at the FC, but maintains or even reduces the secrecy outage probability at the eavesdropper. Therefore, we conclude that if a probabilistic power allocation policy is the optimal solution of problem (4.7), it can be transformed into the same form as $\hat{\mathbf{P}}\left(\mathbf{G}\right)$.

Next, we construct from $\hat{\mathbf{P}}\left(\mathbf{G}\right)$ another probabilistic power scheme $\mathbf{P}'\left(\mathbf{G}\right)$ which randomises among three deterministic power allocations $\{\mathbf{p}_i\left(\mathbf{G}\right)\}$ with time-sharing factors $\{\omega_i\left(\mathbf{G}\right)\}$, i.e.,

$$\mathbf{P}'\left(\mathbf{G}\right) = \sum_{i=1}^{3} \mathbf{p}_i\left(\mathbf{G}\right) 1\left(X\left(\mathbf{G}\right) = i\right), \tag{4.12}$$

where $X\left(\mathbf{G}\right)$ is defined as

$$X\left(\mathbf{G}\right) = \begin{cases} 1, & \text{with probability } \omega_1\left(\mathbf{G}\right), \\ 2, & \text{with probability } \omega_2\left(\mathbf{G}\right), \\ 3, & \text{with probability } \omega_3\left(\mathbf{G}\right). \end{cases} \tag{4.13}$$

The deterministic power schemes $\{\mathbf{p}_i\left(\mathbf{G}\right)\}$ are defined by averaging the powers in each of the regions (4.11), i.e.,

$$\begin{aligned} \mathbf{p}_1\left(\mathbf{G}\right) &= \mathbb{E}\left[\hat{\mathbf{P}}\left(\mathbf{G}\right) \middle| \mathbf{p}\left(\mathbf{G}\right) \in \mathcal{B}_1\left(\mathbb{D}, \mathbb{D}_e, \mathbf{G}\right), \mathbf{G}\right], \\ \mathbf{p}_2\left(\mathbf{G}\right) &= \mathbb{E}\left[\hat{\mathbf{P}}\left(\mathbf{G}\right) \middle| \mathbf{p}\left(\mathbf{G}\right) \in \mathcal{B}_2\left(\mathbb{D}, \mathbb{D}_e, \mathbf{G}\right), \mathbf{G}\right], \\ \mathbf{p}_3\left(\mathbf{G}\right) &= \mathbb{E}\left[\hat{\mathbf{P}}\left(\mathbf{G}\right) \middle| \mathbf{p}\left(\mathbf{G}\right) \in \mathcal{B}_3\left(\mathbb{D}, \mathbb{D}_e, \mathbf{G}\right), \mathbf{G}\right] = \mathbf{0}. \end{aligned} \tag{4.14}$$

The corresponding weighting functions $\{\omega_i\left(\mathbf{G}\right)\}$ are defined as the probability of using

each deterministic power strategy $\{\mathbf{p}_i(\mathbf{G})\}$, i.e.,

$$\begin{aligned}\omega_1(\mathbf{G}) =& \Pr\left[\mathbf{p}(\mathbf{G}) \in \mathcal{B}_1(\mathbb{D}, \mathbb{D}_e, \mathbf{G}) | \mathbf{G}\right],\\ \omega_2(\mathbf{G}) =& \Pr\left[\mathbf{p}(\mathbf{G}) \in \mathcal{B}_2(\mathbb{D}, \mathbb{D}_e, \mathbf{G}) | \mathbf{G}\right],\\ \omega_3(\mathbf{G}) =& \Pr\left[\mathbf{p}(\mathbf{G}) \in \mathcal{B}_3(\mathbb{D}, \mathbb{D}_e, \mathbf{G}) | \mathbf{G}\right]. \end{aligned} \tag{4.15}$$

Remark: From the definition of the power regions given in (4.11), we know that $\mathcal{B}_1(\mathbb{D}, \mathbb{D}_e, \mathbf{G})$ is a set of transmit powers resulting in non-outage at both the FC and eavesdropper, while $\mathcal{B}_2(\mathbb{D}, \mathbb{D}_e, \mathbf{G})$ is the region resulting in outage at the eavesdropper and non-outage at the FC. In addition, $\mathcal{B}_3(\mathbb{D}, \mathbb{D}_e, \mathbf{G})$ represents the power region leading to outage at the FC and non-outage at the eavesdropper. Given the fact that all powers in $\mathcal{B}_3(\mathbb{D}, \mathbb{D}_e, \mathbf{G})$ are zero, we know that in this case the distortion at both the FC and the eavesdropper has the largest possible value of σ_θ^2. Furthermore, for a given channel state $\mathbf{G}$, if $\mathcal{B}_1(\mathbb{D}, \mathbb{D}_e, \mathbf{G}) = \emptyset$, then we must have $\omega_1(\mathbf{G}) = 0$, as there are no powers in $\mathcal{B}_1(\mathbb{D}, \mathbb{D}_e, \mathbf{G})$ satisfying $D(\mathbf{G}, \mathbf{p}(\mathbf{G})) \leq \mathbb{D}$ and $D_e(\mathbf{G}, \mathbf{p}(\mathbf{G})) \geq \mathbb{D}_e$ simultaneously.

Lemma 1: There exists an optimal solution to problem (4.7) of the form $\mathbf{P}^*(\mathbf{G}) = \sum_{i=1}^{3} \mathbf{p}_i(\mathbf{G}) 1(X(\mathbf{G}) = i)$, where $\{\mathbf{p}_i(\mathbf{G})\}$ and $X(\mathbf{G})$ are respectively defined in (4.14) and (4.13), and

- $\omega_1(\mathbf{G}) D_e(\mathbf{G}, \mathbf{p}_1(\mathbf{G})) + \omega_3(\mathbf{G}) D_e(\mathbf{G}, \mathbf{p}_3(\mathbf{G})) - (\omega_1(\mathbf{G}) + \omega_3(\mathbf{G})) \mathbb{D}_e \geq 0$,
- $\omega_1(\mathbf{G}) D(\mathbf{G}, \mathbf{p}_1(\mathbf{G})) + \omega_2(\mathbf{G}) D(\mathbf{G}, \mathbf{p}_2(\mathbf{G})) - (\omega_1(\mathbf{G}) + \omega_2(\mathbf{G})) \mathbb{D} \leq 0$,
- $\sum_{i=1}^{3} \omega_i(\mathbf{G}) = 1$,
- $\mathbb{E}[\omega_2(\mathbf{G})] \leq \delta$,
- $\mathbb{E}\left[\left\langle \sum_{i=1}^{3} \omega_i(\mathbf{G}) \mathbf{p}_i(\mathbf{G}) \right\rangle\right] \leq \mathcal{P}_{\text{tot}}$.

The proof is given in Appendix 4.7.1.

Applying *Lemma 1*, problem (4.7) can be reformulated into another optimization prob-

lem, shown as:

$$\min_{\{\omega_j(\mathbf{G})\},\{\mathbf{p}_j(\mathbf{G})\}} \quad 1-\mathbb{E}\left[\omega_1(\mathbf{G})+\omega_2(\mathbf{G})\right]$$

$$s.t.\ \ \mathbb{E}\left[\omega_2(\mathbf{G})\right] \le \delta, \tag{4.16a}$$

$$\mathbb{E}\left[\langle\omega_1(\mathbf{G})\,\mathbf{p}_1(\mathbf{G})\rangle + \langle\omega_2(\mathbf{G})\,\mathbf{p}_2(\mathbf{G})\rangle\right] \le \mathcal{P}_{\text{tot}}, \tag{4.16b}$$

$$\omega_1(\mathbf{G})\,D_e(\mathbf{G},\mathbf{p}_1(\mathbf{G})) - \omega_1(\mathbf{G})\,\sigma_\theta^2 + \omega_2(\mathbf{G})\left(\mathbb{D}_e-\sigma_\theta^2\right) \ge \mathbb{D}_e-\sigma_\theta^2, \tag{4.16c}$$

$$\omega_1(\mathbf{G})\,D(\mathbf{G},\mathbf{p}_1(\mathbf{G})) + \omega_2(\mathbf{G})\,D(\mathbf{G},\mathbf{p}_2(\mathbf{G})) - (\omega_1(\mathbf{G})+\omega_2(\mathbf{G}))\,\mathbb{D} \le 0, \tag{4.16d}$$

$$\omega_1(\mathbf{G})+\omega_2(\mathbf{G}) \le 1, \tag{4.16e}$$

$$0 \le \omega_j(\mathbf{G}) \le 1, \qquad j=1,2. \tag{4.16f}$$

The functional optimization problem (4.16) is in general non-convex. Let γ, λ, $\nu_e(\mathbf{G})$, $\nu(\mathbf{G})$, and $s(\mathbf{G})$ denote the nonnegative Lagrange multipliers for the constraints (4.16a)-(4.16e) respectively. The generalised Karush-Kuhn-Tucker (KKT) conditions [63] are:

$$\frac{\partial l(\dots)}{\partial p_{jk}^*(\mathbf{G})} \begin{cases} =0, & p_{jk}^*(\mathbf{G})>0 \\ \ge 0, & p_{jk}^*(\mathbf{G})=0 \end{cases} \qquad k=1,\dots,K \tag{4.17}$$

$$\frac{\partial l(\dots)}{\partial \omega_j^*(\mathbf{G})} \begin{cases} =0, & 0<\omega_j^*(\mathbf{G})<1 \\ \ge 0, & \omega_j^*(\mathbf{G})=0 \\ \le 0, & \omega_j^*(\mathbf{G})=1 \end{cases} \tag{4.18}$$

$$\gamma^*\left(\mathbb{E}\left[\omega_2^*(\mathbf{G})-\delta\right]\right)=0, \qquad \gamma^*\ge 0, \tag{4.19}$$

$$\lambda^*\left(\mathbb{E}\left[\left\langle\sum_{j=1}^{2}\omega_j^*(\mathbf{G})\,\mathbf{p}_j^*(\mathbf{G})\right\rangle\right]-\mathcal{P}_{\text{tot}}\right)=0, \qquad \lambda^*\ge 0, \tag{4.20}$$

$$\nu_e^*(\mathbf{G})\left[\left(\mathbb{D}_e-\sigma_\theta^2\right)(1-\omega_2^*(\mathbf{G})) - \omega_1^*(\mathbf{G})\,D_e(\mathbf{G},\mathbf{p}_1^*(\mathbf{G})) + \omega_1^*(\mathbf{G})\,\sigma_\theta^2\right]=0,$$
$$\nu_e^*(\mathbf{G})\ge 0, \tag{4.21}$$

$$\nu^*(\mathbf{G})\left[\omega_1^*(\mathbf{G})\,D(\mathbf{G},\mathbf{p}_1^*(\mathbf{G})) + \omega_2^*(\mathbf{G})\,D(\mathbf{G},\mathbf{p}_2^*(\mathbf{G})) - (\omega_1^*(\mathbf{G})+\omega_2^*(\mathbf{G}))\,\mathbb{D}\right]=0,$$
$$\nu^*(\mathbf{G})\ge 0, \tag{4.22}$$

$$s^*(\mathbf{G})\left[\omega_1^*(\mathbf{G})+\omega_2^*(\mathbf{G})-1\right]=0, \qquad s^*(\mathbf{G})\ge 0. \tag{4.23}$$

where γ^*, λ^*, $\nu_e^*(\mathbf{G})$, $\nu^*(\mathbf{G})$, $s^*(\mathbf{G})$ are the optimal Lagrange multipliers,

and $\left\{\mathbf{p}_j^*(\mathbf{G})\right\}$ $\left\{\omega_j^*(\mathbf{G})\right\}$ are the optimal primal variables, and $l\left(\gamma,\lambda,\nu_e(\mathbf{G}),\nu(\mathbf{G}),s(\mathbf{G}),\left\{\mathbf{p}_j(\mathbf{G})\right\},\left\{\omega_j(\mathbf{G})\right\}\right)$ is defined as

$$\begin{aligned}
&l\left(\gamma,\lambda,\nu_e(\mathbf{G}),\nu(\mathbf{G}),s(\mathbf{G}),\left\{\mathbf{p}_j(\mathbf{G})\right\},\left\{\omega_j(\mathbf{G})\right\}\right)\\
=&-\sum_{j=1}^{2}\omega_j(\mathbf{G})+\gamma\omega_2(\mathbf{G})+\lambda\left\langle\sum_{j=1}^{2}\omega_j(\mathbf{G})\,\mathbf{p}_j(\mathbf{G})\right\rangle\\
&+\nu_e(\mathbf{G})\left[\omega_1(\mathbf{G})\,\sigma_\theta^2-\omega_1(\mathbf{G})\,D_e(\mathbf{G},\mathbf{p}_1(\mathbf{G}))-\omega_2(\mathbf{G})\left(\mathbb{D}_e-\sigma_\theta^2\right)\right]\\
&+\nu(\mathbf{G})\left[\omega_1(\mathbf{G})\,D(\mathbf{G},\mathbf{p}_1(\mathbf{G}))+\omega_2(\mathbf{G})\,D(\mathbf{G},\mathbf{p}_2(\mathbf{G}))-(\omega_1(\mathbf{G})+\omega_2(\mathbf{G}))\,\mathbb{D}\right]\\
&+s(\mathbf{G})\left[\omega_1(\mathbf{G})+\omega_2(\mathbf{G})\right].
\end{aligned}\tag{4.24}$$

From (4.17), we know that for any nonnegative $p_{1k}^*(\mathbf{G})$ and $p_{2k}^*(\mathbf{G})$, they must satisfy, respectively,

$$\lambda^*\omega_1^*(\mathbf{G})\left(\sigma_{\omega k}^2+\sigma_\theta^2\right)-\nu_e^*(\mathbf{G})\,\omega_1^*(\mathbf{G})\,\frac{\partial D_e(\mathbf{G},\mathbf{p}_1^*(\mathbf{G}))}{\partial p_{1k}^*(\mathbf{G})}+\nu^*(\mathbf{G})\,\omega_1^*(\mathbf{G})\,\frac{\partial D(\mathbf{G},\mathbf{p}_1^*(\mathbf{G}))}{\partial p_{1k}^*(\mathbf{G})}=0,\qquad k=1,\ldots,K,\tag{4.25}$$

and

$$\lambda^*\omega_2^*(\mathbf{G})\left(\sigma_{\omega k}^2+\sigma_\theta^2\right)-\nu^*(\mathbf{G})\,\omega_2^*(\mathbf{G})\,\frac{\partial D(\mathbf{G},\mathbf{p}_2^*(\mathbf{G}))}{\partial p_{2k}^*(\mathbf{G})}=0,\qquad k=1,\ldots,K.\tag{4.26}$$

Furthermore, from (4.21)-(4.24) we can obtain the Lagrangian at the optimal points for each channel state $\mathbf{G}$ as

$$\begin{aligned}
&l\left(\gamma^*,\lambda^*,\nu_e^*(\mathbf{G}),s^*(\mathbf{G}),\left\{\mathbf{p}_j^*(\mathbf{G})\right\},\left\{\omega_j^*(\mathbf{G})\right\}\right)\\
=&-\sum_{j=1}^{2}\omega_j^*(\mathbf{G})+\gamma^*\omega_2^*(\mathbf{G})+\lambda^*\left\langle\sum_{j=1}^{2}\omega_j^*(\mathbf{G})\,\mathbf{p}_j^*(\mathbf{G})\right\rangle-\nu_e^*(\mathbf{G})\left(\mathbb{D}_e-\sigma_\theta^2\right)+s^*(\mathbf{G}),
\end{aligned}\tag{4.27}$$

from which we can obtain

$$\frac{\partial l(\ldots)}{\partial\omega_1^*(\mathbf{G})}=-1+\lambda^*\left\langle\mathbf{p}_1^*(\mathbf{G})\right\rangle,\tag{4.28}$$

and

$$\frac{\partial l\left(\ldots\right)}{\partial \omega_2^*\left(\mathbf{G}\right)} = -1 + \lambda^* \left\langle \mathbf{p}_2^*\left(\mathbf{G}\right)\right\rangle + \gamma^*. \tag{4.29}$$

Note that if the channel distributions of both the eavesdropper and the FC are continuous, then the events $\lambda^* \left\langle \mathbf{p}_1^*\left(\mathbf{G}\right)\right\rangle = 1$ or $\lambda^* \left\langle \mathbf{p}_2^*\left(\mathbf{G}\right)\right\rangle = 1-\gamma^*$ have zero probability. Thus, from condition (4.18) and (4.28)-(4.29) we obtain the following result:

$$\omega_j^*\left(\mathbf{G}\right) = \begin{cases} 1, & \frac{\partial l(\ldots)}{\partial \omega_j^*(\mathbf{G})} \leq 0, \\ 0, & \frac{\partial l(\ldots)}{\partial \omega_j^*(\mathbf{G})} > 0. \end{cases} \quad j = 1,2. \tag{4.30}$$

Remark: From the structure of the power allocation in (4.12) and (4.30), we see that for continuous fading channel distributions, the optimal power allocation policies are deterministic.

Theorem 4.1. *Consider the following optimization problems* (4.31) *and* (4.32)*:*

$$\begin{aligned} \min_{\boldsymbol{p}} \quad & \left\langle \boldsymbol{p}\left(\boldsymbol{G}\right)\right\rangle \\ s.t. \quad & D_e\left(\boldsymbol{G},\boldsymbol{p}\left(\boldsymbol{G}\right)\right) \geq \mathbb{D}_e, \\ & D\left(\boldsymbol{G},\boldsymbol{p}\left(\boldsymbol{G}\right)\right) \leq \mathbb{D}, \end{aligned} \tag{4.31}$$

and

$$\begin{aligned} \min_{\boldsymbol{p}} \quad & \left\langle \boldsymbol{p}\left(\boldsymbol{G}\right)\right\rangle \\ s.t. \; & D\left(\boldsymbol{G},\boldsymbol{p}\left(\boldsymbol{G}\right)\right) = \mathbb{D}, \end{aligned} \tag{4.32}$$

with optimal solutions $\boldsymbol{p}_a^*\left(\boldsymbol{G}\right)$ *and* $\boldsymbol{p}_b^*\left(\boldsymbol{G}\right)$ *respectively. Then a locally optimal solution to problem* (4.16) *is given by:*

$$\boldsymbol{P}^*\left(\boldsymbol{G}\right) = \begin{cases} \boldsymbol{p}_a^*\left(\boldsymbol{G}\right), & \text{if } \omega_1^*\left(\boldsymbol{G}\right) = 1 \\ \boldsymbol{p}_b^*\left(\boldsymbol{G}\right), & \text{if } \omega_2^*\left(\boldsymbol{G}\right) = 1 \text{ and } D_e\left(\boldsymbol{G},\boldsymbol{p}_b^*\left(\boldsymbol{G}\right)\right) < \mathbb{D}_e \\ \boldsymbol{0}, & \text{otherwise.} \end{cases} \tag{4.33}$$

The proof is given in Appendix 4.7.2.

Remark: We may have no feasible solutions for problem (4.31), which corresponds to the channel conditions where there are no power allocations satisfying non-outage at both the FC and the eavesdropper, i.e., $\mathcal{B}_1(\mathbb{D}, \mathbb{D}_e, \mathbf{G}) = \emptyset$. In this case, we have $\omega_1^*(\mathbf{G}) = 0$.

Consider the channel states where problem (4.31) has solution $\mathbf{p}_a^*(\mathbf{G})$, and $\mathbf{p}_b^*(\mathbf{G})$ satisfying $D_e\left(\mathbf{G}, \mathbf{p}_b^*(\mathbf{G})\right) < \mathbb{D}_e$. As both $D(\mathbf{G}, \mathbf{p}(\mathbf{G}))$ and $D_e(\mathbf{G}, \mathbf{p}(\mathbf{G}))$ are convex over $\mathbf{p}(\mathbf{G})$, we obtain $D(\mathbf{G}, \mathbf{p}_a^*(\mathbf{G},)) = \mathbb{D}$ and $D_e(\mathbf{G}, \mathbf{p}_a^*(\mathbf{G})) = \mathbb{D}_e$, and $\langle \mathbf{p}_a^*(\mathbf{G}) \rangle \geq \langle \mathbf{p}_b^*(\mathbf{G}) \rangle$ since problem (4.31) has a smaller feasible region than problem (4.32). As a consequence, there is a trade off between choosing $\mathbf{p}_a^*(\mathbf{G})$ or $\mathbf{p}_b^*(\mathbf{G})$ to transmit at each channel instance; $\mathbf{p}_a^*(\mathbf{G})$ leads to non-outage at both the FC and the eavesdropper, whereas $\mathbf{p}_b^*(\mathbf{G})$ results in outage at the eavesdropper but consumes less power.

Define a non-negative transmit power difference $p_{\text{diff}}(\mathbf{G}) = \langle \mathbf{p}_a^*(\mathbf{G}) \rangle - \langle \mathbf{p}_b^*(\mathbf{G}) \rangle$. We may then further categorise the power transmission policy into two different types, depending on the given secrecy outage probability threshold δ and power budget $\mathcal{P}_{\text{tot}}$.

- When $\lambda^* p_{\text{diff}}(\mathbf{G}) \geq \gamma^*$, we obtain either $\mathbf{P}^*(\mathbf{G}) = \mathbf{p}_b^*(\mathbf{G})$ or $\mathbf{P}^*(\mathbf{G}) = \mathbf{0}$. In these channel states, the transmission policies are chosen to use less transmit power by sacrificing either an outage at the eavesdropper, i.e., to use $\mathbf{p}_b^*(\mathbf{G})$, or not transmit leading to an outage at the FC. By doing this, transmit power can be saved for future 'higher potential' channel states where outage occurs neither at the FC nor at the eavesdropper. Furthermore, when $\gamma^* = 0$, from (4.18) we have $\mathbb{E}[\omega_2^*(\mathbf{G})] \leq \delta$, which indicates that we either have a small total power budget or a loose security requirement at the eavesdropper, i.e., a large δ. Intuitively, the optimal transmit policy under such circumstances should be more energy conservative and aim to meet the maximum acceptable distortion level $\mathbb{D}$ at the FC.

- When $\lambda^* p_{\text{diff}}(\mathbf{G}) < \gamma^*$, which implies that either $\langle \mathbf{p}_a^*(\mathbf{G}) \rangle$ is fairly close to $\langle \mathbf{p}_b^*(\mathbf{G}) \rangle$ or we have a relatively small λ^*, we should have $\mathbf{P}^*(\mathbf{G}) = \mathbf{p}_a^*(\mathbf{G})$ or $\mathbf{P}^*(\mathbf{G}) = \mathbf{0}$. When $p_{\text{diff}}(\mathbf{G})$ is small or the transmit power budget is large, instead of using $\mathbf{p}_b^*(\mathbf{G})$, which would result in outage at the eavesdropper, using $\mathbf{p}_a^*(\mathbf{G})$ guarantees non-outage at both the FC and the eavesdropper. If the incremental power $p_{\text{diff}}(\mathbf{G})$

is too large, the sensors will stop transmitting to save power.

4.2.2 Partial CSI

Due to the practical difficulties in obtaining the full channel information of the eavesdropper, in this subsection we will assume that the FC only has statistical knowledge of the eavesdropper. We first explore the power allocation problem that minimises the distortion outage probability at the FC via the Lagrange multiplier technique. To reduce computational cost we then consider suboptimal power allocation policies.

From the analysis in Section 4.2.1 we notice that the optimal transmit power policies are deterministic if both the FC's and eavesdropper's fading channels have continuous distributions, based on which, in this part of the work we aim to develop deterministic transmit power policies with full knowledge of only the sensor-to-FC channels. Using a similar setup as problem (4.7), the Lagrangian in the partial CSI case can be constructed as

$$l(\mathbf{g},\nu,\lambda) = \int_{\mathbf{g}} \Big[1\left\{D\left(\mathbf{g},\mathbf{p}\left(\mathbf{g}\right)\right) > \mathbb{D}\right\} + \lambda \left\langle \mathbf{p}\left(\mathbf{g}\right)\right\rangle + \nu \int_{\mathbf{g}_e} 1\left\{D_e\left(\mathbf{g}_e,\mathbf{p}\left(\mathbf{g}\right)\right) < \mathbb{D}_e\right\} dF\left(\mathbf{g}_e\right)\Big] dF\left(\mathbf{g}\right), \tag{4.34}$$

where λ and ν are non-negative Lagrange multipliers satisfying the following equations at the optimal point:

$$\begin{aligned} &\lambda^* \left(\mathcal{P}_{\text{tot}} - \mathbb{E}\left[\left\langle \mathbf{p}^*\left(\mathbf{g}\right)\right\rangle\right]\right) = 0, \\ &\nu^* \left(\delta - \Pr\left[D_e\left(\mathbf{g}_e,\mathbf{p}^*(\mathbf{g})\right) < \mathbb{D}_e\right]\right) = 0. \end{aligned} \tag{4.35}$$

To minimise the Lagrangian given in (4.34), we need to find the optimal power allocation for each channel state at the FC such that $1\left\{D\left(\mathbf{g},\mathbf{p}\left(\mathbf{g}\right)\right) > \mathbb{D}\right\} + \lambda \left\langle \mathbf{p}\left(\mathbf{g}\right)\right\rangle + \nu \int_{\mathbf{g}_e} 1\left\{D_e\left(\mathbf{g}_e,\mathbf{p}\left(\mathbf{g}\right)\right) < \mathbb{D}_e\right\} dF\left(\mathbf{g}_e\right)$ is minimised.

