OPTIMAL PARAMETER ENCODING STRATEGIES FOR ESTIMATION THEORETIC SECURE COMMUNICATIONS

A DISSERTATION SUBMITTED TO

THE GRADUATE SCHOOL OF ENGINEERING AND SCIENCE

OF BILKENT UNIVERSITY

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR

THE DEGREE OF

DOCTOR OF PHILOSOPHY

IN

ELECTRICAL AND ELECTRONICS ENGINEERING

By

Çağrı Göken

December 2019

OPTIMAL PARAMETER ENCODING STRATEGIES FOR ESTIMATION THEORETIC SECURE COMMUNICATIONS

By Çağrı Göken

December 2019

We certify that we have read this dissertation and that in our opinion it is fully adequate, in scope and in quality, as a dissertation for the degree of Doctor of Philosophy.

Sinan Gezici (Advisor)

Orhan Arıkan

Berkan Dülek

Tolga Mete Duman

Ayşe Melda Yüksel Turgut

Approved for the Graduate School of Engineering and Science:

Ezhan Karaşan
Director of the Graduate School

ABSTRACT

OPTIMAL PARAMETER ENCODING STRATEGIES FOR ESTIMATION THEORETIC SECURE COMMUNICATIONS

Çağrı Göken
Ph.D. in Electrical and Electronics Engineering
Advisor: Sinan Gezici
December 2019

Physical layer security has gained a renewed interest with the advances in modern wireless communication technologies. In estimation theoretic security, secrecy levels are measured via estimation theoretic tools and metrics, such as mean-squared error (MSE), where the objective is to perform accurate estimation of the parameter at the intended receiver while keeping the estimation error at the eavesdropper above a certain level. This framework proves useful both for analyzing the achievable performance under security constraints in parameter estimation problems, and for designing low-complexity, practical methods to enhance security in communication systems. In this dissertation, we investigate optimal deterministic encoding of random scalar and vector parameters in the presence of an eavesdropper, who is unaware of the encoding operation. We also analyze optimal stochastic encoding of a random parameter under secrecy constraints in a Gaussian wiretap channel model, where the eavesdropper is aware of the encoding strategy at the transmitter. In addition, we perform optimal parameter design for secure broadcast of a parameter to multiple receivers with fixed estimators.

First, optimal deterministic encoding of a scalar parameter is investigated in the presence of an eavesdropper. The aim is to minimize the expectation of the conditional Cramér-Rao bound (ECRB) at the intended receiver while keeping the MSE at the eavesdropper above a certain threshold. The eavesdropper is modeled to employ the linear minimum mean-squared error (LMMSE) estimator based on the encoded version of the parameter. First, the optimal encoding function is derived in the absence of secrecy constraints for any given prior distribution on the parameter. Next, an optimization problem is formulated under a secrecy constraint and various solution approaches are proposed. Also, theoretical results on the form of the optimal encoding function are provided. Furthermore, a robust parameter encoding approach is developed. In this case, the objective is

to maximize the worst-case Fisher information of the parameter at the intended receiver while keeping the MSE at the eavesdropper above a certain level. The optimal encoding function is derived when there exist no secrecy constraints. Next, to obtain the solution of the problem in the presence of the secrecy constraint, the form of the encoding function that maximizes the MSE at the eavesdropper is explicitly derived for any given level of worst-case Fisher information. Then, based on this result, a low-complexity algorithm is provided to calculate the optimal encoding function for the given secrecy constraint. Numerical examples are presented to illustrate the theoretical results for both the ECRB and worst-case Fisher information based designs.

Second, optimal deterministic encoding of a vector parameter is investigated in the presence of an eavesdropper. The objective is to minimize the ECRB at the intended receiver while satisfying an individual secrecy constraint on the MSE of estimating each parameter at the eavesdropper. The eavesdropper is modeled to employ the LMMSE estimator based on the noisy observation of the encoded parameter without being aware of encoding. First, the problem is formulated as a constrained optimization problem in the space of vector-valued functions. Then, two practical solution strategies are developed based on nonlinear individual encoding and affine joint encoding of parameters. Theoretical results on the solutions of the proposed strategies are provided for various scenarios on channel conditions and parameter distributions. Finally, numerical examples are presented to illustrate the performance of the proposed solution approaches.