Lemma 2: Let $\xi\left(\mathbf{p}\left(\mathbf{g}\right)\right) = \lambda \left\langle \mathbf{p}\left(\mathbf{g}\right)\right\rangle + \nu \int_{\mathbf{g}_e} 1\left\{D_e\left(\mathbf{g}_e,\mathbf{p}\left(\mathbf{g}\right)\right) < \mathbb{D}_e\right\} dF\left(\mathbf{g}_e\right)$. Then the optimal $\mathbf{p}^*\left(\mathbf{g}\right)$ must satisfy $0 \leq 1\left\{D\left(\mathbf{g},\mathbf{p}^*\left(\mathbf{g}\right)\right) > \mathbb{D}\right\} + \xi\left(\mathbf{p}^*\left(\mathbf{g}\right)\right) \leq 1$.

The proof is given in Appendix 4.7.3.

In order to minimise $1\{D(\mathbf{g},\mathbf{p}(\mathbf{g})) > \mathbb{D}\} + \xi(\mathbf{p}(\mathbf{g}))$, we either obtain $D(\mathbf{g},\mathbf{p}^*(\mathbf{g})) > \mathbb{D}$ where we declare an outage at the FC, or the distortion at the FC is no larger than $\mathbb{D}$ and so $1\{D(\mathbf{g},\mathbf{p}^*(\mathbf{g})) > \mathbb{D}\} = 0$. To be more specific,

- When $D(\mathbf{g},\mathbf{p}^*(\mathbf{g})) > \mathbb{D}$, we see that $1\{D(\mathbf{g},\mathbf{p}^*(\mathbf{g})) > \mathbb{D}\} = 1$ indicates an outage at the FC. Furthermore, we must have the optimal power allocation at this channel instance being equal to zero for all sensors, since a non-zero power would result in a nonnegative value of $\lambda\langle\mathbf{p}^*(\mathbf{g})\rangle + \nu\int 1\{D_e(\mathbf{g}_e,\mathbf{p}^*(\mathbf{g})) < \mathbb{D}_e\} f(\mathbf{g}_e)\, d\mathbf{g}_e$. Intuitively, knowing that an outage will happen at the FC, the sensors would stop transmitting to save power and to reduce the possibility of information being leaked to the eavesdropper.

- When $D(\mathbf{g},\mathbf{p}^*(\mathbf{g})) \leq \mathbb{D}$, which implies non-outage at the FC. In this situation, either the FC has relatively good channel conditions that a small amount of power would secure non-outage at the FC, or the constraints are quite loose (i.e. a large power budget and/or a loose security requirement at the eavesdropper).

Therefore, for a given channel state at the FC, the sensors either choose to forward the information to the FC (with non-outage at the FC achieved) or keep silent. Hence, by applying Lemma 2 we obtain that the optimal power allocation $\mathbf{p}^*(\mathbf{g})$ has the form

$$\mathbf{p}^*(\mathbf{g}) = \begin{cases} \hat{\mathbf{p}}(\mathbf{g}), & \text{if } \xi(\hat{\mathbf{p}}(\mathbf{g})) < 1 \\ \mathbf{0}, & \text{otherwise,} \end{cases} \tag{4.36}$$

where $\hat{\mathbf{p}}(\mathbf{g})$ is a locally optimal solution of the following problem:

$$\begin{aligned} &\min_{\mathbf{p}(\mathbf{g})} \ \lambda\langle\mathbf{p}(\mathbf{g})\rangle + \nu\int 1\{D_e(\mathbf{g}_e,\mathbf{p}(\mathbf{g})) < \mathbb{D}_e\} f(\mathbf{g}_e)\, d\mathbf{g}_e \\ &s.t. \ \ D(\mathbf{g},\mathbf{p}(\mathbf{g})) \leq \mathbb{D}. \end{aligned} \tag{4.37}$$

Partial CSI Suboptimal Solution

Due to the difficulties of explicitly expressing $\int 1\{D_e(\mathbf{g}_e,\mathbf{p}(\mathbf{g})) < \mathbb{D}_e\} f(\mathbf{g}_e)\, d\mathbf{g}_e$ and deriving a locally optimal solution to problem

(4.37), which has high computational costs, in this part we will look at a suboptimal power allocation scheme based on sensor scheduling.

In a multiple-sensor system, instead of activating all the sensors, we can selectively choose one sensor to forward its measurement to the FC. This may be useful in scenarios where bandwidth is at a premium or there are very strict interference constraints. Let $g_m = \max(g_1, \ldots, g_K)$ where m corresponds to the index of the sensor with the largest channel gain, and g_{e_m} be the corresponding channel power gain from sensor m to the eavesdropper. One possible sensor scheduling policy[4] is that only the sensor with the best channel transmits. The distortion at the FC and the eavesdropper then become:

$$D = \left(\frac{1}{\sigma_\theta^2} + \frac{g_m p_m}{g_m p_m \sigma_{\omega m}^2 + \sigma_{n m}^2} \right)^{-1}, \tag{4.38a}$$

$$D_e = \left(\frac{1}{\sigma_\theta^2} + \frac{g_{e_m} p_m}{g_{e_m} p_m \sigma_{\omega m}^2 + \sigma_{e\, m}^2} \right)^{-1}. \tag{4.38b}$$

To explicitly illustrate the power policies in this scheme, we will assume that the channel power gains are exponentially distributed at both the FC and the eavesdropper with means $\hat{\lambda}$ and $\hat{\lambda}_e$ respectively. We can then obtain the probability density function of g_m as $\frac{K}{\hat{\lambda}} \left(1 - e^{-\frac{g_m}{\hat{\lambda}}}\right)^{K-1} e^{-\frac{g_m}{\hat{\lambda}}}$.

Following similar techniques as in Section 4.2.2, problem (4.37) is then reduced to

$$\begin{aligned} \min_{p(g_m)} \ & \lambda \langle p(g_m) \rangle + \nu \int 1\{D_e(g_{e_m}, p(g_m)) < \mathbb{D}_e\} \, dF(g_{e_m}) \\ s.t. \ & D(g_m, p(g_m)) \leq \mathbb{D}, \end{aligned} \tag{4.39}$$

from which we can then compute the optimal solution as

$$\hat{p}(g_m) = \frac{\sigma_{n m}^2 \left(\sigma_\theta^2 - \mathbb{D}\right)}{\mathbb{D}\left(\sigma_\theta^2 + \sigma_{\omega m}^2\right) - \sigma_\theta^2 \sigma_{\omega m}^2} \frac{1}{g_m}. \tag{4.40}$$

Knowing that the eavesdropper's channel is exponentially distributed, we can derive the

[4]From the distortion expression at equation (4.5), it would be more intuitive to pick the sensor with the best measurement sensitivity and best channel SNR at the FC; however, in order derive analytical results, only the sensor with the best channel condition is chosen to transmit.

outage probability at the eavesdropper for a given FC channel state as

$$\begin{aligned}
&\Pr\left[D_e\left(g_{e_m}, \hat{p}\left(g_m\right)\right) < \mathbb{D}_e \middle| g_m\right] \\
=&1 - \Pr\left[g_{e_m} \leq \frac{\sigma_{e\,m}^2\left(\sigma_\theta^2 - \mathbb{D}_e\right)}{\mathbb{D}_e\sigma_{\omega m}^2 - \sigma_\theta^2\left(\sigma_{\omega m}^2 - \mathbb{D}_e\right)} \frac{1}{\hat{p}\left(g_m\right)} \middle| g_m\right] \\
=&e^{-\frac{D_{\text{th}}}{\lambda_e} g_m},
\end{aligned} \tag{4.41}$$

where $D_{\text{th}} = \frac{\sigma_{e\,m}^2}{\sigma_{n\,m}^2} \frac{\left(\sigma_\theta^2 - \mathbb{D}_e\right)\left[\mathbb{D}\left(\sigma_\theta^2 + \sigma_{\omega m}^2\right) - \sigma_\theta^2\sigma_{\omega m}^2\right]}{\left(\sigma_\theta^2 - \mathbb{D}\right)\left[\mathbb{D}_e\left(\sigma_\theta^2 + \sigma_{\omega m}^2\right) - \sigma_\theta^2\sigma_{\omega m}^2\right]}$.

Combining the results of (4.36), (4.40) and (4.41) we obtain the transmit power policy:

$$p^*\left(g_m\right) = \begin{cases} \frac{\sigma_{n\,m}^2\left(\sigma_\theta^2 - \mathbb{D}\right)}{\mathbb{D}\left(\sigma_\theta^2 + \sigma_{\omega m}^2\right) - \sigma_\theta^2\sigma_{\omega m}^2} \frac{1}{g_m}, & \text{if } g_m > g_{m\_th}, \\ 0, & \text{otherwise}, \end{cases} \tag{4.42}$$

where g_{m_th} satisfies $\nu^* e^{-\frac{D_{\text{th}}}{\lambda_e} g_{m\_th}} + \frac{P_{\text{t}}\lambda^*}{g_{m\_th}} = 1$, with $P_{\text{t}} = \frac{\sigma_{n\,m}^2\left(\sigma_\theta^2 - \mathbb{D}\right)}{\mathbb{D} - \frac{\sigma_\theta^2\sigma_{\omega m}^2}{\sigma_\theta^2 + \sigma_{\omega m}^2}}$, and with λ^* and ν^* being the optimal Lagrange multipliers chosen to satisfy the power constraint and secrecy outage constraint at the eavesdropper.

Notice that as g_m is continuous and $\nu^* e^{-\frac{D_{\text{th}}}{\lambda_e} g_m} + \frac{P_{\text{t}}\lambda^*}{g_m}$ is monotonic decreasing with g_m, we obtain the 'on-off' transmit power policy in (4.42), where if $g_m > g_{m\_th}$ the sensor uses $\hat{p}\left(g_m\right)$ to transmit with non-outage at the FC achieved, and the sensor does not transmit when $g_m \leq g_{m\_th}$ which leads to an outage to occur at the FC. In addition, the overall outage probability at the FC can be expressed as

$$\begin{aligned}
&\Pr\left[D\left(g_m, p^*\left(g_m\right)\right) > \mathbb{D}\right] \\
=&\frac{K}{\hat{\lambda}} \int_0^{g_{m_th}} \left(1 - e^{-\frac{g_m}{\hat{\lambda}}}\right)^{K-1} e^{-\frac{g_m}{\hat{\lambda}}} dg_m \\
=&\left(1 - e^{-\frac{g_{m_th}}{\hat{\lambda}}}\right)^K.
\end{aligned} \tag{4.43}$$

From (4.41) and (4.42), which are two monotonic decreasing functions with respect to g_m, we obtain that $g_{m\_th}\left(\lambda^*, \nu^*\right)$ must satisfy either $\int_{g_{m\_th}(\lambda^*,\nu^*)}^{\infty} e^{-\frac{D_{\text{th}}}{\lambda_e} g_m} f(g_m) dg_m = \delta$ or $\frac{\sigma_{n\,m}^2\left(\sigma_\theta^2 - \mathbb{D}\right)}{\mathbb{D}\left(\sigma_\theta^2 + \sigma_{\omega m}^2\right) - \sigma_\theta^2\sigma_{\omega m}^2} \int_{g_{m\_th}(\lambda^*,\nu^*)}^{\infty} \frac{1}{g_m} f(g_m) dg_m = \mathcal{P}_{\text{tot}}$, where $f(g_m) = \frac{K}{\hat{\lambda}}\left(1 - e^{-\frac{g_m}{\hat{\lambda}}}\right)^{K-1} e^{-\frac{g_m}{\hat{\lambda}}}$. This is because for a given total power budget and outage probability threshold at the

eavesdropper, there is zero probability of finding a g_{m_th} to meet both constraints with equality. From the KKT conditions for the optimal points we then derive that either

$$P_t \int_{P_t\lambda^*}^{\infty} \frac{1}{g_m} f(g_m) dg_m = \mathcal{P}_{\text{tot}}, \;\; \nu^* = 0, \tag{4.44}$$

or

$$\int_{\frac{\lambda_e log \nu^*}{D_{\text{th}}}}^{\infty} e^{-\frac{D_{\text{th}}}{\lambda_e} g_m} f(g_m) dg_m = \delta, \;\; \lambda^* = 0. \tag{4.45}$$

4.3 Single Sensor with Multiple Antennas Scenario

In order to compare with the multiple-sensor scenario as well as for analytical tractability, in this part of work we consider a situation where only one sensor with multiple-antenna is in the network observing the source.[5] In this scenario, similar performance gains as in having multiple sensors can be achieved. In fact, with multiple antennas additional techniques can be used to further enhance the system performance.

A schematic diagram is shown in Figure 4.2. We assume that the same single point Gaussian source θ as defined in Section 4.2 is observed by a sensor with N_t transmit antennas, which employs the analog amplify and forward technique to scale the observed signal with a complex vector $\boldsymbol{\beta} \in \mathbb{C}^{N_t \times 1}$, before sending it to the FC via a set of complex fading channels $\mathbf{H} \in \mathbb{C}^{Nr \times Nt}$. The observed signal x is also listened to by the eavesdropper after passing through another set of channels $\mathbf{H}_e \in \mathbb{C}^{Ne \times Nt}$, where we assume that the FC and the eavesdropper are equipped with N_r and N_e receive antennas respectively.

The signals received by the FC and the eavesdropper are, respectively,

$$\mathbf{y} = \mathbf{H}\boldsymbol{\beta}\theta + \mathbf{H}\boldsymbol{\beta}\omega + \mathbf{z}, \tag{4.46a}$$

$$\mathbf{y}_e = \mathbf{H}_e\boldsymbol{\beta}\theta + \mathbf{H}_e\boldsymbol{\beta}\omega + \mathbf{z}_e, \tag{4.46b}$$

where both $\mathbf{z} \in \mathbb{C}^{Nr \times 1}$ and $\mathbf{z}_e \in \mathbb{C}^{Ne \times 1}$ are complex Gaussian channel noise at the FC and

[5] Multiple-sensor-multiple-antenna (MSMA) is not studied in the work due to its complexity involving non-convex optimization problems and will be investigated in future work. However, techniques that are explored in the partial CSI case of this section can be also implemented in MSMA case.

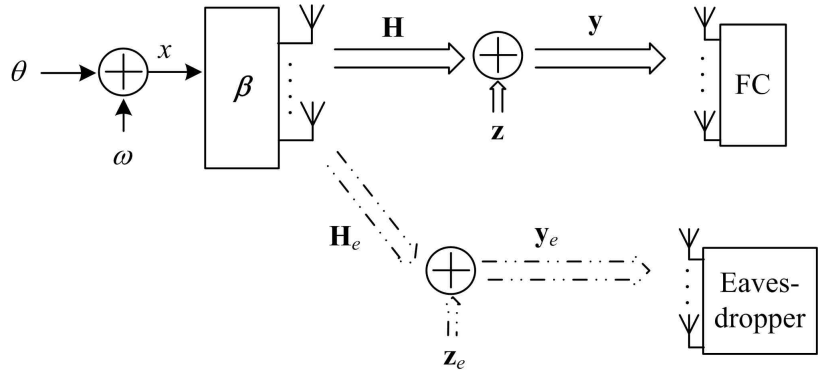

Figure 4.2: Diagram of a multiple-antenna single sensor scenario with the presence of an eavesdropper.

the eavesdropper with covariance $\sigma_n^2\mathbf{I}_{N_r}$ and $\sigma_e^2\mathbf{I}_{N_e}$ respectively.

The optimal linear minimum mean square error estimator, is used at both the FC and the adversary to measure θ. For a given channel instance, the distortion D at the FC can be obtained as

$$\begin{aligned} D &= \left(\frac{1}{\sigma_\theta^2} + (\mathbf{H}\boldsymbol{\beta})^{\mathrm{H}}\Sigma^{-1}\mathbf{H}\boldsymbol{\beta}\right)^{-1}, \\ &\overset{(b)}{=}\sigma_\theta^2\left(1-(\mathbf{H}\boldsymbol{\beta})^{\mathrm{H}}\left[\mathbf{H}\boldsymbol{\beta}\,(\mathbf{H}\boldsymbol{\beta})^{\mathrm{H}}\left(\sigma_\theta^2+\sigma_\omega^2\right)+\sigma_n^2\mathbf{I}_{N_r}\right]^{-1}\mathbf{H}\boldsymbol{\beta}\sigma_\theta^2\right), \\ &\overset{(c)}{=}\sigma_\theta^2\left[1-\frac{\sigma_\theta^2}{\sigma_\theta^2+\sigma_\omega^2}\left(1-\frac{\alpha}{\alpha+(\mathbf{H}\boldsymbol{\beta})^{\mathrm{H}}\mathbf{H}\boldsymbol{\beta}}\right)\right], \end{aligned} \tag{4.47}$$

where $\Sigma \triangleq \mathbf{H}\boldsymbol{\beta}\,(\mathbf{H}\boldsymbol{\beta})^{\mathrm{H}}\sigma_\omega^2+\sigma_n^2\mathbf{I}_{N_r}$ is the covariance matrix of $\mathbf{H}\boldsymbol{\beta}\omega+\mathbf{z}$, $\alpha \triangleq \frac{\sigma_n^2}{\sigma_\omega^2+\sigma_\theta^2}$, and (b)-(c) result from applying the Matrix Inversion Lemma [34]. Similarly, the mean squared error or *distortion* at the eavesdropper is given as

$$D_e=\sigma_\theta^2\left[1-\frac{\sigma_\theta^2}{\sigma_\theta^2+\sigma_\omega^2}\left(1-\frac{\alpha_e}{\alpha_e+(\mathbf{H}_e\boldsymbol{\beta})^{\mathrm{H}}\mathbf{H}_e\boldsymbol{\beta}}\right)\right], \tag{4.48}$$

where $\alpha_e \triangleq \frac{\sigma_e^2}{\sigma_\omega^2+\sigma_\theta^2}$.

We set a maximum acceptable distortion level $\mathbb{D}$ at the FC and define *the distortion*

outage probability as

$$\mathrm{Pr}_{\text{outage\_FC}} = \mathrm{Pr}\left[D > \mathbb{D}\right] = \mathrm{Pr}\left[\frac{1}{\left(\mathbf{H}\boldsymbol{\beta}\right)^{\mathrm{H}}\mathbf{H}\boldsymbol{\beta}} > \mathbb{S}\right], \tag{4.49}$$

where $\mathbb{S} \triangleq \frac{\sigma_\omega^2+\sigma_\theta^2}{\sigma_n^2}\left[\frac{\sigma_n^4}{\sigma_n^4-\mathbb{D}\left(\sigma_\omega^2+\sigma_\theta^2\right)+\sigma_\omega^2\sigma_\theta^2}-1\right]$. We also set a minimum acceptable distortion level $\mathbb{D}_e$ at the eavesdropper and assume that if $D_e < \mathbb{D}_e$ the measurement information at this channel instance can be successfully retrieved by the eavesdropper leading to a security breach. Letting $\mathbb{S}_e \triangleq \frac{\sigma_\omega^2+\sigma_\theta^2}{\sigma_e^2}\left[\frac{\sigma_e^4}{\sigma_e^4-\mathbb{D}_e\left(\sigma_\omega^2+\sigma_\theta^2\right)+\sigma_\omega^2\sigma_\theta^2}-1\right]$, the *secrecy outage probability* at the eavesdropper can be expressed as

$$\mathrm{Pr}_{\text{outage\_EVE}} = \mathrm{Pr}\left[D_e < \mathbb{D}_e\right] = \mathrm{Pr}\left[\frac{1}{\left(\mathbf{H}_e\boldsymbol{\beta}\right)^{\mathrm{H}}\mathbf{H}_e\boldsymbol{\beta}} < \mathbb{S}_e\right].$$

With a given power budget at the sensor, our objective is to minimise the distortion outage probability at the FC, while keeping the secrecy outage probability at the eavesdropper below δ. Hence the optimization problem can be cast as:

$$\begin{aligned} \min_{\boldsymbol{\beta}} \;\; & \mathrm{Pr}\left[\frac{1}{\left(\mathbf{H}\boldsymbol{\beta}\right)^{\mathrm{H}}\mathbf{H}\boldsymbol{\beta}} > \mathbb{S}\right] \\ s.t. \;\; & \mathrm{Pr}\left[\frac{1}{\left(\mathbf{H}_e\boldsymbol{\beta}\right)^{\mathrm{H}}\mathbf{H}_e\boldsymbol{\beta}} < \mathbb{S}_e\right] \le \delta, \quad \mathbb{E}\left[\boldsymbol{\beta}^{\mathrm{H}}\boldsymbol{\beta}\right] \le \frac{\mathcal{P}_{\text{tot}}}{\sigma_\theta^2+\sigma_\omega^2}. \end{aligned} \tag{4.50}$$

In the following, we focus on the full CSI scenario where both the FC and eavesdropper's channel information are available, and the partial CSI scenario where we assume only the FC's channel states are perfectly known. In both scenarios, we first focus on finding the best $\boldsymbol{\beta}$ that minimises the objective while satisfying all the constraints. We then consider other techniques that can be used in the multiple-antenna systems to further enhance the performance.

4.3.1 Full CSI

With full knowledge of the eavesdropper's channel information, problem (4.50) can be solved using similar techniques as in Section 4.2.1, where we start from an arbitrary fea-

sible probabilistic power allocation scheme, from which it can be used to construct another feasible power allocation scheme that provides no worse performance, and based on which we construct three deterministic schemes $\boldsymbol{\beta}_1$, $\boldsymbol{\beta}_2$, and $\boldsymbol{\beta}_3 = \mathbf{0}$. We then show that the optimal $\boldsymbol{\beta}^*$, which is a function of $\mathbf{H}$ and $\mathbf{H}_e$, can be found by considering a probabilistic power allocation scheme that randomises among the three deterministic schemes $\boldsymbol{\beta}_1$, $\boldsymbol{\beta}_2$, and $\boldsymbol{\beta}_3$ with corresponding weighting factors $\{\omega_i\}_{i=1}^3$. The problem (4.50) then becomes

$$\begin{aligned}
\min_{\left\{\boldsymbol{\beta}_j(\mathbf{H},\mathbf{H}_e)\right\},\,\left\{\omega_j(\mathbf{H},\mathbf{H}_e)\right\}} \quad & 1-\mathbb{E}\left[\omega_1+\omega_2\right] \\
s.t.\ \ & \mathbb{E}\left[\omega_2\right] \leq \delta, && \text{(4.51a)}\\
& \mathbb{E}\left[\sum_{j=1}^{2}\omega_j\boldsymbol{\beta}_j^{\mathrm{H}}\boldsymbol{\beta}_j\right] \leq \frac{\mathcal{P}_{\mathrm{tot}}}{\sigma_\theta^2+\sigma_\omega^2}, && \text{(4.51b)}\\
& \frac{\omega_1}{\left(\mathbf{H}_e\boldsymbol{\beta}_1\right)^{\mathrm{H}}\mathbf{H}_e\boldsymbol{\beta}_1} \geq \left(1-\omega_2\right)\mathbb{S}_e, && \text{(4.51c)}\\
& \frac{\omega_1}{\left(\mathbf{H}\boldsymbol{\beta}_1\right)^{\mathrm{H}}\mathbf{H}\boldsymbol{\beta}_1}+\frac{\omega_2}{\left(\mathbf{H}\boldsymbol{\beta}_2\right)^{\mathrm{H}}\mathbf{H}\boldsymbol{\beta}_2} \leq \left(\omega_1+\omega_2\right)\mathbb{S}, && \text{(4.51d)}\\
& 0 \leq \omega_1+\omega_2 \leq 1, \quad 0 \leq \omega_1,\omega_2 \leq 1, && \text{(4.51e)}
\end{aligned}$$

where the derivation is similar to that of problem (4.16) and is thus omitted to avoid repetition. As problem (4.51) is again a non-convex problem the result we derive is a locally optimal solution. With the assumption that both $\mathbf{H}$ and $\mathbf{H}_e$ are continuously distributed, the solution of problem (4.50) in the case of full CSI is given as

$$\boldsymbol{\beta}^*\left(\mathbf{H},\mathbf{H}_e\right) = \begin{cases} \boldsymbol{\beta}_1^*\left(\mathbf{H},\mathbf{H}_e\right), & \text{if } \lambda^*\boldsymbol{\beta}_1^{*\mathrm{H}}\boldsymbol{\beta}_1^* \leq 1, \\ \boldsymbol{\beta}_2^*\left(\mathbf{H},\mathbf{H}_e\right), & \text{if } \lambda^*\boldsymbol{\beta}_2^{*\mathrm{H}}\boldsymbol{\beta}_2^* \leq 1-\gamma^* \text{ and } \left(\mathbf{H}_e\boldsymbol{\beta}_2^*\right)^{\mathrm{H}}\mathbf{H}_e\boldsymbol{\beta}_2^* > \mathbb{S}_e^{-1}, \\ \mathbf{0}, & \text{otherwise}, \end{cases} \qquad \text{(4.52)}$$

where λ^* and $\gamma^*\left(\mathbf{H},\mathbf{H}_e\right)$ are the optimal Lagrange multipliers chosen to satisfy the constraints $\mathbb{E}\left[\sum_{j=1}^{2}\omega_j^*\boldsymbol{\beta}_j^{\mathrm{H}*}\boldsymbol{\beta}_j^*\right] \leq \frac{\mathcal{P}_{\mathrm{tot}}}{\sigma_\theta^2+\sigma_\omega^2}$ and $\mathbb{E}\left[\omega_2^*\right] \leq \delta$ respectively; and $\boldsymbol{\beta}_1^*$ and $\boldsymbol{\beta}_2^*$ are re-

spectively the optimal solutions of the following two problems:

$$\begin{aligned} \min_{\boldsymbol{\beta}_1(\mathbf{H},\,\mathbf{H}_e)} \ & \boldsymbol{\beta}_1^{\mathrm{H}}\boldsymbol{\beta}_1 \\ s.t. \ & (\mathbf{H}\boldsymbol{\beta}_1)^{\mathrm{H}}\mathbf{H}\boldsymbol{\beta}_1 \geq \mathbb{S}^{-1}, \\ & (\mathbf{H}_e\boldsymbol{\beta}_1)^{\mathrm{H}}\mathbf{H}_e\boldsymbol{\beta}_1 \leq \mathbb{S}_e^{-1}; \end{aligned} \tag{4.53}$$

and

$$\begin{aligned} \min_{\boldsymbol{\beta}_2(\mathbf{H})} \ & \boldsymbol{\beta}_2^{\mathrm{H}}\boldsymbol{\beta}_2 \\ s.t. \ & (\mathbf{H}\boldsymbol{\beta}_2)^{\mathrm{H}}\mathbf{H}\boldsymbol{\beta}_2 \geq \mathbb{S}^{-1}. \end{aligned} \tag{4.54}$$

Note that as all the constraints and objective functions in problem (4.51), (4.53), (4.54) are real-valued over the complex field, we need to consider both the real and imaginary parts when applying the KKT conditions for the optimal points [70, 94]. Furthermore, because problem (4.53) is a non-convex optimization problem, while (4.54) is a convex problem; we obtain $\boldsymbol{\beta}_1^*$ being a locally optimal solution of the problem (4.53) and $\boldsymbol{\beta}_2^*$ being the globally optimal solution of the problem (4.54).

Zero Outage Probability at the Eavesdropper

If the sensor has more transmit antennas than the number of receive antennas at the eavesdropper, i.e., $N_t > N_e$, then it can transmit the observation signal x onto the null space of the eavesdropper's channel, thus leaking no useful information to the eavesdropper. To be more specific, let the singular value decomposition of $\mathbf{H}_e$ be $\mathbf{H}_e = \mathbf{U}\mathbf{S}\mathbf{V}^{\mathrm{H}}$. Then we can express the eavesdropper's channel null space as $\tilde{\mathbf{V}}\tilde{\mathbf{V}}^{\mathrm{H}}$, where $\tilde{\mathbf{V}}$ contains the last $N_t - N_e$ columns of $\mathbf{V}$ [108]. Define a precoding matrix $\mathbf{W} = \tilde{\mathbf{V}}\tilde{\mathbf{V}}^{\mathrm{H}} \in \mathbb{C}^{N_t}$. The signals received by the FC and the eavesdropper are then given by, respectively,

$$\mathbf{y} = \mathbf{H}\mathbf{W}\boldsymbol{\beta}\theta + \mathbf{H}\mathbf{W}\boldsymbol{\beta}\omega + \mathbf{z}, \tag{4.55}$$

$$\mathbf{y}_e = \mathbf{H}_e\mathbf{W}\boldsymbol{\beta}\theta + \mathbf{H}_e\mathbf{W}\boldsymbol{\beta}\omega + \mathbf{z}_e = \mathbf{z}_e. \tag{4.56}$$

On the eavesdropper side, as no information about x is received, we obtain the secrecy outage probability $\mathrm{Pr}_{\text{outage\_EVE}} = 0$.

The outage minimization problem can then be given as

$$\begin{aligned} \min_{\boldsymbol{\beta}(\mathbf{H})} \ & \mathrm{Pr}\left[\frac{1}{(\mathbf{HW}\boldsymbol{\beta})^{\mathrm{H}}\mathbf{HW}\boldsymbol{\beta}} > \mathrm{S}\right] \\ s.t. \ & \mathbb{E}\left[\boldsymbol{\beta}^{\mathrm{H}}\boldsymbol{\beta}\right] \leq \frac{\mathcal{P}_{\text{tot}}}{\sigma_\theta^2+\sigma_\omega^2}. \end{aligned} \tag{4.57}$$

Similar techniques as used in Section 4.2.1 can be employed to solve problem (4.57), and it can be shown that the globally optimal $\boldsymbol{\beta}^*$ is constructed by randomizing among two deterministic power schemes $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2 = \mathbf{0}$ with corresponding weighting factors ω and $1-\omega$. Furthermore, problem (4.57) can be reformulated into the following problem:

$$\begin{aligned} \min_{\boldsymbol{\beta}_1(\mathbf{H}),\, \omega(\mathbf{H})} \ & 1 - \mathbb{E}\left[\omega\right] \\ s.t. \ & \mathbb{E}\left[\omega\boldsymbol{\beta}_1^{\mathrm{H}}\boldsymbol{\beta}_1\right] \leq \frac{\mathcal{P}_{\text{tot}}}{\sigma_\theta^2+\sigma_\omega^2} && \text{(4.58a)} \\ & (\mathbf{HW}\boldsymbol{\beta}_1)^{\mathrm{H}}\mathbf{HW}\boldsymbol{\beta}_1 \geq \mathrm{S}^{-1}. && \text{(4.58b)} \end{aligned}$$

The solution is given as

$$\boldsymbol{\beta}^*(\mathbf{H}) = \begin{cases} \boldsymbol{\beta}_1^*(\mathbf{H}), & \text{if } \lambda^*\left(\sigma_\theta^2+\sigma_\omega^2\right)\boldsymbol{\beta}_1^{\mathrm{H}^*}\boldsymbol{\beta}_1^* < 1 \\ \mathbf{0}, & \text{otherwise,} \end{cases} \tag{4.59}$$

where λ^* is the optimal Lagrange multiplier associated with the power constraint (4.58a) which is obtained numerically, and $\boldsymbol{\beta}_1^*$ is the globally optimal solution of the problem:

$$\begin{aligned} \min_{\boldsymbol{\beta}_1(\mathbf{H})} \ & \left(\sigma_\theta^2+\sigma_\omega^2\right)\boldsymbol{\beta}_1^{\mathrm{H}}\boldsymbol{\beta}_1, \\ s.t. \ & (\mathbf{HW}\boldsymbol{\beta}_1)^{\mathrm{H}}\mathbf{HW}\boldsymbol{\beta}_1 \leq \mathrm{S}^{-1}. \end{aligned}$$

Remark: With this scheme, the FC's effective channel is $\mathbf{HW}$, which is the projection of $\mathbf{H}$ onto the null space of $\mathbf{H}_e$ via the precoding matrix $\mathbf{W}$. Moreover, if the FC has only one receive antenna, i.e., $N_r = 1$, we obtain the beamforming vector $\boldsymbol{\beta}_1^*(\mathbf{H}) =$

$\sqrt{\frac{\mathbb{S}^{-1}}{(\mathbf{HW})^{\mathrm{H}}\mathbf{HW}}}\frac{(\mathbf{HW})^{\mathrm{H}}}{\|\mathbf{HW}\|}$ (where the notation $\|\mathbf{x}\|$ refers to the Euclidean norm of the vector $\mathbf{x}$), which lines up with the effective channel $\mathbf{HW}$ while satisfying the power constraint.[6]

4.3.2 Partial CSI

In this part of the work, we consider a case where the FC can acquire its channel information but only has statistical knowledge of the eavesdropper's. From the full CSI case, we know that a deterministic power allocation is optimal for continuously distributed fading channels. Therefore, applying the results derived in Section 4.2.2, we can obtain a locally optimal $\boldsymbol{\beta}^*$ at each FC channel instance as:

$$\boldsymbol{\beta}^*(\mathbf{H}) = \begin{cases} \hat{\boldsymbol{\beta}}(\mathbf{H}), & \text{if } \nu(\mathbf{H}) \int_{\mathbf{H}_e} 1\left\{\left(\mathbf{H}_e\hat{\boldsymbol{\beta}}\right)^{\mathrm{H}}\mathbf{H}_e\hat{\boldsymbol{\beta}} > \frac{1}{\mathbb{S}_e}\right\} dF(\mathbf{H}_e) + \lambda\hat{\boldsymbol{\beta}}^{\mathrm{H}}\hat{\boldsymbol{\beta}} < 1 \\ \mathbf{0}, & \text{otherwise,} \end{cases}$$

where $\hat{\boldsymbol{\beta}}$ is a locally optimal solution to the problem:

$$\begin{aligned} &\min_{\boldsymbol{\beta}(\mathbf{H})} \ \lambda\boldsymbol{\beta}^{\mathrm{H}}\boldsymbol{\beta} + \nu(\mathbf{H}) \int_{\mathbf{H}_e} 1\left\{(\mathbf{H}_e\boldsymbol{\beta})^{\mathrm{H}}\mathbf{H}_e\boldsymbol{\beta} > \mathbb{S}_e^{-1}\right\} dF(\mathbf{H}_e) \\ &s.t. \ \ (\mathbf{H}\boldsymbol{\beta})^{\mathrm{H}}\mathbf{H}\boldsymbol{\beta} \geq \mathbb{S}^{-1}, \end{aligned} \tag{4.60}$$

with λ and $\nu(\mathbf{H})$ being nonnegative Lagrange multipliers corresponding to respectively the power constraint and the secrecy outage constraint at the eavesdropper.

Artificial Noise

Assuming that the sensor is equipped with more transmit antennas than the number of receive antennas at the FC, we can employ the technique of artificial noise [27, 74] to enhance the system performance. The idea is to increase the noise level seen by the adversary in a way that its channel is degraded while the channel of the legitimate receiver is not. With this method, the artificial noise is generated by the sensor and transmitted onto the null space of the FC, thus it does not impact the message received by the FC but

[6] One could also use the techniques in [64] to solve the problem, which will give the same result.

increase the noise level at the eavesdropper.

Let $[\mathbf{W}_1, \mathbf{W}_2]$ be an orthonormal basis of $\mathbb{C}^{N_t}$ with $\mathbf{W}_1 \in \mathbb{C}^{N_t \times N_r}$ and $\mathbf{W}_2 \in \mathbb{C}^{N_t \times (N_t - N_r)}$ representing respectively the signal space and the null space of $\mathbf{H}$. The signals received by the FC and the eavesdropper are, respectively,

$$\begin{aligned} \mathbf{y} &= \mathbf{H}\mathbf{W}_1\boldsymbol{\beta}x + \mathbf{H}\mathbf{W}_2\mathbf{v} + \mathbf{z} \\ &= \mathbf{H}\mathbf{W}_1\boldsymbol{\beta}\theta + \mathbf{H}\mathbf{W}_1\boldsymbol{\beta}\omega + \mathbf{z}, \end{aligned} \tag{4.61a}$$

$$\begin{aligned} \mathbf{y}_e &= \mathbf{H}_e\mathbf{W}_1\boldsymbol{\beta}x + \mathbf{H}_e\mathbf{W}_2\mathbf{v} + \mathbf{z}_e \\ &= \mathbf{H}_e\mathbf{W}_1\boldsymbol{\beta}\theta + \mathbf{H}_e\mathbf{W}_1\omega + \mathbf{H}_e\mathbf{W}_2\mathbf{v} + \mathbf{z}_e. \end{aligned} \tag{4.61b}$$

where the artificial noise $\mathbf{v} \in \mathbb{C}^{(N_t - N_r)\times 1}$ has $N_t - N_r$ i.i.d. complex Gaussian elements with zero mean and variance p_a.

It can be seen from (4.61) that the sensor transmits observation information $\mathbf{W}_1\boldsymbol{\beta}x$ plus a 'noise' term $\mathbf{W}_2\mathbf{v}$, which is chosen to be a random vector in the null space of $\mathbf{H}$, to reduce the possibility of small noise being seen by the eavesdropper. As $[\mathbf{W}_1, \mathbf{W}_2]$ is a unitary matrix, we obtain that $\mathbf{H}_e\mathbf{W}_1$ is independent of $\mathbf{H}_e\mathbf{W}_2$, giving the effective noise at the eavesdropper as $\mathbf{H}_e\mathbf{W}_2\mathbf{v} + \mathbf{z}_e$. The transmit power in each fading block is given as $\left(\sigma_\theta^2 + \sigma_\omega^2\right)\boldsymbol{\beta}^{\mathrm{H}}\boldsymbol{\beta} + (N_t - N_r)\, p_a$.

We want to minimise the distortion outage probability at the FC, by finding the best $\boldsymbol{\beta}^*(\mathbf{H})$ and $p_a^*(\mathbf{H})$ to meet the long-term power constraint and the secrecy outage constraint at the eavesdropper. Assuming that both the FC and the eavesdropper use the MMSE estimator, the optimization problem can be written as

$$\begin{aligned} &\min_{p_a(\mathbf{H}),\, \boldsymbol{\beta}(\mathbf{H})} \Pr\left[(\mathbf{H}\mathbf{W}_1\boldsymbol{\beta})^{\mathrm{H}}\mathbf{H}\mathbf{W}_1\boldsymbol{\beta} < \mathbb{S}^{-1}\right] \\ &s.t.\ \Pr\Big[(\mathbf{H}_e\mathbf{W}_1\boldsymbol{\beta})^{\mathrm{H}}\Big(\left(\sigma_\theta^2+\sigma_\omega^2\right)\mathbf{H}_e\mathbf{W}_1\boldsymbol{\beta}\,(\mathbf{H}_e\mathbf{W}_1\boldsymbol{\beta})^{\mathrm{H}} \\ &\qquad\qquad +\sigma_e^2\mathbf{I}_{N_e}p_a\mathbf{H}_e\mathbf{W}_2\,(\mathbf{H}_e\mathbf{W}_2)^{\mathrm{H}}\Big)^{-1}\mathbf{H}_e\mathbf{W}_1\boldsymbol{\beta} > \frac{\sigma^2 - \mathbb{D}_e}{\sigma^4}\Big] \le \delta, \\ &\quad \mathbb{E}\left[\left(\sigma_\theta^2 + \sigma_\omega^2\right)\boldsymbol{\beta}^{\mathrm{H}}\boldsymbol{\beta} + (N_t - N_r)\, p_a\right] \le \mathcal{P}_{\mathrm{tot}}. \end{aligned} \tag{4.62}$$

In order to solve problem (4.62), we can employ similar techniques as described in

Section 4.2.2, which are omitted for brevity. For the special case of a single receive antenna at both the FC and the eavesdropper, the problem is reduced to finding p and p_a, where $p \triangleq \boldsymbol{\beta}^{\mathrm{H}}\boldsymbol{\beta} \in \mathbb{R}$. Let $\hat{p}_a$ be the solution of

$$\hat{p}_a\left(\mathbf{H}\right) = \arg\min_{\hat{p}_a \geq 0} \lambda\left(N_t - 1\right)\hat{p}_a + \nu d\left(\mathbf{H}, \hat{p}_a\right), \tag{4.63}$$

where λ and ν are the corresponding Lagrange multipliers for the long-term power constraint and secrecy outage constraint of problem (4.62), and $d\left(\mathbf{H}, \hat{p}_a\right) = \int_{\mathbf{H}_e} 1\left\{\hat{p}_a < \frac{\mathbb{S}^{-1}|\mathbf{H}_e\mathbf{W}_1|^2}{|\mathbf{H}\mathbf{W}_1|^2|\mathbf{H}_e\mathbf{W}_2|^2}\frac{\mathbb{D}_e\left(\sigma_\theta^2+\sigma_\omega^2\right)-\sigma_\omega^2\sigma_\theta^2}{\sigma_\theta^2-\mathbb{D}_e} - \frac{\sigma_e^2}{|\mathbf{H}_e\mathbf{W}_2|^2}\right\} dF\left(\mathbf{H}_e\right)$. We then derive the locally optimal $p^*\left(\mathbf{H}\right)$ and $p_a^*\left(\mathbf{H}\right)$ as

$$\begin{cases} p^*\left(\mathbf{H}\right) = \frac{\mathbb{S}^{-1}}{|\mathbf{H}\mathbf{W}_1|^2},\ p_a^*\left(\mathbf{H}\right) = \hat{p}_a\left(\mathbf{H}\right); & \text{if } \frac{\lambda\left(\sigma_\theta^2+\sigma_\omega^2\right)\mathbb{S}^{-1}}{|\mathbf{H}\mathbf{W}_1|^2} + \nu d\left(\mathbf{H}, \hat{p}_a\right) + \lambda\left(N_t - 1\right)\hat{p}_a\left(\mathbf{H}\right) < 1 \\ p^*\left(\mathbf{H}\right) = p_a^*\left(\mathbf{H}\right) = 0, & \text{otherwise.} \end{cases}$$

4.4 Alternative Formulations

In Section 4.2 and Section 4.3 we considered problems that minimise the distortion outage probability at the FC while maintaining the secrecy outage probability at the eavesdropper and overall power consumption to be below certain thresholds. Alternative problems can also be formulated. For instance, we can minimise the secrecy outage probability at the eavesdropper, with a distortion outage constraint at the FC and a long-term power constraint among sensors, given as

$$\begin{aligned} &\min_{\mathbf{P}(\mathbf{G})} \Pr\left[D_e\left(\mathbf{G}, \mathbf{P}\left(\mathbf{G}\right)\right) < \mathbb{D}_e\right] \\ &s.t.\ \Pr\left[D\left(\mathbf{G}, \mathbf{P}\left(\mathbf{G}\right)\right) > \mathbb{D}\right] \leq \phi, \\ &\quad \mathbb{E}\left[\langle \mathbf{P}\left(\mathbf{G}\right)\rangle\right] \leq \mathcal{P}_{\mathrm{tot}}, \end{aligned} \tag{4.64}$$

where ϕ is the distortion outage probability threshold at the FC. Another potential problem would be to minimise the long-term expected estimation error at the FC subject to a secrecy outage constraint at the eavesdropper and a long-term power constraint among

the sensors, written as

$$\begin{aligned} &\min_{\mathbf{P}(\mathbf{G})} \mathbb{E}\left[D\left(\mathbf{G}, \mathbf{P}\left(\mathbf{G}\right)\right)\right] \\ &s.t.\ \Pr\left[D_e\left(\mathbf{G}, \mathbf{P}\left(\mathbf{G}\right)\right) < \mathbb{D}_e\right] \leq \delta, \\ &\quad \mathbb{E}\left[\langle \mathbf{P}\left(\mathbf{G}\right)\rangle\right] \leq \mathcal{P}_{\text{tot}}. \end{aligned} \tag{4.65}$$

For both problems, we could consider the full CSI and the partial CSI cases, which can both be solved using similar techniques as in Section 4.2. Note that problems (4.64), (4.65) are formulated for the multiple-sensor scenario. Similar problem formulations could also be constructed for a multiple-antenna scenario.

4.5 Numerical Results

We first consider a situation with three sensors. For simplicity, we consider the source σ_θ^2 to be distributed as $N\left(0,1\right)$, and all three sensors share the same measurement sensitivity of $\sigma_{\omega k}^2 = 10^{-3}, \forall k$. We assume that the distances from each sensor to the eavesdropper are 125m, 127m and 129m, whereas it is 125m, 130m and 135m to the FC respectively. Furthermore, we consider the path-loss of the signal power at the FC and the eavesdropper as following the free-space path-loss model [29]

$$PL = 20\log_{10}(d) + 20\log_{10}(f) - 27.55, \tag{4.66}$$

where $d \in \{d_k, d_{ek}\}$ is the distance between sensor k and the FC or the eavesdropper in meters, and f is the signal frequency in megahertz (we assume the network uses an operation frequency of 800MHz). Then, the channel power gain follows an exponential distribution with mean $10^{-\frac{PL}{10}}$mW. In addition, the total power budget range is set to 1 mW $\leq \mathcal{P}_{\text{tot}} \leq$ 11 mW, to ensure that the secrecy outage probability requirement at the eavesdropper is achievable. The maximum acceptable distortion level $\mathbb{D}$ at the FC is set to 0.007 while the required minimum distortion level $\mathbb{D}_e$ at the eavesdropper is 0.01.

Figure 4.3 shows the distortion outage probability at the FC with two antennas at the

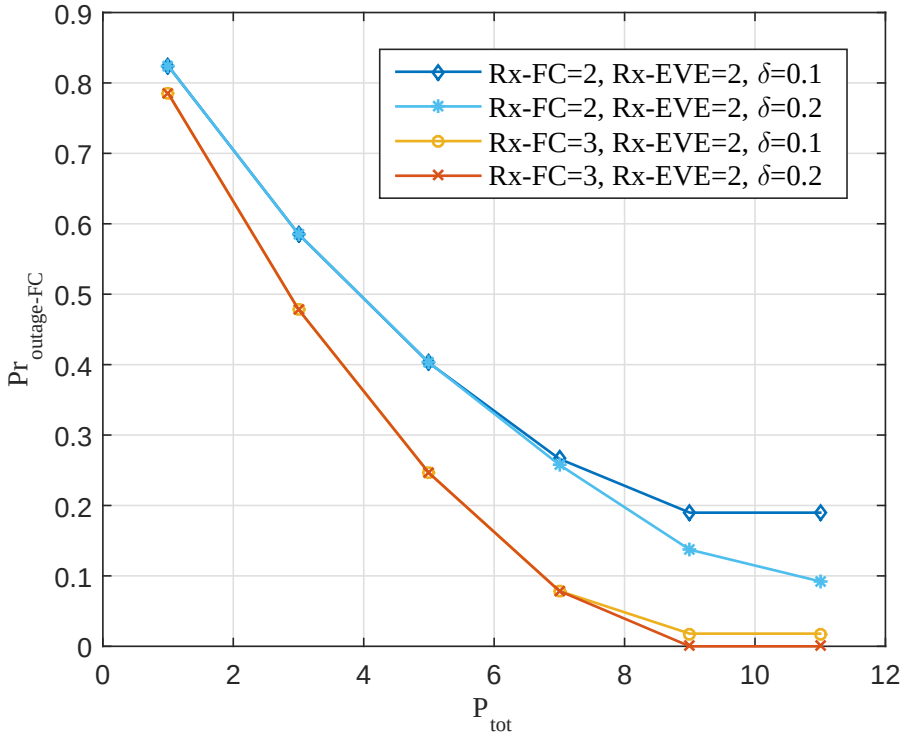

Figure 4.3: Performance comparison in a three-sensor network with $N_e = 2$ and full CSI of both the FC and the eavesdropper.

eavesdropper, under different secrecy outage probability requirements at the eavesdropper, namely 0.1 and 0.2. When the number of receive antennas at the FC is fixed, it is seen that for both sets (i.e., $N_r = 2$ and $N_r = 3$) the outage probability at the FC behaves similarly for the two different outage requirements at the eavesdropper when $\mathcal{P}_{\text{tot}}$ is small. As we increase the total power budget, they start to decrease until saturation. This is because when $\mathcal{P}_{\text{tot}}$ is small, the sensors are more likely to choose small power consumption policies that only guarantee non-outage at the FC, or the sensors would simply stop transmitting to save power. As the transmission power budget increases, sensors begin to transmit in channel states where outage happens neither at the FC nor at the eavesdropper, until a point where more incremental power would lead to the secrecy outage probability at the eavesdropper being greater than the security requirement δ, at which the distortion outage probability at the FC saturates.

In Figure 4.4, we compare the distortion outage probability at the FC with the sensor

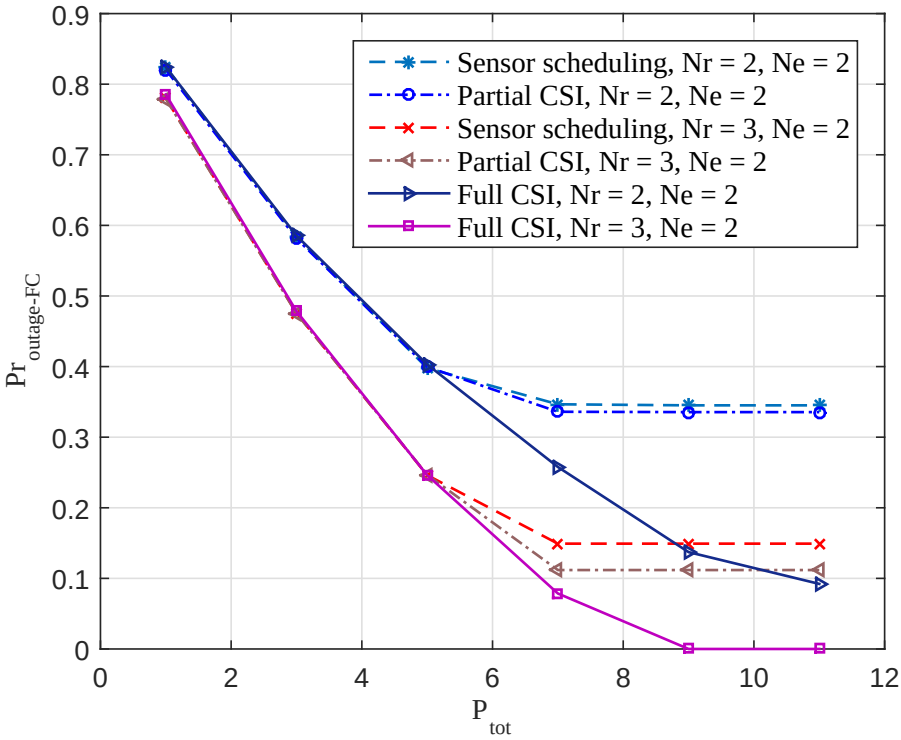

Figure 4.4: Performance comparison in a three-sensor network with $N_e = 2$ and $\delta = 0.2$.

scheduling scheme, partial CSI, and full CSI schemes in a three-sensor network, with the FC having two or three antennas. As we can see, similar to Figure 4.3, the outage probability at the FC is smaller when the FC is equipped with more antennas for all three cases. In addition, the performance of sensor scheduling follows closely the partial CSI case, and it even has similar performance as the full CSI case when the transmit power budget is small.