Third, estimation theoretic secure transmission of a scalar random parameter is investigated in the presence of an eavesdropper. The aim is to minimize the estimation error at the receiver under a secrecy constraint at the eavesdropper; or, alternatively, to maximize the estimation error at the eavesdropper for a given estimation accuracy limit at the receiver. In the considered setting, the encoder at the transmitter is allowed to use a randomized mapping between two one-to-one and continuous functions and the eavesdropper is fully aware of the encoding strategy at the transmitter. For small numbers of observations, both the eavesdropper and the receiver are modeled to employ LMMSE estimators, and for large numbers of observations, the ECRB metric is employed for both the receiver and the eavesdropper. Optimization problems are formulated and various theoretical results are provided in order to obtain the optimal solutions and to analyze the effects of encoder randomization. In addition, numerical examples are presented to corroborate the theoretical results. It is observed that stochastic

encoding can bring significant performance gains for estimation theoretic secrecy problems.

Finally, estimation theoretic secure broadcast of a random parameter is investigated. In the considered setting, each receiver device employs a fixed estimator and carries a certain security risk such that its decision can be available to a malicious third party with a certain probability. The encoder at the transmitter is allowed to use a random mapping to minimize the weighted sum of the conditional Bayes risks of the estimators under secrecy and average power constraints. After formulating the optimal parameter design problem, it is shown that the optimization problem can be solved individually for each parameter value and the optimal mapping at the transmitter involves a randomization among at most three different signal levels. Sufficient conditions for improvability and nonimprovability of the deterministic design via stochastic encoding are obtained. Numerical examples are provided to corroborate the theoretical results.

Keywords: Parameter estimation, physical layer security, secrecy, Cramér-Rao bound (CRB), optimization, mean-squared error, Fisher information matrix (FIM), Gaussian wiretap channel, broadcast channel.

ÖZET

KESTİRİM KURAMSAL GÜVENLİ HABERLEŞME İÇİN OPTİMAL PARAMETRE KODLAMA STRATEJİLERİ

Çağrı Göken
Elektrik ve Elektronik Mühendisliği, Doktora
Tez Danışmanı: Sinan Gezici
Aralık 2019

Modern kablosuz haberleşme teknolojilerindeki gelişmelerle beraber fiziksel katman güvenliğine duyulan ilgi yeniden artmıştır. Kestirim kuramsal güvenlikte gizlilik seviyeleri, ortalama karesel hata (OKH) gibi kestirim kuramı araç ve metrikleriyle ölçülmekte, buradaki amaç ise parametrenin gerçek alıcıda doğru bir şekilde kestirimi sağlanırken, gizli dinleyici tarafında oluşacak hatayı da belirli bir seviyenin üstünde tutabilmektir. Böyle bir model, hem parametre kestirim problemlerinde güvenlik kısıtlamaları altında ulaşılabilecek performansı analiz etme açısından, hem de düşük karmaşıklı, pratik yöntemlerle haberleşme sistemlerindeki güvenliği artırabilme açısından faydalıdır. Bu tezde, rastgele skaler ve vektör parametrelerin optimal deterministik kodlaması, kodlama işleminden habersiz olan bir gizli dinleyici varlığı altında araştırılmaktadır. Ayrıca, rastgele bir parametrenin optimal stokastik kodlaması, gizlilik kısıtlamaları ve Gauss dinleme kanalı modeli altında analiz edilmektedir. Burada dinleyicinin göndericideki kodlama stratejisini tam olarak bildiği varsayılmaktadır. Ek olarak, bir parametrenin sabit kestiricilere sahip birden fazla alıcıya güvenli bir şekilde yollanması için optimal parametre tasarımı gerçekleştirilmektedir.