The distortion outage probability at the FC versus different transmit power budgets is plotted in Figure 4.5, where we compare the performance of sensor scheduling to the partial CSI case with the secrecy outage probability constraint at the eavesdropper set to 0.14, 0.18 and 0.22. The first thing to be noticed is the close performance of sensor-scheduling and partial CSI power allocation in all three scenarios (i.e, $\delta = 0.14$, $\delta = 0.18$ and $\delta = 0.22$) when the power budget $\mathcal{P}_{\text{tot}}$ is relatively small. In addition, the results stated in (4.44) and (4.45) can be easily verified from the behaviour of sensor-scheduling. When we have a small power budget, $\text{Pr}_{\text{outage\_FC}}$ performs the same for all scenarios re-

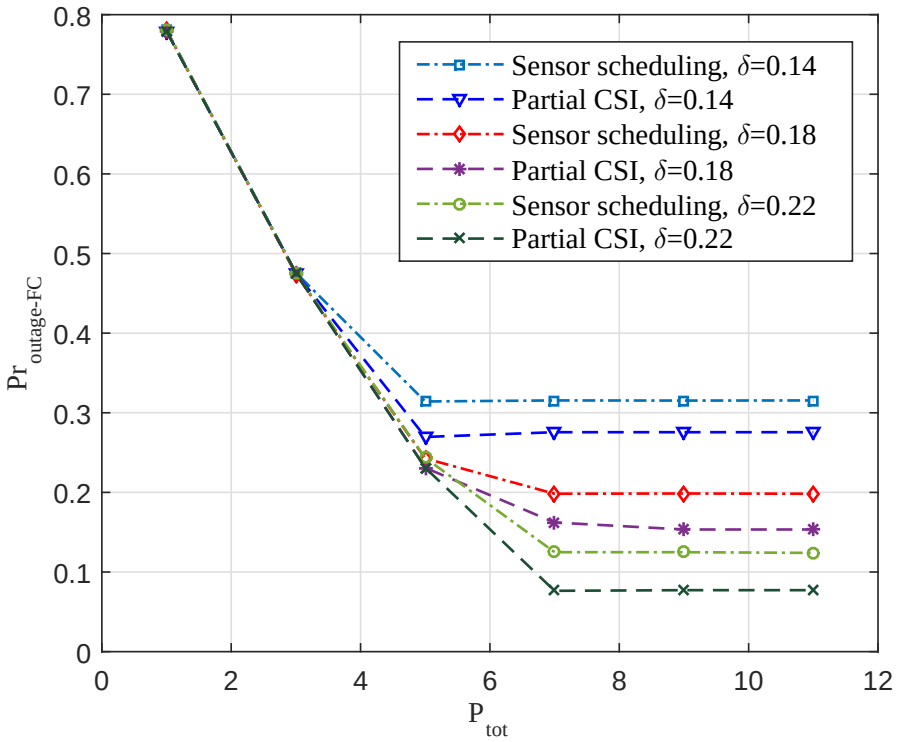

Figure 4.5: Performance comparison in a three-sensor network with $N_r = 3$ and $N_e = 2$.

gardless of the different secrecy outage requirements at the eavesdropper, which implies that the total power constraint satisfies equality at the optimal points, whereas the secrecy outage constraint is loose. As we keep increasing the power budget, $\text{Pr}_{\text{outage\_FC}}$ settles down to a point at which the secrecy outage constraint is satisfied with equality but the power constraint is loose, since any power increment makes no improvement.

Next, we study the distortion outage probability at the FC for the multiple-antenna single sensor scenario, where we assume that the sensor is 127m away from the FC, and 130m away from the eavesdropper. For simplicity, we assume that the sensor is equipped with three antennas, whereas there is only one antenna at the FC and one or two antennas at the eavesdropper. We consider the minimum required distortion level $\mathbb{D}_e$ at the eavesdropper being set to 0.013, which is twice as large as the maximum acceptable distortion level $\mathbb{D}$ at the FC. We assume the same noise level for both the FC and the eavesdropper, where $\sigma_n^2 = \sigma_e^2 = 10^{-8}$ mW.

In Figure 4.6, the distortion outage probability at the FC versus the long-term power

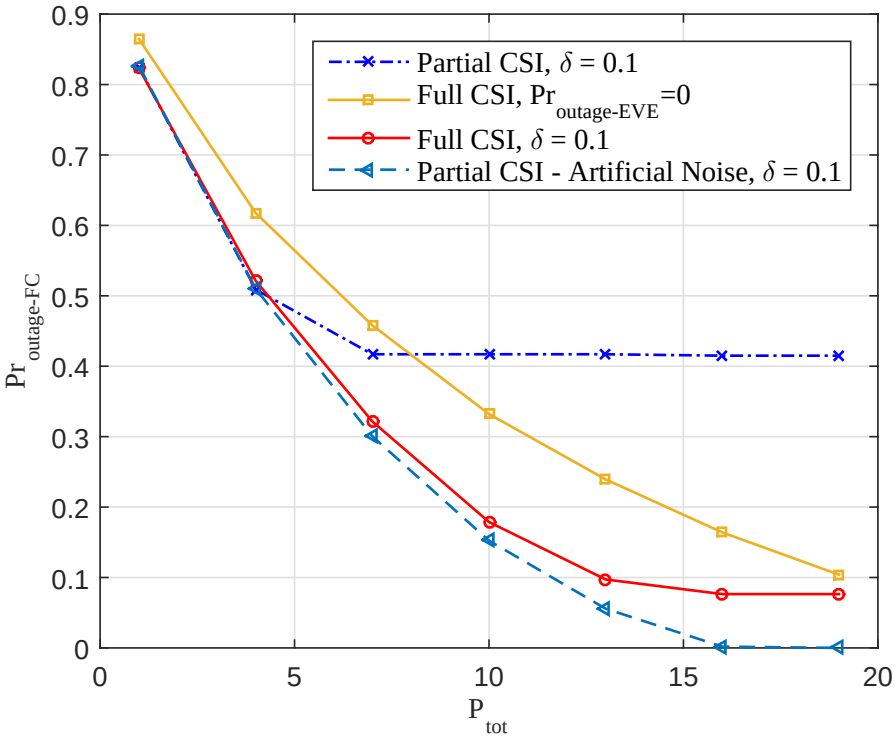

Figure 4.6: Performance comparison for a single sensor multiple-antenna scenario with $N_r = N_e = 1$ and $\delta = 0.1$.

budget is plotted for full CSI, full CSI with $\text{Pr}_{\text{outage\_EVE}} = 0$, partial CSI and partial CSI-Artificial Noise schemes. As we can see, the full CSI case outperforms the partial CSI case, and in both cases the distortion outage probability at the FC saturates. By contrast, the full CSI with $\text{Pr}_{\text{outage\_EVE}} = 0$ and partial CSI-Artificial Noise schemes perform better when we have a relatively large transmit power budget, where $\text{Pr}_{\text{outage\_FC}}$ keeps decreasing as $\mathcal{P}_{\text{tot}}$ increases. More interestingly, it is seen from that the full CSI $\text{Pr}_{\text{outage\_EVE}} = 0$ scheme performs no better than the partial CSI-Artificial Noise scheme across the entire power range. This is owing to the fact that the effective channel gains of the FC are largely reduced when projecting it onto the eavesdropper's channel null space, whereas in the case of partial CSI-Artificial Noise, only a small portion of the transmit power is used to generate 'noise'.

To closely observe the performance of $\text{Pr}_{\text{outage\_FC}}$ using artificial noise, in Figure 4.7 we look at scenarios where the eavesdropper has more receive antennas than the FC, and

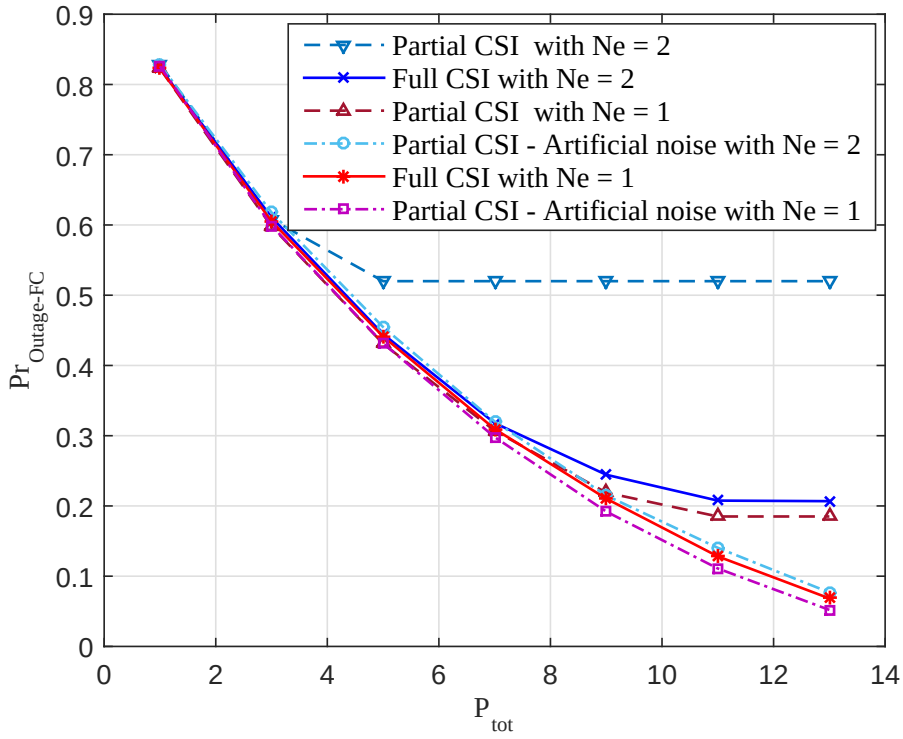

Figure 4.7: Performance comparison for a single sensor multiple-antenna scenario with $N_r = 1$ and $\delta = 0.2$.

we also plot the full CSI and partial CSI cases for comparison. It is noticeable that when the eavesdropper has more antennas, $\Pr_{\text{outage\_FC}}$ in the partial CSI case quickly saturates which is then followed by the full CSI case, as at certain channel states the sensor has to stop transmitting in order to maintain the required secrecy outage probability at the eavesdropper. Whereas in the case of partial CSI-Artificial Noise, because the sensor can intentionally generate noise to degrade the eavesdropper's channel, it can explore more channel states to transmit the observation signals to the FC. Similar behaviour is seen when the eavesdropper has the same number of antennas as the FC, where the partial CSI-Artificial Noise gives better performance, as less 'noise' needs to be produced which means more power can be used to forward the observations. Therefore, the simulation results in Figure 4.6 and Figure 4.7 indicate that injecting artificial noise into the eavesdropper's channel appears to be a better solution for the single sensor multiple-antenna scenario.

4.6 Conclusion

In this chapter, we have considered the problem of transmit power allocation for distortion outage probability minimization in the presence of an eavesdropper. We studied the distortion outage probability performance for both full CSI and partial CSI under two different scenarios: multiple-sensor single antenna scenario and multiple-antenna single sensor scenario. We proposed a suboptimal solution (for the partial CSI case) to overcome the high computational cost in the multiple-sensor scenario. With multiple transmit antennas at the sensor, we investigated techniques that can achieve zero outage at the eavesdropper. Simulation results showed that better performance can be achieved with additional receive antennas at the FC for the multiple-sensor scenario, and in the multiple-antenna single sensor scenario the distortion outage probability at the FC can be reduced to zero if the transmit power budget is sufficiently large.

4.7 Appendix

4.7.1 *Proof of Lemma 1*

We will show that the power allocation policy given in (4.12) is feasible, i.e., $\mathbf{P}'(\mathbf{G})$ satisfies the secrecy outage constraint at the eavesdropper (4.7a) and the total transmit power constraint (4.7b); and that $\mathbf{P}'(\mathbf{G})$ performs at least as well as $\hat{\mathbf{P}}(\mathbf{G})$.

Since $\hat{\mathbf{P}}(\mathbf{G})$ is feasible, $\hat{\mathbf{P}}(\mathbf{G})$ must satisfy all the constraints, i.e.,

$$\begin{aligned}&\Pr\left[D_e\left(\mathbf{G},\hat{\mathbf{P}}(\mathbf{G})\right)<\mathbb{D}_e\right]\\=&\mathbb{E}_{\mathbf{G}}\left[\Pr\left[D_e\left(\mathbf{G},\hat{\mathbf{P}}(\mathbf{G})\right)<\mathbb{D}_e\middle|\mathbf{G}\right]\right]\\=&\mathbb{E}_{\mathbf{G}}\left[\omega_2(\mathbf{G})\right]\leq\delta,\end{aligned}$$

and

$$\begin{aligned}&\mathbb{E}_{\mathbf{G},\mathbf{p}}\left[\left\langle\hat{\mathbf{P}}(\mathbf{G})\right\rangle\right]\\=&\mathbb{E}_{\mathbf{G}}\left[\left\langle\sum_{i=1}^{3}\mathbb{E}_{\mathbf{p}}\left(\hat{\mathbf{P}}(\mathbf{G})\middle|\mathbf{p}(\mathbf{G})\in\mathcal{B}_i(\mathbb{D},\mathbb{D}_e,\mathbf{G}),\mathbf{G}\right)\Pr\left[\mathbf{p}(\mathbf{G})\in\mathcal{B}_i(\mathbb{D},\mathbb{D}_e,\mathbf{G})\middle|\mathbf{G}\right]\right\rangle\right]\\=&\mathbb{E}_{\mathbf{G}}\left[\left\langle\sum_{i=1}^{3}\mathbf{p}_i(\mathbf{G})\,\omega_i(\mathbf{G})\right\rangle\right]\leq\mathcal{P}_{\text{tot}}.\end{aligned}\tag{4.67}$$

Remark: As $\hat{\mathbf{P}}(\mathbf{G})$ has three non-overlapping regions as defined in (4.11), with all powers in $\mathcal{B}_3(\mathbb{D},\mathbb{D}_e,\mathbf{G})$ being zero, we know that the only power region leading to outage at the eavesdropper is $\mathcal{B}_2(\mathbb{D},\mathbb{D}_e,\mathbf{G})$. In addition, the probability of choosing a power in $\mathcal{B}_2(\mathbb{D},\mathbb{D}_e,\mathbf{G})$ is given as $\Pr\left[\mathbf{p}(\mathbf{G})\in\mathcal{B}_2(\mathbb{D},\mathbb{D}_e,\mathbf{G})\middle|\mathbf{G}\right]$, which is the same as the time-sharing factor $\omega_2(\mathbf{G})$ defined in (4.15).

As the new probabilistic power allocation $\mathbf{P}'(\mathbf{G})$ is randomised among the three deterministic power policies given in (4.14), we can find the long-term average power consumption of $\mathbf{P}'(\mathbf{G})$ as

$$\mathbb{E}\left[\left\langle\mathbf{p}(\mathbf{G})\right\rangle\right]=\mathbb{E}_{\mathbf{G}}\left[\left\langle\mathbb{E}_{\mathbf{p}}\left[\mathbf{p}(\mathbf{G})\middle|\mathbf{G}\right]\right\rangle\right]$$

$$= \mathbb{E}_{\mathbf{G}} \left[\left\langle \sum_{i=1}^{3} \mathbf{p}_i (\mathbf{G}) \, \omega_i (\mathbf{G}) \right\rangle \right], \tag{4.68}$$

and hence $\mathbf{P}'(\mathbf{G})$ satisfies the power constraint (4.7b).

In addition, as both $D(\mathbf{G}, \mathbf{p}(\mathbf{G}))$ and $D_e(\mathbf{G}, \mathbf{p}(\mathbf{G}))$ are continuous and convex over $\mathbf{p}(\mathbf{G})$, by the Mean Value Theorem (MVT) for integration [41], we know that, for a given channel realization $\mathbf{G}$, there exists a $\hat{\mathbf{p}}_1(\mathbf{G}) \in \mathcal{B}_1(\mathbb{D}, \mathbb{D}_e, \mathbf{G})$ such that $\hat{\mathbf{p}}_1(\mathbf{G}) = \mathbb{E}\left[\hat{\mathbf{P}}(\mathbf{G}) \middle| \mathbf{p}(\mathbf{G}) \in \mathcal{B}_1(\mathbb{D}, \mathbb{D}_e, \mathbf{G}), \mathbf{G}\right]$. Together with the definition of $\mathbf{p}_1(\mathbf{G})$ in (4.14), we know that $D(\mathbf{G}, \mathbf{p}_1(\mathbf{G})) \leq \mathbb{D}$ and $D_e(\mathbf{G}, \mathbf{p}_1(\mathbf{G})) \geq \mathbb{D}_e$. Similarly, only when $\mathbf{P}'(\mathbf{G}) = \mathbf{p}_2(\mathbf{G})$ does outage occur at the eavesdropper. Therefore, we can compute the secrecy outage probability at the eavesdropper when using the probabilistic power policy $\mathbf{P}'(\mathbf{G})$ as

$$\begin{aligned} \Pr\left[D_e\left(\mathbf{G}, \mathbf{P}'(\mathbf{G})\right) < \mathbb{D}_e\right] &= \mathbb{E}_{\mathbf{G}}\left[\Pr\left[\mathbf{P}'(\mathbf{G}) = \mathbf{p}_2(\mathbf{G}) \middle| \mathbf{G}\right]\right] \\ &= \mathbb{E}_{\mathbf{G}}\left[\omega_2(\mathbf{G})\right] \leq \delta. \end{aligned} \tag{4.69}$$

Remark: Note that for the channel states where $\mathcal{B}_1(\mathbb{D}, \mathbb{D}_e, \mathbf{G}) = \emptyset$, the result given in (4.69) can be also established, since for those channel states we have $\omega_1(\mathbf{G}) = 0$. By following the above arguments and applying the MVT, we see that outage occurs at the eavesdropper only when $\mathbf{P}'(\mathbf{G}) = \mathbf{p}_2(\mathbf{G})$.

The feasibility of $\mathbf{P}'(\mathbf{G})$ has thus been proved. In order to see that the probabilistic power policy $\mathbf{P}'(\mathbf{G})$ performs no worse than $\hat{\mathbf{P}}(\mathbf{G})$, we first show that for each channel realisation, the distortion outage probability at the FC when using $\mathbf{P}'(\mathbf{G})$ is at least as small as when using $\hat{\mathbf{P}}(\mathbf{G})$. We then conclude that for a fixed maximum acceptable distortion level $\mathbb{D}$ at the FC, $\mathbf{P}'(\mathbf{G})$ would result in the same or smaller outage probability at the FC. Given the channel realisation $\mathbf{G}$, the distortion outage probability at the FC is:

$$\begin{aligned} &\Pr\left[\mathrm{D}\left(\mathbf{G}, \hat{\mathbf{P}}(\mathbf{G})\right) > \mathbb{D} \middle| \mathbf{G}\right] \\ &= \sum_{i=1}^{3} \Pr\left[\mathrm{D}\left(\mathbf{G}, \hat{\mathbf{P}}(\mathbf{G})\right) > \mathbb{D} \middle| \mathbf{p}(\mathbf{G}) \in \mathcal{B}_{\mathrm{i}}, \mathbf{G}\right] \Pr\left[\mathbf{p}(\mathbf{G}) \in \mathcal{B}_{\mathrm{i}} \middle| \mathbf{G}\right] \\ &\overset{(a)}{=} \sum_{i=1}^{3} \mathbb{E}_{\mathbf{p}}\left[1\left\{D\left(\mathbf{G}, \hat{\mathbf{P}}(\mathbf{G})\right) > \mathbb{D} \middle| \mathbf{p}(\mathbf{G}) \in \mathcal{B}_i, \mathbf{G}\right\}\right] \omega_i(\mathbf{G}) \end{aligned}$$

$$\begin{aligned}&\overset{(b)}{\geq}\sum_{i=1}^{3}1\left\{D\left(\mathbf{G},\mathbb{E}_{\mathbf{p}}\left[\hat{\mathbf{P}}\left(\mathbf{G}\right)\middle|\,\mathbf{p}\left(\mathbf{G}\right)\in\mathcal{B}_i,\mathbf{G}\right]\right)>\mathbb{D}\right\}\omega_i\left(\mathbf{G}\right)\\&=\Pr\left[D\left(\mathbf{G},\mathbf{P}'\left(\mathbf{G}\right)\right)>\mathbb{D}\middle|\,\mathbf{G}\right],\end{aligned}\tag{4.70}$$

where (a) follows from the definition of $\{\omega_i(\mathbf{G})\}$ given in (4.15) and (b) follows from Jensen's inequality, since $D(\mathbf{G},\mathbf{p}(\mathbf{G}))$ is a convex function over $\mathbf{p}(\mathbf{G})$, and the last equality follows from (4.12). Therefore, the resulting distortion outage probability at the FC from using $\mathbf{P}'(\mathbf{G})$ is no worse than using $\hat{\mathbf{P}}(\mathbf{G})$, i.e.,

$$\Pr\left[D\left(\mathbf{G},\mathbf{P}'\left(\mathbf{G}\right)\right)>\mathbb{D}\right]\leq\Pr\left[D\left(\mathbf{G},\hat{\mathbf{P}}\left(\mathbf{G}\right)\right)>\mathbb{D}\right].\tag{4.71}$$

Combining (4.68), (4.69) and (4.71), we conclude that a probabilistic power allocation scheme $\mathbf{P}'(\mathbf{G})$ with the form (4.12) is feasible and gives the same or smaller outage probability at the FC compared to an arbitrary probabilistic power allocation. Furthermore, from the definition of $\{\mathbf{p}_i(\mathbf{G})\}$ given in (4.14), we have the following:

$$\begin{aligned}\mathbb{D}_e\leq&\mathbb{E}_{\mathbf{p}}\left[D_e\left(\mathbf{G},\mathbf{P}'\left(\mathbf{G}\right)\right)\middle|\,D_e\left(\mathbf{G},\mathbf{p}\left(\mathbf{G}\right)\right)\geq\mathbb{D}_e,\mathbf{G}\right]\\\overset{(c)}{=}&\frac{\omega_3\left(\mathbf{G}\right)\sigma_\theta^2}{\omega_1\left(\mathbf{G}\right)+\omega_3\left(\mathbf{G}\right)}+\frac{\omega_1\left(\mathbf{G}\right)D_e\left(\mathbf{G},\mathbf{p}_1\left(\mathbf{G}\right)\right)}{\omega_1\left(\mathbf{G}\right)+\omega_3\left(\mathbf{G}\right)}\end{aligned}\tag{4.72}$$

$$\begin{aligned}\mathbb{D}\geq&\mathbb{E}_{\mathbf{p}}\left[D\left(\mathbf{G},\mathbf{P}'\left(\mathbf{G}\right)\right)\middle|\,D\left(\mathbf{G},\mathbf{p}\left(\mathbf{G}\right)\right)\leq\mathbb{D},\mathbf{G}\right]\\\overset{(d)}{=}&\frac{\omega_1\left(\mathbf{G}\right)D\left(\mathbf{G},\mathbf{p}_1\left(\mathbf{G}\right)\right)}{\omega_1\left(\mathbf{G}\right)+\omega_2\left(\mathbf{G}\right)}+\frac{\omega_2\left(\mathbf{G}\right)D\left(\mathbf{G},\mathbf{p}_2\left(\mathbf{G}\right)\right)}{\omega_1\left(\mathbf{G}\right)+\omega_2\left(\mathbf{G}\right)},\end{aligned}\tag{4.73}$$

where (c) and (d) are obtained by applying conditional expectations.

4.7.2 *Proof of Theorem* 4.1

We will consider the case $\omega_j^*(\mathbf{G})=1$, as when $\omega_j^*(\mathbf{G})=0$, the solution of $\mathbf{p}_j^*(\mathbf{G})$ has no impact on the optimization problem.

(1) When $\nu^*(\mathbf{G})=0$: From the KKT condition (4.22), we need to have $\omega_1^*(\mathbf{G})D(\mathbf{G},\mathbf{p}_1^*(\mathbf{G}))+\omega_2^*(\mathbf{G})D(\mathbf{G},\mathbf{p}_2^*(\mathbf{G}))-(\omega_1^*(\mathbf{G})+\omega_2^*(\mathbf{G}))\mathbb{D}=0$. If $\omega_1^*(\mathbf{G})=1$,

we must have $\omega_2^*(\mathbf{G}) = 0$, and so $D(\mathbf{G}, \mathbf{p}_1^*(\mathbf{G})) = \mathbb{D}$. However, we also know that $\frac{\partial l(\ldots)}{\partial p_{1k}^*(\mathbf{G})} = \lambda^* \omega_1^*(\mathbf{G}) \left(\sigma_{\omega k}^2 + \sigma_\theta^2\right) - \nu_e^*(\mathbf{G}) \omega_1^*(\mathbf{G}) \frac{\partial D_e(\mathbf{G}, \mathbf{p}_1^*(\mathbf{G}))}{\partial p_{1k}^*(\mathbf{G})} \geq 0$. Combining with (4.17) we see that $\mathbf{p}_1^*(\mathbf{G}) = \mathbf{0}$, which contradicts the requirement that $D(\mathbf{G}, \mathbf{p}_1^*(\mathbf{G})) = \mathbb{D}$. Similar arguments apply for the case when $\omega_2^*(\mathbf{G}) = 1$. Therefore, we conclude that if $\nu^*(\mathbf{G}) = 0$ we must have $\omega_1^*(\mathbf{G}) = \omega_2^*(\mathbf{G}) = 0$.

(2) When $\nu^*(\mathbf{G}) > 0$ and $\nu_e^*(\mathbf{G}) = 0$: Here, one should have $\omega_1^*(\mathbf{G}) \left[D_e(\mathbf{G}, \mathbf{p}_1^*(\mathbf{G})) - \sigma_\theta^2\right] + \omega_2^*(\mathbf{G}) \left(\mathbb{D}_e - \sigma_\theta^2\right) \geq \mathbb{D}_e - \sigma_\theta^2$ and $\omega_1^*(\mathbf{G}) D(\mathbf{G}, \mathbf{p}_1^*(\mathbf{G})) + \omega_2^*(\mathbf{G}) D(\mathbf{G}, \mathbf{p}_2^*(\mathbf{G})) - (\omega_1^*(\mathbf{G}) + \omega_2^*(\mathbf{G})) \mathbb{D} = 0$. If $\omega_1^*(\mathbf{G}) = 1$, we obtain $\omega_2^*(\mathbf{G}) = 0$, $D(\mathbf{G}, \mathbf{p}_1^*(\mathbf{G})) = \mathbb{D}$ and $D_e(\mathbf{G}, \mathbf{p}_1^*(\mathbf{G})) \geq \mathbb{D}_e$. In addition, from (4.25) we see that $p_{1k}^*(\mathbf{G})$ satisfies $\left(\sigma_{\omega k}^2 + \sigma_\theta^2\right) - \frac{\nu^*(\mathbf{G})}{\lambda^*} \frac{\partial D(\mathbf{G}, \mathbf{p}_1^*(\mathbf{G}))}{\partial p_{1k}^*(\mathbf{G})} = 0$. For problem (4.31), from the KKT conditions, we know that the optimal solution $\mathbf{p}_a^*(\mathbf{G})$ must satisfy $D(\mathbf{G}, \mathbf{p}_a^*(\mathbf{G})) = \mathbb{D}$, $D_e(\mathbf{G}, \mathbf{p}_a^*(\mathbf{G})) \geq \mathbb{D}_e$, and $\left(\sigma_{\omega k}^2 + \sigma_\theta^2\right) - \hat{\nu}^*(\mathbf{G}) \frac{\partial D(\mathbf{G}, \mathbf{p}_a^*(\mathbf{G}))}{\partial p_{ak}^*(\mathbf{G})} = 0, \forall k$, which shares the same form as $\mathbf{p}_1^*(\mathbf{G})$, where $\hat{\nu}^*(\mathbf{G})$ is the optimal Lagrange multiplier corresponding to the distortion constraint at the FC for problem (4.31). Therefore, if $\omega_1^*(\mathbf{G}) = 1$ we have $\mathbf{p}_1^*(\mathbf{G}) = \mathbf{p}_a^*(\mathbf{G})$. Similarly, for $\omega_2^*(\mathbf{G}) = 1$, we obtain $\mathbf{p}_2^*(\mathbf{G}) = \mathbf{p}_b^*(\mathbf{G})$ if $D_e\left(\mathbf{G}, \mathbf{p}_b^*(\mathbf{G})\right) < \mathbb{D}_e$.

(3) When $\nu^*(\mathbf{G}) > 0$ and $\nu_e^*(\mathbf{G}) > 0$: The same results can be derived by using similar arguments as for case **(2)**.

4.7.3 *Proof of Lemma 2*

First, from (4.34) we know that the optimal power $\mathbf{p}^*(\mathbf{G})$ should minimise $1\{D(\mathbf{g}, \mathbf{p}^*(\mathbf{g})) > \mathbb{D}\} + \xi(\mathbf{p}^*(\mathbf{g}))$ at each channel instance. When $\mathbf{p}(\mathbf{G}) = \mathbf{0}$, we obtain $\xi(\mathbf{p}(\mathbf{g})) = 0$ and $1\{D(\mathbf{g}, \mathbf{p}(\mathbf{g})) > \mathbb{D}\} + \xi(\mathbf{p}(\mathbf{g})) = 1$, which indicates that $1\{D(\mathbf{g}, \mathbf{p}^*(\mathbf{g})) > \mathbb{D}\} + \xi(\mathbf{p}^*(\mathbf{g}))$ is upper bounded by 1. If $\mathbf{p}^*(\mathbf{G})$ is a nonzero vector, we must have $\xi(\mathbf{p}^*(\mathbf{g})) > 0$. Furthermore, as $1\{D(\mathbf{g}, \mathbf{p}^*(\mathbf{g})) > \mathbb{D}\} + \xi(\mathbf{p}^*(\mathbf{g})) \leq 1$, we obtain $\xi(\mathbf{p}^*(\mathbf{g})) \leq 1$ and $1\{D(\mathbf{g}, \mathbf{p}^*(\mathbf{g})) > \mathbb{D}\} = 0$.

0

Chapter 5

A Game-Theoretic Approach to DoS Attacks in Distributed Estimation

This chapter investigates the counter-attacking strategies of distributed estimation under the denial-of-service (DoS) attack in a single sensor network. The sensor transmits observations to the fusion center (FC), which then reconstructs a minimum mean square error (MMSE) estimate of the physical quantity observed. Meanwhile, the attacker transmits a jamming signal to disrupt the signal received by the FC. We first consider a zero-sum repeated game in non-fading scenarios, where the same static game is played over time. To find a strategy pair in a Nash equilibrium, modified backward induction and Nash Q-learning techniques are applied. In fading channels, we look at both full channel state information (CSI) and partial CSI cases, where Bayesian games are explored. Based on the knowledge of the player's own channel information and the belief of the opponent's channel distribution, we study the type-contingent power allocation strategy at a Bayesian Nash equilibrium. Numerical examples are provided to demonstrate the 'optimality' of the strategy pair.

5.1 Introduction

IN previous chapters, we looked at secure power allocation algorithms where a sensor network is under passive eavesdropping. Various transmission policies are considered to either minimize the estimate distortion level at the fusion center (FC) or to reduce distortion outage probability for different wireless sensor network (WSN) models. In this chapter, we shift our focus to denial-of-service (DoS) attacks[1] in distributed estimation, where an adversary attempts to diminish the quality of the estimates at the FC by injecting noise to the communication channels.

[1] A DoS attack is defined as an event that attempts to reduce a network's capability to operate as expected [102].

In DoS attacks, an adversary, even with small jamming power, is capable of disrupting the legitimate signal reception by broadcasting an interference signal[2]. Frequency hopping and spread-spectrum are two standard defense techniques against DoS attacks; however, these techniques largely reduce the sensors' battery life and are more likely to limit the WSN to single-frequency use [102]. Ideally, in DoS attacks, a sensor should not only take into consideration its own situation, but also the possible jamming signals from the attacker. Therefore, instead of only looking at the sensor's power allocation strategy, a system model capturing the interactions between two conflicting parties is more suitable for DoS attacks.

Game theory, which describes the conflict or cooperation between intelligent rational decision-makers, is generally employed to model the interactive behavior of multiple parties in a 'game' with limited resources or conflicting interests. Resource allocation in fading multiple access channels (MACs) is studied within a game-theoretic framework in [53]. Assuming that all users are selfish and rational, the authors obtained the maximum sum-rate point on the boundary of the MAC capacity region. With each player only having 'incomplete information' about others, the authors in [1,36] considered the power allocation problems in Bayesian games, where a player is aware of its own channel gain, but does not know the channel gains of others. The zero-sum mutual information game in correlated jamming is studied in [47,69,89], where the jammer may correlate its signal to that of the legitimate user. More recently, in DoS attacks, a game-theoretic framework is developed in [58] to analyze SINR-based DoS attacks on remote state estimation, where the players are assumed to be rational and limited by average power constraints. The authors build a Markov game to model the interactive decision-making process and used a modified Nash Q-learning algorithm to obtain the strategies at a Nash equilibrium.

In addition to DoS attacks, the sensors typically have limited energy resources and are geographically widespread, hence replacing batteries is considered costly. Many works have considered and studied cross-layer optimization to enhance the energy efficient transmission from the sensors to the FC. Applying analog amplify-and-forward technique, which is shown to be asymptotically optimal in estimating a Gaussian source

[2] A more severe type of attacks is called Byzantine attack, where the adversary may capture and subvert a limited number of sensors; the compromised sensors are said to suffer a Byzantine fault [66,85].

for a coherent MAC [7,25] and exactly optimal in [24] under certain situations, the authors in [104] studied the optimal power scheduling problem in an inhomogeneous sensor network. The same group of authors extended the results to a vector source and investigated the optimal power allocation policies in [105].

In this chapter, we study a game theoretic approach for distributed estimation in a single sensor network under DoS attacks. The sensor measures a single point Gaussian source, and then transmits the noisy measurements to the FC using an uncoded analog scheme. At the same time, the attack broadcasts a jamming signal attempting to interfere with signals received by the FC, who then attempts to reconstruct a minimum mean square error (MMSE) estimate of the observations. We assume that both the sensor and attacker are operating under a limited power budget. Hence, the sensor wants to reduce the estimate distortion at the FC while satisfying an overall power constraint. The attacker, on the other hand, is seeking to degrade the estimation quality at the FC. In this chapter, we aim to study the optimal power allocation strategies for both the sensor and attacker such that no one gains more benefits by unilateral deviation. The main contributions of the chapter are:

- We consider the interactive behavior between the transmission policies at the sensor and the jamming powers at the attacker in both non-fading and fading scenarios, where the game is played simultaneously. We also study the existence of a Nash equilibrium for different cases.

- In non-fading channels, we look at dynamic games where both players can 'interact' with each other many times hence learning the opponent's strategies and updating its own actions[3]. For a finitely repeated game, a power allocation algorithm is proposed to recursively find the optimal strategy pairs while meeting the power constraints at each player. The Nash Q-learning technique [38] is employed for the case of infinitely repeated games to obtain a stationary Nash equilibrium strategy pair.

- When fading is present, apart from the full CSI scenario we consider the partial

[3]Learning is a technique used to find the 'optimal' strategies. Techniques like backwards induction or Nash Q-learning is implemented by each player in an offline manner.

CSI case or incomplete information games, where each player has perfect knowledge about its own channel gains and only a *belief* about the opponent's statistical channel information. Under this setting, we study the Bayesian Nash equilibrium, where the 'optimal' transmission schemes at each player depend only on its own channel information[4].

The rest of the chapter is organised as follows. In Section 5.2, we present the system model, where the FC reconstructs MMSE estimates of the physical quantity observed while the network is under DoS attacks. In Section 5.3, we investigate transmission strategies at a Nash equilibrium in non-fading channel scenarios, in which two different repeated games are discussed. Fading channel scenarios are explored in Section 5.4, where we consider Bayesian games when only partial CSI is available at the players. Illustrative numerical results are provided in Section 5.6, followed by concluding remarks in Section 5.7.

5.2 System Model

We consider a wireless network with one sensor, one adversary and a FC as shown in Figure 5.1. The sensor is observing a single point Gaussian source, denoted by θ, which has zero mean and variance σ_θ^2. Applying analog amplify and forward techniques [24,25] with amplifying factor $\beta_S[t]$, at time t the sensor transmits the measurement

$$x[t] \triangleq \theta[t] + \omega[t],$$

which is the corrupted source with ω being i.i.d. zero mean variance σ_ω^2 complex Gaussian noise, to the FC over a noisy channel. The malicious node attempts to attack the transmitted observations reaching at the FC with a jamming signal β_A.

To model fading in the received power, we consider h_S and h_A as complex fading

[4] 'optimality' and 'optimal' are written within quotes in this chapter. This is because unlike one-player games which the optimality has an unambiguous meaning; the optimality in multi-person decision making is a not well-defined concept; therefore, the Nash equilibrium solution is considered as a specific form of 'optimality' [8].

random variables. Therefore, the signal received by the FC at time t is modeled as [89]

$$y = h_S[t]\beta_S[t]x[t] + h_A[t]\beta_A[t] + z[t], \tag{5.1}$$

where $z[t]$ is independent and identically distributed (i.i.d.) complex Gaussian noise with zero mean and variance σ_n^2. All fading random variables are assumed to be i.i.d. in time. The sensor node and the adversary are power constrained by the average power budget $\mathcal{P}_{S\_\text{avg}}$ and $\mathcal{P}_{A\_\text{avg}}$ respectively, this indicates that the power consumption averaged across all transmission slots should be no greater than $\mathcal{P}_{S\_\text{avg}}$ for the sensor and $\mathcal{P}_{A\_\text{avg}}$ for the attacker.

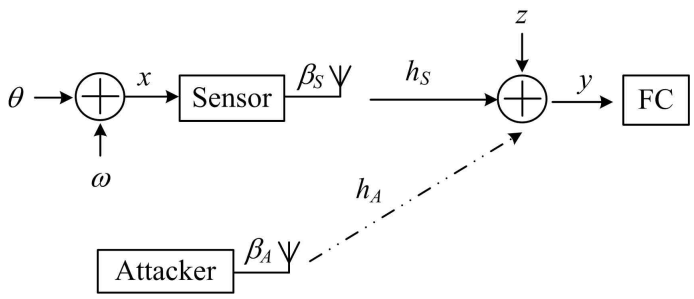

Figure 5.1: Diagram of a wireless sensor network with the presence of an attacker.

At the FC, it reconstructs a minimum mean square error (MMSE) estimate of the physical quantity observed. Hence the mean squared error (MSE) or *distortion* at the FC is derived as

$$D[t] = \left(\frac{1}{\sigma_\theta^2} + \frac{g_S[t]p_S[t]}{g_S[t]p_S[t]\sigma_\omega^2 + g_A[t]p_A[t] + \sigma_n^2}\right)^{-1}, \tag{5.2}$$

where $g_S[t] \triangleq h_S[t]^{\mathrm{H}}h_S[t] \in \mathbb{R}$ and $g_A[t] \triangleq h_A[t]^{\mathrm{H}}h_A[t] \in \mathbb{R}$ are respectively the channel power gains from the sensor and the attacker to the FC, and $p_S[t] \triangleq \beta_S[t]^{\mathrm{H}}\beta_S[t] \in \mathbb{R}$ is the power scaling factors at the sensor and $p_A[t] \triangleq \beta_A[t]^{\mathrm{H}}\beta_A[t] \in \mathbb{R}$ is the jamming signal power from the attacker[5].

Remark: Notice that $p_S[t]$ and $p_A[t]$ are real valued, which means that for a given $p_S[t]$, any $\beta_S[t]$ satisfying $\beta_S[t]^{\mathrm{H}}\beta_S[t] = p_S[t]$ would result in the same distortion (similar

[5]As the sensor forwards the observation signal x with amplify factor β, the transmit power at time t is $(\sigma_\theta^2 + \sigma_\omega^2)\, p_S[t]$; whereas it is $p_A[t]$ for the attacker.

results applied to the attacker that any $\beta_A[t]$ satisfying $\beta_A[t]^{\mathrm{H}}\beta_A[t] = p_A[t]$ lead to the same effects). Therefore, we mainly focus on $p_S[t]$ and $p_A[t]$ in the rest of the work.

Depending on the channel knowledge at each player, we study how to use game theory to analyze the transmission strategy at the sensor and the attacker in a Nash equilibrium, where each player's choices are the best response to the other player's choices. We assume both players are rational, meaning they will always choose the action that offers them the highest expected payoff. Therefore, throughout this chapter, our objective is to obtain the 'optimal' strategies for both players such that no one gains more benefits by unilateral deviation. Before proceeding further, we define the following terms:

- **Players**: The players are the sensor node and the attacker.

- **Action sets**: The action sets are a collection of power levels that can be used to transmit, which are denoted by $\mathbb{P}_S$ for the sensor and $\mathbb{P}_A$ for the attacker. If there are infinitely many elements in an action set, it is called infinite action set; otherwise it is a finite action set.

- **Strategies**: A strategy is defined as a plan of actions intended to complete a specific goal, which are the power control policies in our case. A pure strategy for a player is a deterministic plan of action and is denoted by $\mathbf{p}_S$ for the sensor and $\mathbf{p}_A$ for the attacker; and we use $\boldsymbol{\pi}_S$ and $\boldsymbol{\pi}_A$ to indicate mixed strategies for the sensor and attacker, which are defined as a set of probability distributions over the corresponding pure strategies [95]. For a single-act game, a pure strategy consists of only one element from the action set. Whereas in repeated games, $\mathbf{p}_S$ is collection of actions for each stage of the game defined as $\mathbf{p}_S \triangleq [p_{S_0}, \ldots, p_{S_T}]$ where T is the time horizon and $p_{S_t} \in \mathbb{P}_A$. As for the mixed strategy, $\boldsymbol{\pi}_S \triangleq [\pi_{S0}, \ldots, \pi_{ST}]$ with π_{St} being a probability distribution over the action set $\mathbb{P}_S$ and interpreted as the decision rules at time t. Similar definitions apply to the attacker.

- **Payoffs**: A payoff or reward is the objective of the players. If the sum of players' payoffs is zero, we have a zero-sum game.

5.3 AWGN Channels

In this part of the work, we look at additive white Gaussian noise (AWGN) channels where the channel states of both the sensor and attacker are time invariant. In such environments, the channel states depend only on the distance between the transmitter and the receiver; hence it does not suffer from fading or shadowing. The AWGN channel models accurately describe deep space communication links and satellite sensing channels [68]. In such environments, we consider repeated games, where the same static game is played many times. In our setup, we consider both parties having imperfect information of each other, meaning that the players play simultaneously and they do not know the opponent's action until the game is completed, but they can observe the other's immediate payoffs and actions taken previously. We assume that the players are sophisticated enough that, instead of simply repeating the same static strategy, they are able to link their behavior in a given game to the outcomes of previous games. The idea is that each player may learn the opponent's strategies, hence to update its own strategies to counteract the actions of its opponent[6].

We first investigate finitely repeated games where the same static game is repeated over a finite number of periods. In this game, the payoff function is defined by looking at the overall system performance across the whole time horizon. For infinitely repeated games, in order to quantify the performance over an infinite time-horizon, we employ the discounted sum of rewards as payoff functions.

5.3.1 Finitely Repeated Games with Infinite Action Sets

First, we look at scenarios where both players' action sets comprise an infinite number of elements. We consider the repeated game with both the attacker and the sensor having an average power constraint over a finite time horizon. For the sensor, the goal is to minimize the overall distortion at the FC within a power budget $\mathcal{P}_{S\_\text{avg}}$; whereas the attacker wants to find the best attacking strategy without violating an average power constraint $\mathcal{P}_{A\_\text{avg}}$.

[6]Learning techniques used in this section such as backward induction and Nash Q-learning are implemented by each player offline.

As both the sensor and the attacker act simultaneously at each time, for the sensor and the attacker the optimization problems are

$$\begin{aligned}
&\texttt{Sensor's Game}: \\
&\min_{\{p_{S_t}\}} \max_{\{p_{A_t}\}} \sum_{t=0}^{T} D\left(p_{S_t}, p_{A_t}\right) \\
&\quad s.t. \quad \frac{1}{T}\sum_{t=1}^{T} p_{S_t} = \frac{\mathcal{P}_{S_\text{avg}}}{\sigma_\omega^2 + \sigma_\theta^2}, \\
&\qquad \frac{1}{T}\sum_{t=1}^{T} p_{A_t} = \mathcal{P}_{A_\text{avg}},
\end{aligned} \tag{5.3}$$

$$\begin{aligned}
&\texttt{Attacker's Game}: \\
&\max_{\{p_{A_t}\}} \min_{\{p_{S_t}\}} \sum_{t=0}^{T} D\left(p_{S_t}, p_{A_t}\right) \\
&\quad s.t. \quad \frac{1}{T}\sum_{t=1}^{T} p_{S_t} = \frac{\mathcal{P}_{S_\text{avg}}}{\sigma_\omega^2 + \sigma_\theta^2}, \\
&\qquad \frac{1}{T}\sum_{t=1}^{T} p_{A_t} = \mathcal{P}_{A_\text{avg}}.
\end{aligned} \tag{5.4}$$

Remark: It is noticeable, from (5.2), that the distortion level at the FC is a convex function over p_S that gradually decreases to the minimum value of $\frac{\sigma_\omega^2\sigma_\theta^2}{\sigma_\omega^2+\sigma_\theta^2}$ as the power p_S approaches infinity, and that it reaches the maximum distortion σ_θ^2 when the sensor stops transmitting. Similar features can be found for D w.r.t. p_A. The sensor intends to have high quality estimates at the FC (or a small MSE); whereas, the attacker attempts to lower the quality of the estimates, with both sides having a transmission power budget.

It is straightforward to verify that the objective functions are continuous for the game framework described in (5.3) and (5.4). Because $\mathbf{p}_S$ is compact and convex over the feasible region defined in (5.3) and $\sum_{t=1}^{T} D\left(p_{S_t}, p_{A_t}\right)$ is convex in $\mathbf{p}_S$ (similar results can be observed for (5.4) regarding $\mathbf{p}_A$), we conclude that the game defined in (5.3) and (5.4) has

a pure-strategy Nash equilibrium, i.e.,

$$\min_{\{p_{S_t}\}} \max_{\{p_{A_t}\}} \sum_{t=1}^{T} D\left(p_{S_t}, p_{A_t}\right) = \max_{\{p_{A_t}\}} \min_{\{p_{S_t}\}} \sum_{t=1}^{T} D\left(p_{S_t}, p_{A_t}\right).$$

Let $\{p_{S_t}^*\}$ be the optimal solution of problem (5.3), and $\{p_{A_t}^*\}$ the optimal solution of problem (5.4). A pure strategy pair $\left\{\mathbf{p}_S^*, \mathbf{p}_A^*\right\}$ at a Nash equilibrium results in

$$\sum_{t=0}^{T} D\left(p_{S_t}^*, p'_{A_t}\right) \leq \sum_{t=0}^{T} D\left(p_{S_t}^*, p_{A_t}^*\right) \leq \sum_{t=1}^{T} D\left(p'_{S_t}, p_{A_t}^*\right),$$

for all $\mathbf{p}'_S \in \left\{\mathbf{p}_S : \frac{1}{T}\sum_{t=0}^{T} p_{S_t} = \frac{\mathcal{P}_{S\_\text{avg}}}{\sigma_\omega^2+\sigma_\theta^2}\right\}$ and $\mathbf{p}'_A \in \left\{\mathbf{p}_A : \frac{1}{T}\sum_{t=0}^{T} p_{A_t} = \mathcal{P}_{A\_\text{avg}}\right\}$.

To obtain $\left\{\mathbf{p}_S^*, \mathbf{p}_A^*\right\}$, that is to solve the constrained nonlinear minimax problems, algorithms proposed in [44] or primal-dual interior-point method [75] can be used; however, a high computational cost is generally involved in deriving the results. Furthermore, in the power control architectures of wireless communication, it is more common to employ quantized power levels to the mobile terminals. Therefore, in favor of the low complexity and being easy to implement in real setups, for the rest of this section, we assume that both the sensor and the attacker only have a finite number of power levels (instead of an infinite amount of power levels) to adapt their power transmission strategy at each time slot.

5.3.2 Finitely Repeated Games with Finite Action Sets

In cases where players have finitely many actions from which to choose, if the action size is small, the game may not admit a Nash equilibrium solution in pure strategies. One way to obtain an equilibrium solution is to enlarge the strategy space [8]. This leads to mixed strategies, which are probability distributions over the finite pure strategy sets.

In finitely repeated games, we also assume that both the sensor and the attacker move simultaneously at each time point, and both of them are able to observe the outcome from the last time step before acting. In other words, at each time, each player has information concerning the current time of play, and each player knows the state of the game (i.e. the outcomes from last move) at every time of play.

Recall that π_{St} and π_{At} stand for the decision rules for the sensor and the attacker at time t respectively, from which the corresponding actions at time t can be determined and denoted by $p^1_{\pi_{St}} \in \mathbb{P}_S$ and $p^2_{\pi_{At}} \in \mathbb{P}_A$ respectively. Our goal is to find a mixed strategy pair $\left\{\boldsymbol{\pi}^*_S, \boldsymbol{\pi}^*_A\right\}$ at a Nash equilibrium, where $\boldsymbol{\pi}^*_S = \left[\pi^*_{S0}, \ldots, \pi^*_{ST}\right]$ and $\boldsymbol{\pi}^*_A = \left[\pi^*_{A0}, \ldots, \pi^*_{AT}\right]$, such that after playing the game T times, the overall estimation distortion is minimized at the sensor; while it is maximized at the attacker with the average power consumption being less than the budget at both the sensor and the attacker, given as $\frac{1}{T}\sum_{t=0}^{T}\mathbb{E}\left[p^1_{\pi^*_{St}}\right] \leq \frac{\mathcal{P}_{S\_\text{avg}}}{\sigma^2_\theta+\sigma^2_\omega}$ and $\frac{1}{T}\sum_{t=0}^{T}\mathbb{E}\left[p^2_{\pi^*_{At}}\right] \leq \mathcal{P}_{A\_\text{avg}}$[7].

To address the problem, we introduce a reward function for the sensor, given as

$$R_S\left(\boldsymbol{\pi}_S, \boldsymbol{\pi}_A\right) = \sum_{t=0}^{T}\mathbb{E}\left[-D\left(p^1_{\pi_{St}}, p^2_{\pi_{At}}\right) - \lambda p^1_{\pi_{St}} + \nu p^2_{\pi_{At}}\right]. \tag{5.5}$$

Alternatively, it can be expressed as

$$R_S\left(\boldsymbol{\pi}_S, \boldsymbol{\pi}_A\right) = \sum_{t=0}^{T}\left(\sum_{p_1\in\mathbb{P}_S}\sum_{p_2\in\mathbb{P}_A}\left[-D\left(p_1,p_2\right) - \lambda p_1 + \nu p_2\right]\pi_{St}\left(p_1\right)\pi_{At}\left(p_2\right)\right), \tag{5.6}$$

where λ and ν control the average power consumption for the sensor and the attacker respectively. The payoff function at the attacker is defined as $R_A \triangleq -R_S$ for the zero-sum game we consider. The following algorithm is applied to derive a best strategy pair $\left\{\boldsymbol{\pi}^*_S, \boldsymbol{\pi}^*_A\right\}$ at a Nash equilibrium point that satisfies the power constraints.

[7]The expectation is taking across all the power realizations produced according to $\left\{\pi^*_{St}, \pi^*_{At}\right\}$ at time t. $\mathbb{E}\left[D\left(p^1_{\pi^*_{St}}, p^2_{\pi^*_{At}}\right)\right]$ can be expressed as $\mathbb{E}\left[D\left(p^1_{\pi^*_{St}}, p^2_{\pi^*_{At}}\right)\right] = \sum_{p_1\in\mathbb{P}_S}\sum_{p_2\in\mathbb{P}_A} D\left(p_1,p_2\right)\pi^*_{St}\left(p_1\right)\pi^*_{At}\left(p_2\right)$. Similar expressions can be derived for $\frac{1}{T}\sum_{t=0}^{T}\mathbb{E}\left[p^1_{\pi^*_{St}}\right]$ and $\frac{1}{T}\sum_{t=0}^{T}\mathbb{E}\left[p^2_{\pi^*_{At}}\right]$.

Algorithm 2 Zero-Sum Game with Power Constraint

1: Initialize λ.
2: **repeat**
3: Initialize ν.
4: **repeat**
5: Recursively obtain a mixed strategy pair $\left\{\boldsymbol{\pi}_S^*(\lambda,\nu), \boldsymbol{\pi}_A^*(\lambda,\nu)\right\}$.
6: Apply bisection method to update ν.
7: **until** convergence: Average power consumption at the attacker is close to $\mathcal{P}_{A\_\text{avg}}$.
8: Apply bisection method to update λ.
9: **until** convergence: Average power consumption at the sensor is close to $\mathcal{P}_{S\_\text{avg}}$.

To be more specific about the recursive procedure[8] [8] in Step 5 of Algorithm 2 above, we first start at the last stage of play, T, solve each single-act game and derive the 'optimal' power policies corresponding to each outcome from $T-1$. This results in a mixed Nash equilibrium strategy pair $\{\pi_{S_T}^*(\lambda,\nu), \pi_{A_T}^*(\lambda,\nu)\}$. Next, cross out the Tth level of play, and consider the resulting $T-1$ level games and find the strategy pair $\left\{\pi_{S_{T-1}}^*(\lambda,\nu), \pi_{A_{T-1}}^*(\lambda,\nu)\right\}$. Repeat the same procedure until we reach $t=0$, to obtain $\{\pi_{S_0}^*(\lambda,\nu), \pi_{A_0}^*(\lambda,\nu)\}$. Then, we could say that given λ and ν, the zero-sum feedback game in extensive form admits a saddle-point solution $\left\{\boldsymbol{\pi}_S^*(\lambda,\nu), \boldsymbol{\pi}_A^*(\lambda,\nu)\right\}$.

Remark: The entities in $\{\pi_{S_t}^*(\lambda,\nu), \pi_{A_t}^*(\lambda,\nu)\}$ depend on the path from stage 0 to stage t; hence, at t, there are a total number of $(|\mathbb{P}_S| \times |\mathbb{P}_A|)^t$ matrix games, where $|\mathbb{P}_S|$ and $|\mathbb{P}_A|$ are the size of action spaces of $\mathbb{P}_S$ and $\mathbb{P}_A$. Therefore, by walking through all the matrix games on level t we are able to find $\{\pi_{S_t}^*(\lambda,\nu), \pi_{A_t}^*(\lambda,\nu)\}$.

5.3.3 Infinitely Repeated Games with Finite Action Sets

In this subsection, we focus on scenarios where the same game is repeatedly played over an infinite time horizon. Different from the games considered previously, we look at an optimal stationary strategy for infinitely repeated games, meaning the best decision rules are fixed over time, i.e., $\pi_{S_t} = \pi_S$ and $\pi_{A_t} = \pi_A$ for all t. This allows players to update

[8] This optimization procedure is also know as dynamic programming or backward induction.

their strategy based on its own information and the information collected from the last time step, until the optimal one is found. To simplify the notation, we replace the strategy pair $\{\boldsymbol{\pi}_S, \boldsymbol{\pi}_A\}$ by $\{\pi_S, \pi_A\}$ for a stationary game.

We use a similar setup as in Section 5.3.2. Assuming that both players move simultaneously and are capable of observing the payoffs from last step, we first set up the game framework. Next, we reach a Nash equilibrium based on the 'interaction' between the sensor and attacker in infinite time horizon, and obtain stationary strategies for both players.

The **Players** are the sensor and the attacker, and the **Action** of each of the players is the transmit power, which are denoted by $p_{S_t} \in \mathbb{P}_S \triangleq [0, \ldots, p_{S_{\max}}]$ and $p_{A_t} \in \mathbb{P}_A \triangleq [0, \ldots, p_{A_{\max}}]$ at time step t. In infinite horizon, we define the **Payoff** or **Reward** for the sensor at the time step t as

$$R_t\left(p_{S_t}, p_{A_t}\right) = -D\left(p_{S_t}, p_{A_t}\right) - \lambda p_{S_t} + \nu p_{A_t}, \tag{5.7}$$

and it is $-R_t\left(p_{S_t}, p_{A_t}\right)$ for the attacker, where λ and ν are non-negative weight parameters as the sensor is seeking to minimize the estimate distortion at the FC by consuming as little power as possible. By contrast, the attacker wants to reduce the estimation quality at the FC. At each time step t, the sensor and the attacker take actions simultaneously and then respectively receive reward R_t and $-R_t$.

Next, we define a function J_S for the sensor to quantify the estimation quality over an infinite time-horizon as a discounted sum of rewards, namely

$$J_S\left(\pi_S, \pi_A\right) \triangleq \sum_{t=0}^{\infty} \gamma^t R_t, \tag{5.8}$$

where $\gamma \in [0, 1)$ is the discount factor. $J_S\left(\pi_S, \pi_A\right)$ can be rewritten as

$$J_S\left(\pi_S, \pi_A\right) = R\left(p_{S_{\pi_S}}, p_{A_{\pi_A}}\right) + \gamma J_S\left(\pi_S, \pi_A\right), \tag{5.9}$$

where $p_{S_\pi} \in \mathbb{P}_S$ and $p_{A_{\pi_A}} \in \mathbb{P}_A$ are determined by the decision rules π_S and π_A.

Remark: In order for the payoff function in (5.8) to be well defined, we should have

the discount factor γ being less than 1; otherwise the $J_S(\pi_S, \pi_A)$ may approach infinity [95]. The discount factor can be interpreted as the uncertainty towards the future, hence players would weight more to the immediately rewards than payoffs in the future.

As both players are rational, the objective of the sensor is to maximize the discounted sum of rewards $J_S(\pi_S, \pi_A)$; and the attacker also wants to maximize the discounted sum of its rewards, defined as

$$J_A(\pi_S, \pi_A) \triangleq -J_S(\pi_S, \pi_A). \tag{5.10}$$

Definition 1: In this zero-sum repeated game, a Nash equilibrium point is a pair of strategies $\{\pi_S^*, \pi_A^*\}$ such that

$$J_S(\pi_S^*, \pi_A^*) \geq J_S(\pi_S, \pi_A^*), \ \forall \pi_S,$$

and

$$J_A(\pi_S^*, \pi_A^*) \geq J_A(\pi_S^*, \pi_A), \ \forall \pi_A.$$

In order to obtain the strategy pair $\{\pi_S^*, \pi_A^*\}$ at the Nash equilibrium, we employ the Nash Q-learning algorithm [38,39], which is one of the model-free reinforcement learning algorithms. It enable players to interact with the environment in incomplete information scenarios, such as when players are not aware of the opponent's payoff functions, thus obtaining the optimal policy at a Nash equilibrium. In our setup, players have complete but imperfect information towards the others. Although algorithms developed in [19] can be applied to obtain an Nash equilibrium strategy, we pick the Nash Q-learning algorithm as it allows us to extend the work to the incomplete information cases in future.

We first define the optimal Q-value at the sensor as

$$\begin{aligned} &Q_S^*(p_1, p_2) \\ \triangleq &R(p_1, p_2) + \gamma \max_{\pi_S} \min_{\pi_A} \sum_{p_1' \in \mathbb{P}_S} \sum_{p_2' \in \mathbb{P}_A} Q_S(p_1', p_2') \pi_S(p_1') \pi_A(p_2'), \end{aligned} \tag{5.11}$$

where $p_1 \in \mathbb{P}_S$, $p_2 \in \mathbb{P}_A$, and $\pi_S(p_1')$, $\pi_A(p_2')$ are the probabilities of choosing the action

p_1' in strategy π_S at the sensor and choosing the action p_2' in strategy π_A at the attacker respectively. The Q-value at the attacker is exactly the opposite for the zero-sum game, which is defined later in (5.12).

$Q_S^*(p_1, p_2)$ can be seen as the expected reward for the sensor taking action p_1 and the attacker executing action p_2, and then following the optimal policy thereafter. In addition, when we know the optimal Q-values, the optimal policies π_S^* can be easily found.

To obtain the optimal Q-values, the sensor needs to maintain and keep updating one $|\mathbb{P}_S| \times |\mathbb{P}_A|$ size table[9], with each element corresponding to a Q-value when adopting a specific transmission power (p_1, p_2) at the players. This is possible since we assume the sensor knows the attacker's action set $\mathbb{P}_A$. Similarly, the attacker computes the optimal Q-values by updating a $|\mathbb{P}_S|$ by $|\mathbb{P}_A|$ size table with each of the entries defined as

$$\begin{aligned} & Q_A^*(p_1, p_2) \\ \triangleq & -R(p_1, p_2) - \gamma \max_{\pi_A} \min_{\pi_S} \sum_{p_1' \in \mathbb{P}_S} \sum_{p_2' \in \mathbb{P}_A} Q_A(p_1', p_2') \pi_S(p_1') \pi_A(p_2'). \end{aligned} \tag{5.12}$$

Because we have a zero-sum game, it is not hard to see that $Q_A = -Q_S$. To simplify the notation, we denote Q_S by Q for the rest of this subsection.

To be more specific about the Q-learning procedure, the players first initialize the values of $Q(p_1, p_2)$ for all $p_1 \in \mathbb{P}_S$ and $p_2 \in \mathbb{P}_A$. At each iteration, an action is chosen, R_t obtained, and the corresponding element in the table is updated based on the following equation:

$$\begin{aligned} Q_{t+1}(p_1, p_2) = & (1 - \alpha_t) Q_t(p_1, p_2) + \alpha_t \Big[R_t(p_1, p_2) \\ & + \max_{\pi_S} \min_{\pi_A} \gamma \sum_{p_1' \in \mathbb{P}_S} \sum_{p_2' \in \mathbb{P}_A} Q_t(p_1', p_2') \pi_S(p_1') \pi_A(p_2') \Big], \end{aligned} \tag{5.13}$$

where $\alpha_t \in [0, 1)$ is the learning rate that decays over time. The Q-learning algorithm depicted in (5.13) is guaranteed to converge to the optimal Q-value as long as [38, 39]:

[9] This is true for zero-sum games, for a general-sum game each player needs to keep two tables for updating, one for each player.

- The learning rate α_t satisfies $\sum_{t=0}^{\infty} \alpha_t = +\infty$ and $\sum_{t=0}^{\infty} \alpha_t^2 < +\infty$.

- Every action has been visited infinitely often.

Hence, the 'optimal' stationary strategy pair $\left\{\pi_S^*, \pi_A^*\right\}$ is found.

5.4 Continuous Fading Channels

Starting from this section, we investigate 'optimal' power allocation policies at a Nash equilibrium in block fading channel scenarios, where the channel states of both the sensor and the attacker may vary in different fading blocks. In this part of the work, we also consider a non-cooperative game with each player having complete but imperfect information; in order words, the channel state information is known everywhere but players are unclear about its opponent's strategies, thus the optimal power policies at a Nash equilibrium depends largely on one's guess toward the others. In the game, the strategy of the sensor is the power control policy p_S, and it is p_A for the attacker. With complete information at both players, p_S and p_A are dependent on $G \triangleq [g_S, g_A]^{\mathrm{T}}$.

In the case of fading channels, we assume that the crucial information lies in the long-term behavior of the estimates, thus it would be more meaningful to set $-\mathbb{E}_G\left[D\left(p_S, p_A\right)\right]$ as the payoff function that the sensor attempts to maximize, while the attacker wants to enlarge its payoff function $\mathbb{E}_G\left[D\left(p_S, p_A\right)\right]$.

It is clear that the payoff function of each player depends on the power policy pair $\{p_S, p_A\}$ with $p_S \in \mathbb{P}_S \triangleq \left\{p_S : \left(\sigma_\omega^2 + \sigma_\theta^2\right) \mathbb{E}_G\left[p_S\left(G\right)\right] \leq \mathcal{P}_{S\_\mathrm{avg}}, p_S\left(G\right) \geq 0\right\}$ and $p_A \in \mathbb{P}_A \triangleq \left\{p_A : \mathbb{E}_G\left[p_A\left(G\right)\right] \leq \mathcal{P}_{\mathrm{A\_avg}}, p_A\left(G\right) \geq 0\right\}$. Given a fixed power scheme of the attacker, the sensor's optimal strategy is derived by solving the following optimization problem:

$$\begin{aligned} &\min_{p_S(G)} \int\int_G \left(\frac{1}{\sigma_\theta^2} + \frac{g_S p_S\left(G\right)}{g_S p_S\left(G\right) \sigma_\omega^2 + g_A p_A\left(G\right) + \sigma_n^2}\right)^{-1} dF(g_S) dF(g_A) \\ &s.t. \quad \int\int_G p_S\left(G\right) dF(g_S) dF(g_A) \leq \frac{\mathcal{P}_{S\_\mathrm{avg}}}{\sigma_\omega^2 + \sigma_\theta^2}, \;\; p_S\left(G\right) \geq 0. \end{aligned} \tag{5.14}$$

The solution is

$$p_S(G) = \left(\frac{\sqrt{\frac{g_S(\sigma_n^2 + p_A(G) g_A)}{\lambda}} - \frac{\sigma_n^2 + p_A(G) g_A}{\sigma_\theta^2}}{g_S \left(\frac{\sigma_\omega^2}{\sigma_\theta^2} + 1 \right)} \right)^+, \tag{5.15}$$

in which $(x)^+ = \max\{x, 0\}$ and λ is the power level satisfying the power constraint in (5.14) with equality. When the sensor transmits, i.e., $p_S(G) > 0$, we obtain $\frac{g_S}{\sigma_n^2 + p_A(G) g_A} > \frac{\lambda^2}{\sigma_\theta^4}$. Since the sensor treats the attacker's signal as a part of the noise in the channel, intuitively, this says that the sensor would start to transmit as long as its Signal-to-Noise Ratio (SNR) lies beyond a threshold.

Given $p_S(G)$, the optimal strategy of the attacker $p_A(G)$ can be obtained in a similar manner; given as

$$p_A(G) = \left(\sqrt{\frac{g_S p_S(G) \sigma_\theta^4 g_A}{\nu g_A^2}} - \frac{\sigma_n^2 + g_S p_S(G) \left(\sigma_\omega^2 + \sigma_\theta^2 \right)}{g_A} \right)^+, \tag{5.16}$$

where ν is the power level satisfying its average power constraint with equality. From (5.16), we can see that $p_S(G) = 0$ results in $p_A(G) = 0$, which makes sense. It means that when the sensor keeps silent or is not communicating with the FC, there is no need for the attacker to jam the channel.

From the expression of the players' best strategy, one can see that the optimal power policy of one player is coupled with the other's. This means that before determining its own policy the sensor has to guess the attacker's transmitting policy; and vice versa. Before moving further, we have the Nash Equilibrium defined as follows.

Definition 2 [8]: A Nash Equilibrium point is a power policy pair $\left\{ p_S^*, p_A^* \right\}$ such that

$$\mathbb{E}\left[D\left(p_S^*, p_A' \right) \right] \leq \mathbb{E}\left[D\left(p_S^*, p_A^* \right) \right] \leq \mathbb{E}\left[D\left(p_S', p_A^* \right) \right] \tag{5.17}$$

with $p_S^*, p_S' \in \mathbb{P}_S$ and $p_A^*, p_A' \in \mathbb{P}_A$.

At a Nash equilibrium, no player can benefit from moving away from the Nash equilibrium point.

Theorem 1: There exists at least one Nash equilibrium for the static non-cooperative

game with complete information.

The proof is not difficult. From (5.2), we see that the payoff function is continuous in both p_S and p_A, and $D(p_S, p_A)$ is convex in p_S for any p_A (same for $D(p_S, p_A)$ with respect to p_A for any given p_S). Together with the fact that the strategy spaces of both players are convex, compact and nonempty, the existence is then proved [21].

5.5 Discrete Fading Channels with Finite Action Sets

Wireless communication channels are time-varying multiple-path channels, which have a strong effect on the performance of digital radio communication systems. When a multi-dimensional distribution, considering the number of multipath components, SNR values, multipath delay, etc, is used to characterize the channel fading statistics, a high computational cost is required in simulation, as a large number of samples need to be drawn in order to produce statistically matched channel gains [43]. To reduce the complexity, finite-state channel in introduced, where the continuous channel gains are quantized into a number of intervals or channel states with probability of being on each state determined by the continuous fading distribution [86]. Such channel model was considered by Goldsmith in [28] and Wolfowitz to derive the time-varying channel capacity in [100].

Therefore, in this section we assume that the sensor is only aware of discrete channel states taking values on a finite set of discrete memoryless channels. We also assume both the sensor and the attacker have a finite number of power levels to adapt their transmission power. In fact, as stated in [83], a quantized power is more frequently used in cellular networks or wireless networks than actual power values. Therefore, the power sets are quantized into finite levels, and $\mathbb{P}_S$ and $\mathbb{P}_A$ consist of a finite number of elements.

As in previous sections, our objective in this part is to investigate the 'best' power allocation policies for both players at a Nash equilibrium, where the sensor and attacker are constrained to the limited power budget $\mathcal{P}_{S\_\text{avg}}$ and $\mathcal{P}_{A\_\text{avg}}$ respectively. We start with an imperfect information scenario, where both parties move simultaneously. Next, we look at the case of incomplete information, in which we assume players only have perfect knowledge about its own channel conditions. For both cases, the optimal strategy

pair $\left\{\pi_S^*, \pi_A^*\right\}$ is discussed. Since both the sensor and attacker have finite action set, the game may not admit a Nash equilibrium in pure strategies. Instead, we look for mixed strategies, which always exists for any finite bimatrix games [72].

5.5.1 Complete Information

First, we consider a complete information scenario. In this case, the sensor and the attacker adjust their strategies depending on the channel status of both players.

In a zero-sum game, as one player's payoff is always negative of the other, we denote R as the payoff function of the attacker, given as

$$R\left(p_S, p_A\right) = D\left(p_S, p_A\right) + \lambda p_S - \nu p_A, \tag{5.18}$$

where $p_S \in \mathbb{P}_S$, $p_A \in \mathbb{P}_A$, and λ and ν control average power consumption at the sensor and the attacker respectively.

Let $\{\pi_S, \pi_A\}$ be a mixed strategy pair. With complete information at the players, π_S and π_A are functions of both players' channel information; hence at each channel instance the sensor uses $\pi_S\left(G\right)$ and the attacker employs $\pi_A\left(G\right)$ to adjust the transmission power.

Definition 3: A mixed strategy at the Nash equilibrium is a pair of $\left\{\pi_S^*, \pi_A^*\right\}$ such that

$$\begin{aligned}\mathbb{E}_G\left[\mathbb{E}\left[R\left(\pi_S^*\left(G\right), \pi_A'\left(G\right)\right)\middle| G\right]\right] &\leq \mathbb{E}_G\left[\mathbb{E}\left[R\left(\pi_S^*\left(G\right), \pi_A^*\left(G\right)\right)\middle| G\right]\right] \\ &\leq \mathbb{E}_G\left[\mathbb{E}\left[R\left(\pi_S'\left(G\right), \pi_A^*\left(G\right)\right)\middle| G\right]\right],\end{aligned} \tag{5.19}$$

where $\pi_S', \pi_S^* \in \left\{\pi_S : \mathbb{E}_G\left[\sum_{p_1\in\mathbb{P}_S} p_1 \pi_S\left(p_1\right)\right] \leq \frac{\mathcal{P}_{S\_\text{avg}}}{\sigma_\theta^2+\sigma_\omega^2}\right\}$; $\pi_A', \pi_A^* \in \left\{\pi_A : \mathbb{E}_G\left[\sum_{p_2\in\mathbb{P}_A} p_2 \pi_A\left(p_2\right)\right] \leq \mathcal{P}_{A\_\text{avg}}\right\}$ and $\mathbb{E}\left[R\left(\pi_S, \pi_A\right)\middle| G\right] \triangleq \sum_{p_1\in\mathbb{P}_S}\sum_{p_2\in\mathbb{P}_A} D\left(p_1, p_2\right)\pi_S\left(p_1\right)\pi_A\left(p_2\right)$ is the expected payoff at channel state G when the sensor and the attacker adopt mixed strategies π_S and π_A respectively.

At a Nash equilibrium, the two power constraints are satisfied with equality. One can employ the following iterative method to obtain $\left\{\pi_S^*, \pi_A^*\right\}$. First, initialize $\lambda^{(1)}$ and then update $\nu^{(1)}$ such that $\mathbb{E}_G\left[\sum_{p_2\in\mathbb{P}_A} p_2 \pi_A^{(1)}\left(p_2\right)\right] = \mathcal{P}_{A\_\text{avg}}$. This can be done by solving a series of matrix games via linear programming. Next, we adjust $\lambda^{(2)}$ to satisfy

$\mathbb{E}_G\left[\sum_{p_1\in\mathbb{P}_S} p_1\pi_S^{(2)}(p_1)\right] = \frac{\mathcal{P}_{S\_\text{avg}}}{\sigma_\theta^2+\sigma_\omega^2}$. Repeat this process until $\mathbb{E}_G\left[\sum_{p_1\in\mathbb{P}_S} p_1\pi_S^{(n)}(p_1)\right] = \frac{\mathcal{P}_{S\_\text{avg}}}{\sigma_\theta^2+\sigma_\omega^2}$ and $\mathbb{E}_G\left[\sum_{p_2\in\mathbb{P}_A} p_2\pi_A^{(n)}(p_2)\right] = \mathcal{P}_{A\_\text{avg}}$ are achieved simultaneously; this gives a mixed strategy pair $\left\{\pi_S^*, \pi_A^*\right\}$ at a Nash equilibrium.

Because each player has complete channel information, the whole approach of deriving $\left\{\pi_S^*, \pi_A^*\right\}$ can be completed locally. When transmission starts at the sensor, depending on the channel state, both players would operate using $\left\{\pi_S^*, \pi_A^*\right\}$ to obtain transmit powers.

5.5.2 Incomplete Information - Bayesian Games

The assumption that both players have complete channel information is difficult to achieve for real setups. Therefore, in this subsection, we model the DoS attack as a Bayesian game, where both players have incomplete information, meaning that the sensor knows its own channel gain g_S, but does not know the channel gains of its opponent, g_A. Similar assumptions apply to the attacker. We further assume that the channel gains are drawn from a fixed distribution that is common knowledge to both parties. Hence, in order to obtain the optimal power allocation, the sensor or the attacker has to adjust its power level based on its own channel gains, in other words its *type*. The *type* of a player is any private information that is not common knowledge to others but is relevant to the player's decision marking [21].

The DoS attack Bayesian game consists of the following elements:

- A set of players: $\{S,\ A\}$, S represents the sensor and A denotes the attacker.
- A set of actions: $\{\mathbb{P}_S, \mathbb{P}_A\}$, $\mathbb{P}_S$ for the sensor and $\mathbb{P}_A$ for the attacker.
- A set of types: $\{\mathbb{G}_S, \mathbb{G}_A\}$, $\mathbb{G}_S$ for the sensor and $\mathbb{G}_A$ for the attacker, which are defined as the different channel power gains for each player, i.e., $g_S \in \mathbb{G}_S$ and $g_A \in \mathbb{G}_A$.
- A set of probability functions: $\{\text{Pr}_S, \text{Pr}_A\}$, Pr_S is the attacker's belief about the type of the sensor; whereas Pr_A is the belief for the attacker's type at the sensor.

- Payoff functions: R_S for the sensor and R_A for the attacker, where $R_S = -R_A$ for the zero-sum game we consider.

In Bayesian games, a pure strategy for a player is a function, $p_S(g_S)$ for the sensor and $p_A(g_A)$ for the attacker, mapping from its type set to the action set, which specifies a pure action that the player will choose when a particular type is observed. In other words, each player knows his type-contingent strategy before he learns his type and then plays the game accordingly.

For player i, where $i \in \{S,\ A\}$, denote g_{-i} as the type of its opponent and P_i as the collection of functions $p_i : \mathbb{G}_i \to \mathbb{P}_i$.

Definition 4 [21]: Given a strategy $p_i(g_i)$ and $p_i'(g_i) \in P_i$, the strategy profile $\{p_i^*, p_{-i}^*\}$ is a pure-strategy Bayesian Nash equilibrium if, for each player $i \in \{S,\ A\}$ and every $g_i \in \mathbb{G}_i$,

$$p_i^*(g_i) = \arg\max_{p_i' \in P_i} \sum_{g_{-i}} R_i\left(p_i', p_{-i}^*(g_{-i}) \middle|\, g_i, g_{-i}\right) \Pr\left(g_{-i} \middle|\, g_i\right). \tag{5.20}$$

That is, regardless of the type realization, no player benefits from changing his strategy $p_i^*(g_i)$; in other words, a Bayesian Nash equilibrium is just a Nash equilibrium in a Bayesian game [95]. The extension to a best mixed strategy pair $\{\pi_S^*, \pi_A^*\}$ can be derived by mapping types into probability distribution over the action set, given as

$$\pi_i^*(g_i) = \arg\max_{\pi_i'} \sum_{g_{-i}} R_i\left(\pi_i', \pi_{-i}^*(g_{-i}) \middle|\, g_i, g_{-i}\right) \Pr\left(g_{-i} \middle|\, g_i\right), \quad \forall i \in \{S,\ A\}. \tag{5.21}$$

In order to solve problem (5.21), we first derive the matrix form of the zero-sum Bayesian game. More specifically, we find the expected payoff for every strategy of the sensor's against each opponent's strategy. For a player with type of size $|\mathbb{G}_S|$ and action set size $|\mathbb{P}_S|$, the strategy space is made up of $|\mathbb{G}_S|^{|\mathbb{P}_S|}$ entries, with each indicating the power schemes to be used for different types of channels. The payoffs are calculated at both the sensor and the attacker using the joint probabilities together with both players' strategies. This leads to a matrix game of size $|\mathbb{G}_S|^{|\mathbb{P}_S|} \times |\mathbb{G}_A|^{|\mathbb{P}_A|}$, which can

be solved via linear programming[10]. The solution of the zero-sum matrix game is then added across each type of both players giving strategies at the Bayesian Nash equilibrium, $\left\{\pi_S^*\left(g_S\right), \pi_A^*\left(g_A\right)\right\}$.

To accommodate the average power constraints $\mathcal{P}_{S\_\text{avg}}$ for the sensor and $\mathcal{P}_{A\_\text{avg}}$ for the attacker, we define the payoff function at the attacker as

$$R_A\left(p_S, p_A\right) = D\left(p_S, p_A\right) + \lambda p_S - \nu p_A, \tag{5.22}$$

where $p_S \in \mathbb{P}_S$ and $p_A \in \mathbb{P}_A$ are the transmit powers at the sensor and attacker respectively, and λ and ν are variables respectively guaranteeing $\mathbb{E}_{g_S}\left[\sum_{p_1 \in \mathbb{P}_S} p_1 \pi_S\left(p_1\right)\right] \leq \frac{\mathcal{P}_{S\_\text{avg}}}{\sigma_\theta^2+\sigma_\omega^2}$ and $\mathbb{E}_{g_A}\left[\sum_{p_2 \in \mathbb{P}_A} p_2 \pi_A\left(p_2\right)\right] \leq \mathcal{P}_{A\_\text{avg}}$, which can be solved numerically. More specifically, we first fix $\lambda^{(1)}$ and adjust ν to $\hat{\nu}$ such that the resulting mixed strategy pair $\left\{\pi_S^*\left(g_S\right), \pi_A^*\left(g_A\right)\right\}$ derived from solving a zero-sum matrix game satisfies $\mathbb{E}_{g_A}\left[\sum_{p_2 \in \mathbb{P}_A} p_2 \pi_A\left(p_2\right)\right] = \mathcal{P}_{A\_\text{avg}}$; we then set $\nu^{(1)} = \hat{\nu}$ and find the best λ satisfying $\mathbb{E}_{g_S}\left[\sum_{p_1 \in \mathbb{P}_S} p_1 \pi_S\left(p_1\right)\right] = \frac{\mathcal{P}_{S\_\text{avg}}}{\sigma_\theta^2+\sigma_\omega^2}$. We keep repeating the same procedure until $\mathbb{E}_{g_S}\left[\sum_{p_1 \in \mathbb{P}_S} p_1 \pi_S\left(p_1\right)\right] = \frac{\mathcal{P}_{S\_\text{avg}}}{\sigma_\omega^2+\sigma_\theta^2}$ and $\mathbb{E}_{g_A}\left[\sum_{p_2 \in \mathbb{P}_A} p_2 \pi_A\left(p_2\right)\right] = \mathcal{P}_{A\_\text{avg}}$ are satisfied simultaneously, and this gives the optimal strategies $\left\{\pi_S^*\left(g_S\right), \pi_A^*\left(g_A\right)\right\}$ at a Bayesian Nash equilibrium.

5.6 Numerical Results

In this section, we provide some numerical examples. For simplicity, we consider the source θ to be Gaussian distributed with zero mean and variance $\sigma_\theta^2 = 1$ mW. The sensor measurement sensitivity is set to $\sigma_\omega^2 = 10^{-3}$ mW. We assume the same noise level for both the FC and the eavesdropper's channel, where $\sigma_n^2 = \sigma_e^2 = 10^{-8}$ mW.

We first look at finitely repeated games in non-fading scenarios where the same static game is played four times with the channel power gains of the sensor and the attacker being 9.2×10^{-9} and 9.7×10^{-8} respectively. Assume that the sensor and the attacker have the same three transmission power levels, given as $\mathbb{P}_S = \mathbb{P}_A = [0.0, 0.5, 1.0]$, and the average power budget are 0.78mW and 1.56mW respectively. Interestingly, the mixed

[10]The Lemke-Howson algorithm [54] can be used to solve non-zero sum matrix games in mixed strategy.

strategies at a Nash equilibrium for the sensor and the attacker remain the same for all stage games. For instance, when $t = 1$ all nine elements[11] in $\{\pi^*_{S1}, \pi^*_{A1}\}$ are identical, given in Table 5.1.

$\{\pi^i_{S1}, \pi^i_{A1}\} \in \{\pi^*_{S1}, \pi^*_{A1}\}, i = 1, \ldots, 9$	
π^i_{S1}	π^i_{A1}
$\Pr(p_S = 0.0) = 0.608$	$\Pr(p_A = 0.0) = 0.214$
$\Pr(p_S = 0.5) = 0.392$	$\Pr(p_A = 0.5) = 0.786$
$\Pr(p_S = 1.0) = 0.000$	$\Pr(p_A = 1.0) = 0.000$

Table 5.1: Mixed strategies in a finitely repeated game at $t = 1$.

Unsurprisingly, similar simulation results are observed for infinitely repeated games, where the action sets and channel power gains are the same as the previous finite repeated game. As shown in Table 5.2, the strategy pair does not update providing the information collected from last step[12]. The simulation results for non-fading scenarios indicate that no useful information can be acquired by playing the same static game multiple times; in other words, each player's power allocation strategy remains unchanged. One possible explanation is that the random source observed by the sensor is i.i.d. over time, which implies that the static game played in one time slot is independent with other games. Therefore, looking at the games played in the past does not help players to choose the current power allocation strategies.

Iteration Index	$\{\pi_S, \pi_A\}$
0	{[0.333 0.333 0.333], [0.333 0.333 0.333]}
1	{[0.000 0.146 0.854], [0.000 0.722 0.278]}
2	{[0.000 0.146 0.854], [0.000 0.722 0.278]}
...	{..., ...}
100	{[0.000 0.146 0.854], [0.000 0.722 0.278]}

Table 5.2: Mixed strategies in an infinitely repeated game.

Next, we look at continuous fading channel scenarios, where we assume the distances from the sensor and attacker to the FC are 329m and 96m respectively. We also con-

[11] As both players have an action set of three elements, overall there are nine paths from $t = 0$ to $t = 1$ level of the game.

[12] The mixed strategy pair $\{\pi_S, \pi_A\}$ at Iteration Index 0 is resulted from the initial value of $Q(p_1, p_2)\ \forall p_1 \in \mathbb{P}_S, \forall p_2 \in \mathbb{P}_A$.

sider the path-loss of the signal power at the FC as the free-space path-loss model [29]: $PL = 20\log_{10}(d) + 20\log_{10}(f) - 27.55$, where $d \in \{d_S, d_A\}$ is the distance between the FC and the sensor or the attacker in meters, and f is the signal frequency in megahertz (we assume the network uses an operation frequency of 800MHz). Then, the channel power gain follows an exponential distribution with mean $10^{-\frac{PL}{10}}$mW. In Figure 5.2, we plot the average estimate distortion at the FC in a Nash equilibrium, i.e. $\mathbb{E}\left[D\left(p_S^*, p_A^*\right)\right]$, versus other two non-optimal cases. For the sake of comparison, we apply the same channels at each testing point but vary the non-optimal strategy pairs, $\{p_S(G), p_A^*(G)\}$ and $\{p_S^*(G), p_A(G)\}$, with the power constraints satisfied. Clearly, from the *Definition 2* in Section 5.4, we see that the power policy pair $\{p_S^*(G), p_A^*(G)\}$ is in a Nash equilibrium.

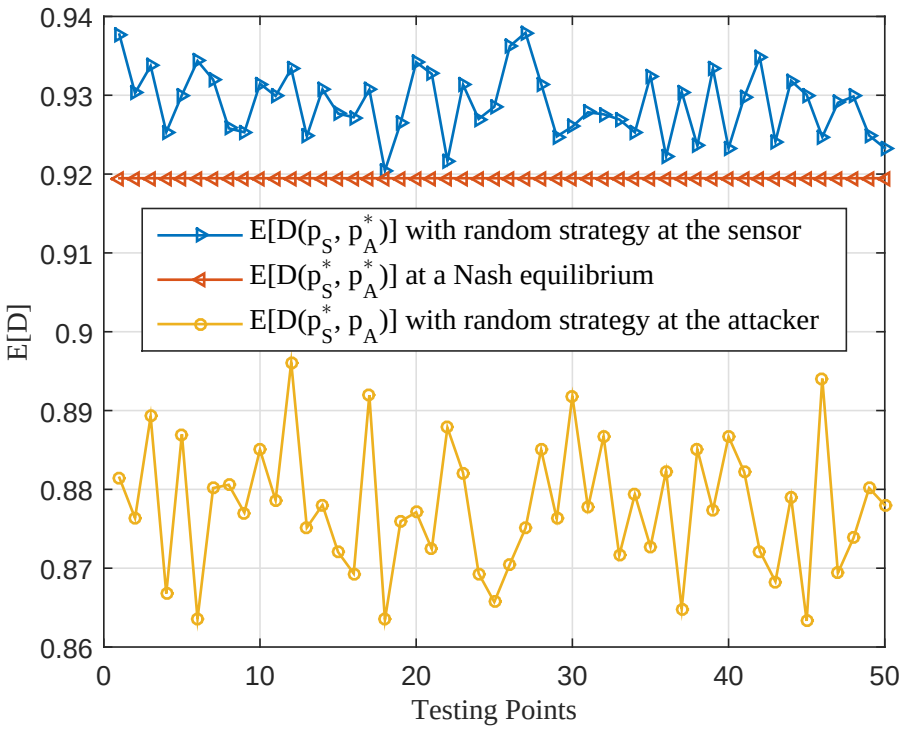

Figure 5.2: Performance comparison for zero-sum games in continuous fading channels with complete channel information.

In discrete fading channels, we assume the attacker generally has better channel conditions than the sensor, and it has a power budget of 0.76mW, which is twice as much

as the sensor's. Suppose both players have three channel types, and the sensor has four power levels whereas the attacker has an action set of three, as given in Table 5.3.

$\mathbb{G}_S$ $\times 10^{-8}$	$\mathbb{P}_S$ mW	$\mathbb{G}_A$ $\times 10^{-7}$	$\mathbb{P}_A$ mW
$g_S^1 = 1.6$	$p_S^1 = 0.0$	$g_A^1 = 1.8$	$p_A^1 = 0.0$
$g_S^2 = 5.4$	$p_S^1 = 0.5$	$g_A^2 = 2.0$	$p_A^2 = 1.0$
$g_S^3 = 7.9$	$p_S^1 = 1.5$	$g_A^3 = 2.8$	$p_A^3 = 2.0$
	$p_S^1 = 2.0$		

Table 5.3: Discrete channel gains and power levels at the sensor and the attacker.

In the case of full CSI, since players are aware of the channel gain of all parties, they can adjust the transmission power based on (g_S, g_A) which consists of nine channel combinations. At the Nash equilibrium, the 'optimal' strategies of sensor and the attacker are given in Figure 5.3 and Figure 5.4 respectively. For instance, when the channel combination index equals to 8 which corresponds to a channel power gain of 7.9×10^{-8} at the sensor and 18×10^{-8} at the attacker, the sensor chooses not to transmit with probability of 0.32 and uses power level 0.5mW to transmit with probability of 0.67; whereas the jamming power at the attacker is 0.5mW with probability of 0.55 and it is 2.0mW with probability of 0.41.

To verify the 'optimal' transmission strategies, in Figure 5.5 we plot the system performance at the Nash equilibrium versus other two non-optimal scenarios. At each testing point, mixed strategies $\pi_S(g_S, g_A)$ and $\pi_A(g_S, g_A)$ are randomly generated while satisfying the power budget $\mathcal{P}_{S\_\text{avg}}$ and $\mathcal{P}_{A\_\text{avg}}$ respectively. $\mathbb{E}[R]$ is evaluated by averaging over 1000000 times for three different mixed strategy pairs, namely, $\{\pi_S, \pi_A^*\}$, $\{\pi_S^*, \pi_A^*\}$ and $\{\pi_S^*, \pi_A\}$. From the definition of a Nash equilibrium point in (5.19) of Section 5.5.1, that $\mathbb{E}_G\left[\mathbb{E}\left[R\left(\pi_S^*(G), \pi_A(G)\right)\middle| G\right]\right] \leq \mathbb{E}_G\left[\mathbb{E}\left[R\left(\pi_S^*(G), \pi_A^*(G)\right)\middle| G\right]\right] \leq \mathbb{E}_G\left[\mathbb{E}\left[R\left(\pi_S(G), \pi_A^*(G)\right)\middle| G\right]\right]$, it is straightforward to tell that the optimal transmission strategy pair is at a Nash equilibrium.

Next, we study the Bayesian Nash equilibrium in incomplete information games, where a player only perfectly knows its own channel types. The power budget for both player are assumed to be the same as the full CSI case, where $\mathcal{P}_{S\_\text{avg}} = 0.38$ mW and $\mathcal{P}_{A\_\text{avg}} = 0.76$ mW. The channel gains and transmission power levels are given in Table

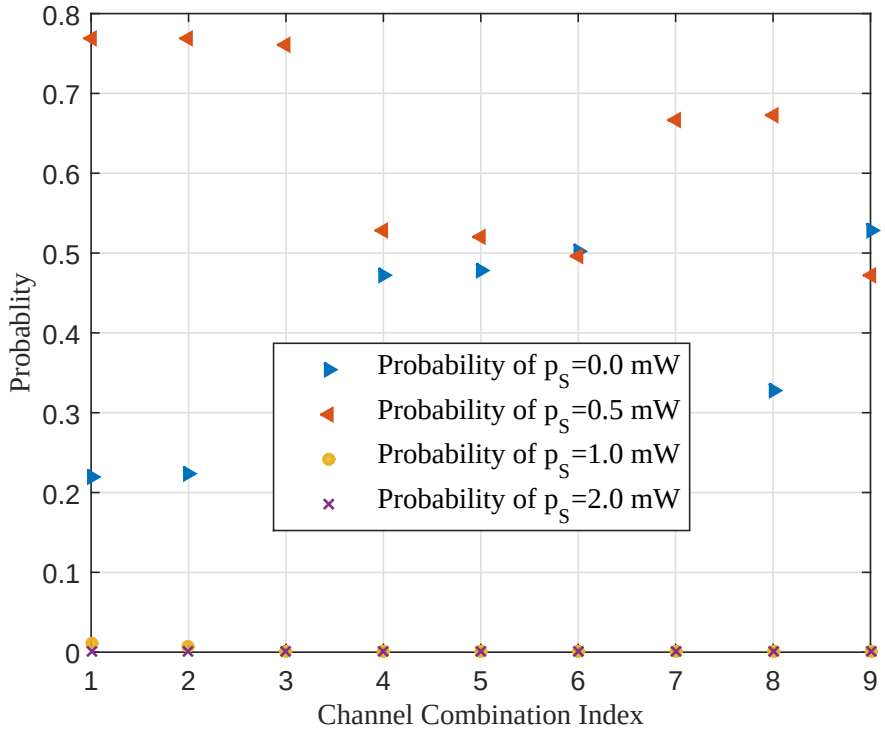

Figure 5.3: Mixed strategies at the sensor.

5.3. Applying the results from Section 5.5.2, the mixed strategies $\pi_S^*(g_S)$ and $\pi_A^*(g_A)$ at the Bayesian Nash equilibrium is respectively given in Table 5.4 and Table 5.5.

	$g_S^i = g_S^1$	$g_S^i = g_S^2$	$g_S^i = g_S^3$
$\Pr(p_S = p_S^1 \mid g_S^i)$	0.189	1.0	1.0
$\Pr(p_S = p_S^2 \mid g_S^i)$	0.811	0.0	0.0
$\Pr(p_S = p_S^3 \mid g_S^i)$	0.000	0.0	0.0
$\Pr(p_S = p_S^4 \mid g_S^i)$	0.000	0.0	0.0

Table 5.4: The sensor's strategy $\pi_S^*(g_S)$.

	$g_A^j = g_A^1$	$g_A^j = g_A^2$	$g_A^j = g_A^3$
$\Pr(p_A = p_A^1 \mid g_A^j)$	0.0	0.786	1.0
$\Pr(p_A = p_A^2 \mid g_A^j)$	1.0	0.214	0.0
$\Pr(p_A = p_A^3 \mid g_A^j)$	0.0	0.000	0.0

Table 5.5: The attacker's strategy $\pi_A^*(g_A)$.

Notice that, from Table 5.4, regardless of the attacker's channel state, the sensor chooses

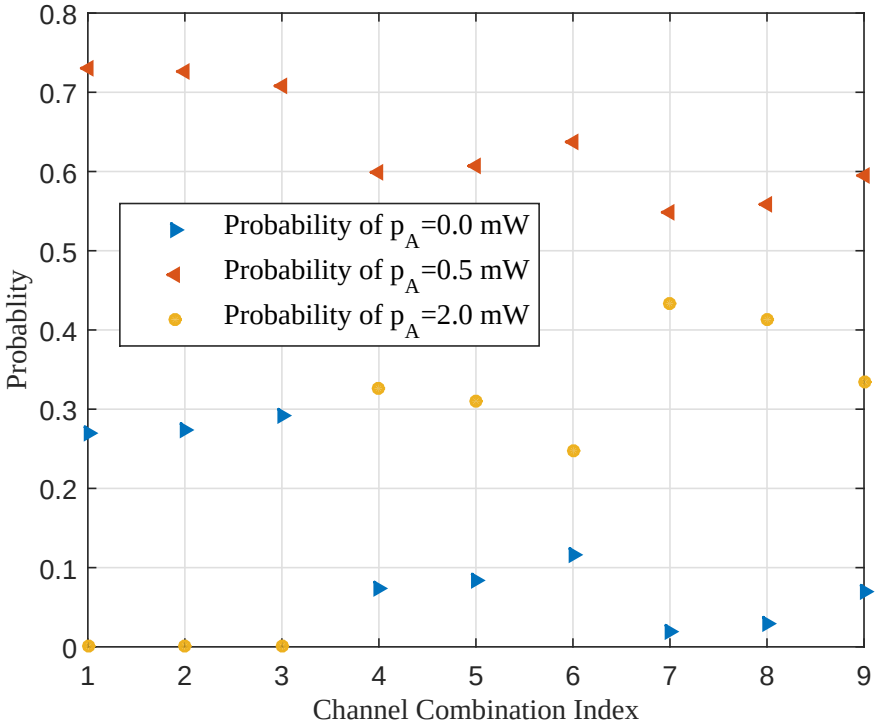

Figure 5.4: Mixed strategies at the attacker.

not to transmit with probability of 1 when g_S^2 and g_S^3 are observed. If the sensor's channel changes to g_S^1, it keeps silent with probability of 0.189; while using p_S^2 to transmit with the probability of 0.811. Similar interpretations can be obtained from the attacker's power allocation strategy $\pi_A^*(g_A)$ in Table 5.5. Assuming at the channel instance (g_S^2, g_A^2), without the sensor's channel information the attacker decides to transmit using power 1.0mW with probability 0.214, which is a kind of a 'silly' move since the sensor would keep silent when the channel is g_S^2. Similar behaviors can be viewed at the channel instance (g_S^1, g_A^1) that although the sensor does not transmit anything to the FC with probability of 0.189, the attacker still keeps jamming the channel with power 1.0mW all the time.

Figure 5.6 shows the optimality of the mixed strategy pair $\{\pi_S^*(g_S), \pi_A^*(g_A)\}$ where players only have access to its own full channel state information. At each testing point, the payoff function is evaluated over 80000 times for three strategy pairs under the same power constraints. It is seen that no player would benefit from moving away from the Bayesian Nash equilibrium point.

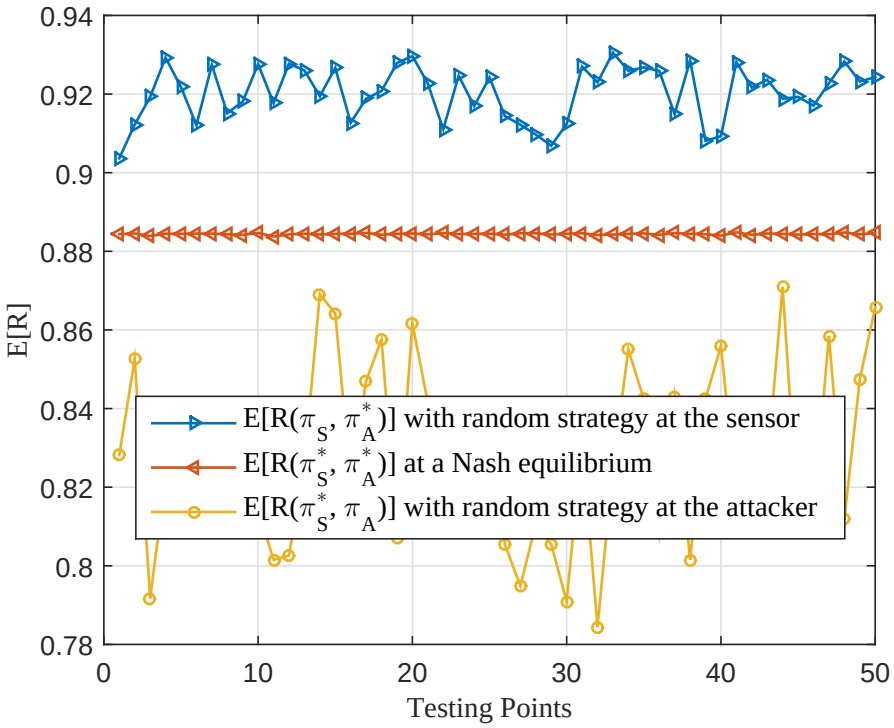

Figure 5.5: Performance comparison for zero-sum games in discrete fading channels with complete channel information.

5.7 Conclusion

In this chapter, we have considered a game theoretic approach for transmit power allocation in distributed estimation under DoS attacks. A two-player zero-sum game is formulated and a Nash equilibrium is investigated for various scenarios. We first looked at repeated games in non-fading channels, where players are assumed being able to receive feedback from the last play and based on which to update its current transmit power. To obtain a strategy pair at a Nash equilibrium, we applied modified backward induction or Nash Q-learning algorithm. In the partial CSI of fading channel scenarios, we considered Bayesian games, where the optimal strategy is formed based on player's knowledge about its own channel and the belief it holds toward the opponent's statistical channel information.

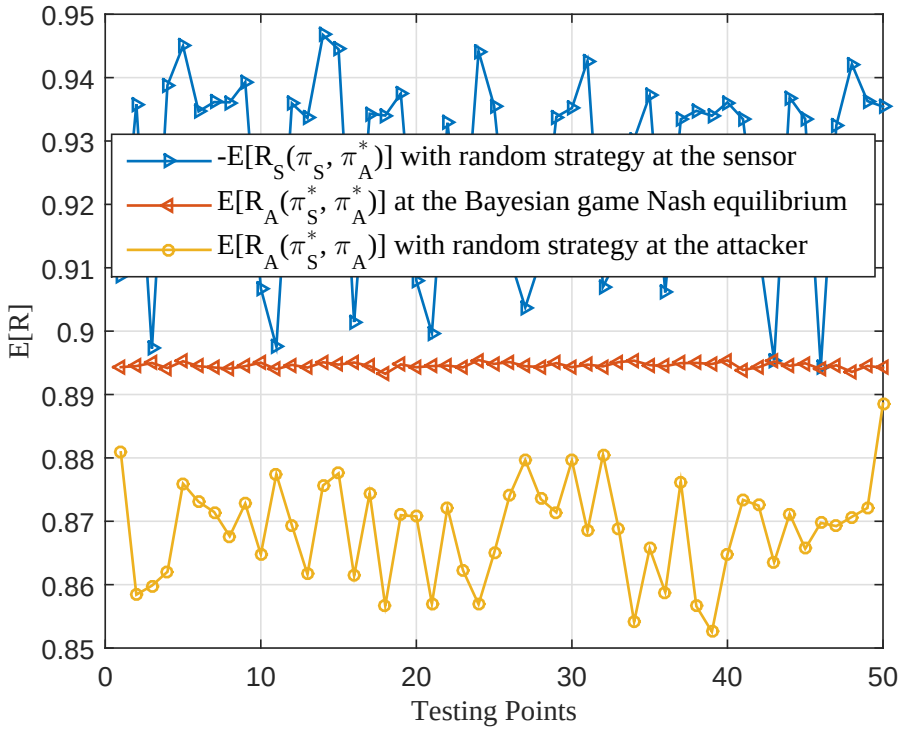

Figure 5.6: Performance comparison for zero-sum Bayesian games in discrete fading channels with incomplete channel information.

Chapter 6
Conclusions

THIS thesis had explored energy-efficient and secure power allocation algorithms in distributed estimation of wireless sensor networks. We focused on two types of security issues in the network − passive eavesdropping and active attacking, and we investigated a number of power transmission policies that minimize the distortion level or distortion outage probability at a remote processor. Below we summarize our work and present some possible future research topics that are related to our work.

6.1 Summary

In Chapter 3 we looked at the performance of distributed estimation in a wireless sensor network with the presence of an eavesdropper. Based on the amplify-and-forward framework, we studied power allocation algorithms that minimize estimation error at the fusion center subject to transmit power constraints at the sensor(s) and secrecy constraints at the adversary in every transmission time slot or over a few fading blocks. If a sensor is equipped with multiple transmit antennas, we derived power allocation schemes that are either able to achieve zero information leakage or significantly enhance the overall system performance depending on the availability of the eavesdropper's channel information. A similar idea is introduced in the multiple sensors scenario where a few sensors serve as friendly relays to broadcast random signals to deceive the adversary.

In Chapter 4 we studied outage-minimizing power allocation algorithms in distributed estimation of a wireless sensor network, where the sensor(s) send their observations to the fusion center via orthogonal multiple access channels (MACs) which are listened to

by a passive eavesdropper via another set of MACs. We obtained the optimal power transmission policies for Rayleigh fading channels assuming that the legitimate receiver knows the channel state information (CSI) of both itself and the eavesdropper. When the number of sensors or number of antennas at the fusion center is large, to overcome a high computational complexity in the partial CSI scenario we proposed a sub-optimal power allocation algorithm that has a low computational complexity. On the other hand, we showed that the distortion outage probability can be dramatically reduced or even driven to zero in some cases for a sensor with two or more transmit antennas.

In Chapter 5 we exploited the sensor's best transmission strategy in distributed estimation when a single sensor system is under denial-of-service attacks, where an adversary sends jamming signals attempting to jeopardize the estimation quality at the fusion center. We applied a game theoretic approach to model the interactions between the sensor and attacker, and derived power allocation strategies at a Nash equilibrium subject to power constraints for non-fading and fading channel environments. Repeated games are studied for non-fading channels where both players are assumed to be able to observe the results from previous plays, based on which to update the current actions. When fading is present, we looked at channel adaptive power policies at a Nash equilibrium. If players only know its own CSI and have a *belief* toward the opponent's statistical channel information, we studied Bayesian games and obtained the 'optimal' power allocation schemes.

6.2 Future Research

For the work on optimal power allocation for distributed estimation in Chapter 3, future research could include the consideration of multiple receive antennas at the fusion center and/or at the eavesdropper. It is of interest to investigate power allocation algorithms that keep a low distortion level at the FC when the eavesdropper is equipped with multiple antennas. One other possible extension is to study the optimal number of sensors serving as friendly relays to broadcast 'noise' in the network. From the simulation results it is clear that when the power budget is relatively large, having a few sensors

transmitting noise can dramatically increase the estimation qualities. But the question is how many sensors, as when the the number of sensors, K, increases, the estimation error decays to a constant at the rate $1/K$. Therefore, instead of fixing the number of relays for all situations, an adaptive network structure considering the power budget and security constraints is more applicable to practical setups.

In Chapter 4, we studied optimal power schemes for distortion outage minimization problems. Owing to the high complexity costs in deriving optimal power strategies in the partial CSI case, we considered a sub-optimal solution in which only the sensor with the best channel conditions transmits its observations to the fusion center. Although this consumes less time to implement, other power allocation algorithms that give better performance and less complexity could be investigated. Additionally, in the multiple-sensor scenarios, the distortion outage probability saturates as the power budget grows large as to keep the security constraint at the eavesdropper satisfied. It is of interest to consider using a small portion of power to generate 'noise' to confuse the adversary, such that the secrecy constraint is always met.

For the power allocation for distributed estimation under denial-of-service attacks in Chapter 5, the work only considers a single sensor system. A more general game theoretic model capturing the conflicting interests of sensors and attacker could be studied for a multiple-sensor network where the fusion center is equipped with multiple receive antennas. One may also be interested in looking at the existence as well as uniqueness of Nash equilibrium for such game theoretic power allocation problems.

Another future work could look at power control problems for different transmission protocols. In Chapter 3 and Chapter 4, we only considered applying orthogonal MACs; we are also interested in investigating the secure and energy-efficient power allocation schemes for a sensor network using coherent MAC protocol. Furthermore, the wireless sensor network considered in this thesis only looks at a single Gaussian source. It is also interesting to study the power allocation policies for a vector source. In regards to the channel model, future research may include different fading distributions.

Bibliography

[1] S. Adlakha, R. Johari, and A. Goldsmith, "Competition in wireless systems via Bayesian interference games," *arXiv preprint arXiv:0709.0516*, 2007.

[2] 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.

[3] G. Anastasi, M. Conti, M. Di Francesco, and A. Passarella, "Energy conservation in wireless sensor networks: A survey," *Ad hoc networks*, vol. 7, no. 3, pp. 537–568, 2009.

[4] T. Aysal and K. Barner, "Sensor data cryptography in wireless sensor networks," *IEEE Transactions on Information Forensics and Security*, vol. 3, no. 2, pp. 273–289, 2008.

[5] G. Bagherikaram and K. N. Plataniotis, "Secure hybrid digital-analog Wyner-Ziv coding," in *2011 IEEE 22nd International Symposium on Personal, Indoor and Mobile Radio Communications*, Sept 2011, pp. 1161–1166.

[6] I. Bahceci and A. K. Khandani, "Linear estimation of correlated data in wireless sensor networks with optimum power allocation and analog modulation," *IEEE Transactions on Communications*, vol. 56, no. 7, pp. 1146–1156, July 2008.

[7] W. Bajwa, A. Sayeed, and R. Nowak, "Matched source-channel communication for field estimation in wireless sensor network," in *IPSN 2005. Fourth International Symposium on Information Processing in Sensor Networks, 2005.*, April 2005, pp. 332–339.

[8] T. Basar and G. J. Olsder, *Dynamic noncooperative game theory*. SIAM, 1995.

[9] S. Boyd and L. Vandenberghe, *Convex Optimization*. New York, NY, USA: Cambridge University Press, 2004.

[10] R. Bustin, R. Liu, H. V. Poor, and S. Shamai, "An MMSE approach to the secrecy capacity of the MIMO Gaussian wiretap channel," *EURASIP Journal on Wireless Communications and Networking*, vol. 2009, p. 3, 2009.

[11] G. Caire, G. Taricco, and E. Biglieri, "Optimum power control over fading channels," *IEEE Transactions on Information Theory*, vol. 45, no. 5, pp. 1468–1489, Jul 1999.

[12] H. Chan and A. Perrig, "Security and privacy in sensor networks," *Computer*, vol. 36, no. 10, pp. 103–105, Oct 2003.

[13] X. Chen, K. Makki, K. Yen, and N. Pissinou, "Sensor network security: a survey," *IEEE Communications Surveys Tutorials*, vol. 11, no. 2, pp. 52–73, 2009.

[14] S. Cui, A. Goldsmith, and A. Bahai, "Energy-constrained modulation optimization," *IEEE Transactions on Wireless Communications*, vol. 4, no. 5, pp. 2349–2360, Sept 2005.

[15] S. Cui, J.-J. Xiao, A. Goldsmith, Z.-Q. Luo, and H. Poor, "Estimation diversity and energy efficiency in distributed sensing," *IEEE Transactions on Signal Processing*, vol. 55, no. 9, pp. 4683–4695, Sept 2007.

[16] L. Dong, Z. Han, A. P. Petropulu, and H. V. Poor, "Improving wireless physical layer security via cooperating relays," *IEEE Transactions on Signal Processing*, vol. 58, no. 3, pp. 1875–1888, 2010.

[17] E. Ekrem and S. Ulukus, "Secure lossy transmission of vector Gaussian sources," *IEEE Transactions on Information Theory*, vol. 59, no. 9, pp. 5466–5487, 2013.

[18] D. Estrin, R. Govindan, J. Heidemann, and S. Kumar, "Next century challenges: Scalable coordination in sensor networks," in *Proceedings of the 5th Annual ACM/IEEE International Conference on Mobile Computing and Networking*, ser. MobiCom '99. New York, NY, USA: ACM, 1999, pp. 263–270.

[19] J. Filar and K. Vrieze, *Competitive Markov decision processes*. Springer Science & Business Media, 2012.

[20] J. W. Friedman, *Game theory with applications to economics*. Oxford University Press New York, 1990.

[21] D. Fudenberg and J. Tirole, *Game theory*. MIT Press, 1991.

[22] J. W. Gardner and V. K. Varadan, *Microsensors, MEMS and smart devices*. John Wiley & Sons, Inc., 2001.

[23] M. Gastpar, B. Rimoldi, and M. Vetterli, "To code, or not to code: lossy source-channel communication revisited," *IEEE Transactions on Information Theory*, vol. 49, no. 5, pp. 1147–1158, 2003.

[24] M. Gastpar, "Uncoded transmission is exactly optimal for a simple Gaussian "sensor" network," *IEEE Transactions on Information Theory*, vol. 54, no. 11, pp. 5247–5251, 2008.

[25] M. Gastpar and M. Vetterli, "Source-channel communication in sensor networks," in *Information Processing in Sensor Networks*. Springer, 2003, pp. 162–177.

[26] S. Gerbracht, A. Wolf, and E. A. Jorswieck, "Beamforming for fading wiretap channels with partial channel information," in *2010 International ITG Workshop on Smart Antennas (WSA)*, Feb 2010, pp. 394–401.

[27] S. Goel and R. Negi, "Guaranteeing secrecy using artificial noise," *IEEE Transactions on Wireless Communications*, vol. 7, no. 6, pp. 2180–2189, June 2008.

[28] A. Goldsmith and P. Varaiya, "Capacity of fading channels with channel side information," *IEEE Transactions on Information Theory*, vol. 43, no. 6, pp. 1986–1992, Nov 1997.

[29] A. Goldsmith, *Wireless communications*. Cambridge university press, 2005.

[30] P. K. Gopala, L. Lai, and H. El-Gamal, "On the secrecy capacity of fading channels," *IEEE Transactions on Information Theory*, vol. 54, no. 10, pp. 4687–4698, 2008.

[31] L. K. Grover, "A fast quantum mechanical algorithm for database search," in *Proceedings of the twenty-eighth annual ACM symposium on Theory of computing*. ACM, 1996, pp. 212–219.

[32] D. Guo, S. Shamai, and S. Verdú, "Mutual information and minimum mean-square error in Gaussian channels," *IEEE Transactions on Information Theory*, vol. 51, no. 4, pp. 1261–1282, 2005.

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

[34] W. W. Hager, "Updating the inverse of a matrix," *SIAM review*, vol. 31, no. 2, pp. 221–239, 1989.

[35] J. C. Harsanyi, "Games with incomplete information played by 'Bayesian' players, I–III: Part I. the basic model," *Management science*, vol. 50, no. 12_supplement, pp. 1804–1817, 2004.

[36] G. He, M. Debbah, and E. Altman, "A Bayesian game-theoretic approach for distributed resource allocation in fading multiple access channels," *EURASIP Journal on Wireless Communications and Networking*, vol. 2010, p. 8, 2010.

[37] J. Hill, R. Szewczyk, A. Woo, S. Hollar, D. Culler, and K. Pister, "System architecture directions for networked sensors," in *ACM SIGOPS operating systems review*, vol. 34, no. 5. ACM, 2000, pp. 93–104.

[38] J. Hu and M. P. Wellman, "Multiagent reinforcement learning: Theoretical framework and an algorithm," in *Proceedings of the Fifteenth International Conference on Machine Learning*, ser. ICML '98. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1998, pp. 242–250.

[39] ——, "Nash Q-learning for general-sum stochastic games," *The Journal of Machine Learning Research*, vol. 4, pp. 1039–1069, 2003.

[40] T.-C. Hu, F. Moricz, and R. Taylor, "Strong laws of large numbers for arrays of row-wise independent random variables," *Acta Mathematica Hungarica*, vol. 54, no. 1, pp. 153–162, 1989.

[41] B. Jacobson, "On the mean value theorem for integrals," *The American Mathematical Monthly*, vol. 89, no. 5, pp. 300–301, 1982.

[42] A. Jeffrey and D. Zwillinger, *Table of integrals, series, and products*. Academic Press, 2007.

[43] M. C. Jeruchim, P. Balaban, and K. S. Shanmugan, *Simulation of communication systems: modeling, methodology and techniques*. Springer Science & Business Media, 2006.

[44] D. C. Jiang, K. L. Teo, and W. Y. Yan, "A new computational method for the functional inequality constrained minimax optimization problem," in *Proceedings of the 34th IEEE Conference on Decision and Control*, vol. 3, Dec 1995, pp. 2310–2315.

[45] N. Jindal, "MIMO broadcast channels with finite-rate feedback," *IEEE Transactions on Information Theory*, vol. 52, no. 11, pp. 5045–5060, 2006.

[46] B. Kailkhura, V. Nadendla, and P. Varshney, "Distributed inference in the presence of eavesdroppers: a survey," *IEEE Transactions on Communications*, vol. 53, no. 6, pp. 40–46, June 2015.

[47] 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.

[48] Y. Kaspi and N. Merhav, "Zero-delay and causal secure source coding," *IEEE Transactions on Information Theory*, vol. PP, no. 99, pp. 1–1, 2015.

[49] S. M. Kay, *Fundamentals of statistical signal processing: Estimation theory*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1993.

[50] A. Khisti and G. W. Wornell, "Secure transmission with multiple antennas - part II: The MIMOME wiretap channel," *IEEE Transactions on Information Theory*, vol. 56, no. 11, pp. 5515–5532, Nov 2010.

[51] A. Khisti, A. Tchamkerten, and G. W. Wornell, "Secure broadcasting over fading channels," *IEEE Transactions on Information Theory*, vol. 54, no. 6, pp. 2453–2469, 2008.

[52] H. Khodakarami and F. Lahouti, "Link adaptation for physical layer security over wireless fading channels," *Communications, IET*, vol. 6, no. 3, pp. 353–362, 2012.

[53] L. Lai and H. El Gamal, "The water-filling game in fading multiple-access channels," *IEEE Transactions on Information Theory*, vol. 54, no. 5, pp. 2110–2122, May 2008.

[54] C. E. Lemke and J. T. Howson, Jr, "Equilibrium points of bimatrix games," *Journal of the Society for Industrial & Applied Mathematics*, vol. 12, no. 2, pp. 413–423, 1964.

[55] A. S. Leong, S. Dey, G. N. Nair, and P. Sharma, "Power allocation for outage minimization in state estimation over fading channels," *IEEE Transactions on Signal Processing*, vol. 59, no. 7, pp. 3382–3397, 2011.

[56] A. Leong and S. Dey, "On scaling laws of diversity schemes in decentralized estimation," *IEEE Transactions on Information Theory*, vol. 57, no. 7, pp. 4740–4759, 2011.

[57] H. Li, L. Lai, and R. Qiu, "A denial-of-service jamming game for remote state monitoring in smart grid," in *2011 45th Annual Conference on Information Sciences and Systems (CISS)*, March 2011, pp. 1–6.

[58] Y. Li, D. Quevedo, and S. Dey, "SINR-based DoS attack on remote state estimation: A game-theoretic approach," *IEEE Transactions on Signal and Information Processing over Networks*.

[59] Y. Liang, H. V. Poor, and S. Shamai, "Secrecy capacity region of fading broadcast channels," in *2007 IEEE International Symposium on Information Theory*, June 2007, pp. 1291–1295.

[60] Y. Liang, H. Poor, and S. Shamai, "Secure communication over fading channels," *IEEE Transactions on Information Theory*, vol. 54, no. 6, pp. 2470–2492, 2008.

[61] W.-C. Liao, T.-H. Chang, W.-K. Ma, and C.-Y. Chi, "QoS based transmit beamforming in the presence of eavesdroppers: An optimized artificial-noise-aided approach," *IEEE Transactions on Signal Processing*, vol. 59, no. 3, pp. 1202–1216, 2011.

[62] T. Liu and S. Shamai, "A note on the secrecy capacity of the multiple-antenna wiretap channel," *IEEE Transactions on Information Theory*, vol. 55, no. 6, pp. 2547–2553, 2009.

[63] D. G. Luenberger, *Optimization by vector space methods*. John Wiley & Sons, 1968.

[64] J. Luo, R. Yates, and P. Spasojevic, "Service outage based power and rate allocation for parallel fading channels," *IEEE Transactions on Information Theory*, vol. 51, no. 7, pp. 2594–2611, 2005.

[65] A. B. MacKenzie and L. A. DaSilva, "Game theory for wireless engineers," *Synthesis Lectures on Communications*, vol. 1, no. 1, pp. 1–86, 2006.

[66] 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, Jan 2009.

[67] S. Marano, V. Matta, and P. Willett, "Distributed detection with censoring sensors under physical layer secrecy," *IEEE Transactions on Signal Processing*, vol. 57, no. 5, pp. 1976–1986, 2009.

[68] J. L. Massey, "Deep-space communications and coding: A marriage made in heaven," in *Advanced Methods for Satellite and Deep Space Communications*. Springer, 1992, pp. 1–17.

[69] M. Médard, "Capacity of correlated jamming channels," in *the Thirty Fifth Allerton Conference*, 2013, pp. 1043–1052.

[70] D. Messerschmitt, "Stationary points of a real-valued function of a complex variable," *EECS Department, University of California, Berkeley, Tech. Rep. UCB/EECS-2006-93*, 2006.

[71] F. Naghibi, S. Salimi, and M. Skoglund, "The CEO problem with secrecy constraints," in *2014 IEEE International Symposium on Information Theory*, June 2014, pp. 756–760.

[72] J. Nash, "Non-cooperative games," *Annals of mathematics*, pp. 286–295, 1951.

[73] R. Negi and J. Cioffi, "Delay-constrained capacity with causal feedback," *IEEE Transactions on Information Theory*, vol. 48, no. 9, pp. 2478–2494, Sep 2002.

[74] R. Negi and S. Goel, "Secret communication using artificial noise," in *2005 IEEE 62nd Vehicular Technology Conference*, vol. 3, Sept 2005, pp. 1906–1910.

[75] E. Obasanjo, G. Tzallas-Regas, and B. Rustem, "An interior-point algorithm for nonlinear minimax problems," *Journal of Optimization Theory and Applications*, vol. 144, no. 2, pp. 291–318, 2009.

[76] F. Oggier and B. Hassibi, "The secrecy capacity of the MIMO wiretap channel," *IEEE Transactions on Information Theory*, vol. 57, no. 8, pp. 4961–4972, Aug 2011.

[77] R. Olfati-Saber, J. A. Fax, and R. M. Murray, "Consensus and cooperation in networked multi-agent systems," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 215–233, Jan 2007.

[78] A. V. Oppenheim and G. C. Verghese. (2010) Class notes for 6.011: Introduction to communication, control and signal processing. Massachusetts Institute of Technology. [Online]. Available: http://ocw.mit.edu/courses/electrical-engineering-and-computer-science/6-011-introduction-to-communication-control-and-signal-processing-spring-2010/readings/MIT6_011S10_notes.pdf

[79] M. J. Osborne, *An introduction to game theory*. Oxford University Press New York, 2004, vol. 3, no. 3.

[80] D. Palomar and S. Verdu, "Gradient of mutual information in linear vector Gaussian channels," *IEEE Transactions on Information Theory*, vol. 52, no. 1, pp. 1515–21 092 015, Jan 2006.

[81] A. Perrig, J. Stankovic, and D. Wagner, "Security in wireless sensor networks," *Communications of the ACM*, vol. 47, no. 6, pp. 53–57, 2004.

[82] G. J. Pottie and W. J. Kaiser, "Wireless integrated network sensors," *Communications of the ACM*, vol. 43, no. 5, pp. 51–58, 2000.

[83] D. E. Quevedo, A. Ahlén, and J. Ostergaard, "Energy efficient state estimation with wireless sensors through the use of predictive power control and coding," *IEEE Transactions on Signal Processing*, vol. 58, no. 9, pp. 4811–4823, 2010.

[84] V. Raghunathan, C. Schurgers, S. Park, and M. B. Srivastava, "Energy-aware wireless microsensor networks," *IEEE Transactions on Signal Processing*, vol. 19, no. 2, pp. 40–50, 2002.

[85] 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, Feb 2011.

[86] P. Sadeghi and P. Rapajic, "Capacity analysis for finite-state Markov mapping of flat-fading channels," *IEEE Transactions on Communications*, vol. 53, no. 5, pp. 833–840, May 2005.

[87] J. Sen, "A survey on wireless sensor network security," *International Journal of Communication Networks and Information Security (IJCNIS)*, vol. 1, no. 2, pp. 59–82, Nov 2010.

[88] S. Shafiee and S. Ulukus, "Correlated jamming in multiple access channels," in *Conference on Information Sciences and Systems*, 2005.

[89] ——, "Mutual information games in multiuser channels with correlated jamming," *IEEE Transactions on Information Theory*, vol. 55, no. 10, pp. 4598–4607, 2009.

[90] C. E. Shannon, "Communication theory of secrecy systems," *Bell system technical journal*, vol. 28, no. 4, pp. 656–715, 1949.

[91] A. N. Shiryaev, *Probability*. Springer-Verlag, New York,, 1996.

[92] Y.-S. Shiu, S. Y. Chang, H.-C. Wu, S.-H. Huang, and H.-H. Chen, "Physical layer security in wireless networks: a tutorial," *IEEE Transactions on Wireless Communications*, vol. 18, no. 2, pp. 66–74, April 2011.

[93] R. Soosahabi and M. Naraghi-Pour, "Scalable PHY-layer security for distributed detection in wireless sensor networks," *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 4, pp. 1118–1126, Aug 2012.

[94] L. Sorber, M. V. Barel, and L. D. Lathauwer, "Unconstrained optimization of real functions in complex variables," *SIAM Journal on Optimization*, vol. 22, no. 3, pp. 879–898, 2012.

[95] S. Tadelis, *Game theory: an introduction*. Princeton University Press, 2013.

[96] J. Villard and P. Piantanida, "Secure multiterminal source coding with side information at the eavesdropper," *IEEE Transactions on Information Theory*, vol. 59, no. 6, pp. 3668–3692, 2013.

[97] C.-H. Wang and S. Dey, "Power allocation for distortion outage minimization in clustered wireless sensor networks," in *2008 IEEE Wireless Communications and Mobile Computing Conference*. IEEE, 2008, pp. 395–400.

[98] ——, "Distortion outage minimization in Rayleigh fading using limited feedback," in *2009 IEEE Global Telecommunications Conference*. IEEE, 2009, pp. 1–8.

[99] C.-H. Wang, A. S. Leong, and S. Dey, "Distortion outage minimization and diversity order analysis for coherent multiaccess," *IEEE Transactions on Signal Processing*, vol. 59, no. 12, pp. 6144–6159, 2011.

[100] J. Wolfowitz, *Coding theorems of information theory*. Springer Science & Business Media, 2012, vol. 31.

[101] R. Wong, *Asymptotic approximations of integrals*. SIAM, 2001, vol. 34.

[102] A. D. Wood and J. Stankovic, "Denial of service in sensor networks," *Computer*, vol. 35, no. 10, pp. 54–62, 2002.

[103] A. D. Wyner, "The wire-tap channel," *Bell System Technical Journal*, vol. 54, no. 8, pp. 1355–1387, 1975.

[104] J.-J. Xiao, S. Cui, Z.-Q. Luo, and A. Goldsmith, "Power scheduling of universal decentralized estimation in sensor networks," *IEEE Transactions on Signal Processing*, vol. 54, no. 2, pp. 413–422, Feb 2006.

[105] ——, "Linear coherent decentralized estimation," *IEEE Transactions on Signal Processing*, vol. 56, no. 2, pp. 757–770, Feb 2008.

[106] J.-J. Xiao and Z.-Q. Luo, "Decentralized estimation in an inhomogeneous sensing environment," *IEEE Transactions on Information Theory*, vol. 51, no. 10, pp. 3564–3575, 2005.

[107] J. Yick, B. Mukherjee, and D. Ghosal, "Wireless sensor network survey," *Computer networks*, vol. 52, no. 12, pp. 2292–2330, 2008.

[108] T. Yoo and A. Goldsmith, "On the optimality of multiantenna broadcast scheduling using zero-forcing beamforming," *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 3, pp. 528–541, 2006.

[109] L. Yuan and G. Qu, "Design space exploration for energy-efficient secure sensor network," in *Proceedings of 2002 IEEE International Conference on Application-Specific Systems, Architectures and Processors*. IEEE, 2002, pp. 88–97.

[110] X. Zhou and M. R. McKay, "Secure transmission with artificial noise over fading channels: Achievable rate and optimal power allocation," *IEEE Transactions on Vehicular Technology*, vol. 59, no. 8, pp. 3831–3842, 2010.