İlk olarak, skaler bir parametrenin deterministik kodlaması gizli dinleyici varlığı altında incelenmektedir. Amaç, gizli dinleyicideki OKH değerini belli bir eşiğin üzerinde tutarken, hedeflenen alıcıdaki ortalama koşullu Cramér-Rao sınırını (OCRS) minimize etmektir. Dinleyici, kodlanmış parametreyi temel alarak hesaplanmış olan doğrusal minimum ortalama karasel hata (DMOKH) kestirici kullanıyor şekilde modellenmektedir. Öncelikle, gizlilik kısıtlamaları olmadığı durumlar için, herhangi verilmiş bir öncül dağılıma sahip parametrenin optimal kodlama fonksiyonu elde edilmektedir. Daha sonra, optimizasyon problemi gizlilik kısıtlaması altında formüle edilmekte, çeşitli çözüm yaklaşımları

önerilmektedir. Ayrıca, optimal kodlama fonksiyonunun yapısı üzerine kuramsal sonuçlar sunulmaktadır. Bunların dışında, gürbüz bir parametre kodlama yaklaşımı geliştirilmektedir. Bu durumda amaç, gizli dinleyicideki OKH değerini belli bir seviyenin üzerinde tutarken, hedeflenen alıcıdaki en kötü Fisher bilgisini maksimize etmektir. Gizlilik kısıtlaması olmadığı durum için optimal kodlama fonksiyonu elde edilmektedir. Daha sonra, verilmiş herhangi bir en kötü Fisher bilgisi seviyesi için, gizli dinleyicideki OKH değerini maksimize eden kodlama fonksiyonu da tam olarak elde edilmektedir. Bu sonuç kullanılarak, verilmiş bir gizlilik kısıtlaması için optimal kodlama fonksiyonunu hesaplayacak düşük karmaşıklı bir algoritma temin edilmektedir. Sayısal sonuçlar, hem OCRS hem de en kötü Fisher bilgisi temelli tasarımlar için elde edilmiş kuramsal sonuçları örneklendirmek için sunulmaktadır.

İkinci olarak, gizli dinleyicinin bulunduğu durumda, bir vektör bir parametrenin optimal deterministik kodlaması araştırılmaktadır. Amaç, gizli dinleyicide her bir parametre için oluşacak bireysel gizlilik kısıtlamalarını sağlarken, hedeflenen alıcıdaki OCRS değerini minimize etmektir. Gizli dinleyici, kodlanmış parametrenin gürültü içeren gözlemlerini kullanan ancak kodlamanın farkında olmadan oluşturulmuş bir DMOKH kestirici kullanacak şekilde modellenmektedir. Öncelikle problem, kısıtlamalı optimizasyon problemi olarak vektör değerli fonksiyonlar uzayı üzerinde formüle edilmektedir. Daha sonra, iki tane pratik çözüm stratejisi geliştirilmektedir. Bunlar doğrusal olmayan bireysel kodlama ve parametrelerin afin ortak kodlaması üzerine kurulmaktadır. Önerilmiş stratejilerin çözümleri üzerine kuramsal sonuçlar, değişik kanal şartları ve parametre dağılımı senaryoları için verilmektedir. Son olarak, sayısal sonuçlar, önerilmiş çözüm yaklaşımlarının performansı üzerine örnekler sunmaktadır.

Üçüncü olarak skaler rastgele bir parametrenin, gizli bir dinleyici varlığı altında, kestirim kuramsal güvenli bir şekilde iletilmesi araştırılmaktadır. Hedef, gizli dinleyicide bir gizlilik kısıtlaması varken alıcıdaki kestirim hatasını minimize etmek veya alternatif olarak, alıcıda bir kestirim doğruluğu sınırı varken, gizli dinleyicide oluşacak kestirim hatasını maksimize etmektir. Çalışılan düzende, göndericide bulunan kodlayıcı, iki adet birebir ve sürekli fonksiyon arasında rastgele bir tercihte bulunabilmektedir ve gizli dinleyici, göndericinin bu kodlama stratejisini tamamen bilmektedir. Küçük sayıdaki gözlemler için, hem alıcının, hem de gizli dinleyicinin DMOKH kestirici kullandıkları varsayılmakta, fazla sayıdaki gözlemler için ise OCRS metriği hem alıcı hem de gizli dinleyici için kullanılmaktadır. Optimizasyon problemleri formüle edilmekte, kuramsal

sonuçlar hem optimal çözümleri bulmak, hem de kodlayıcı rastgeleleştirmesinin etkilerini analiz etmek açısından verilmektedir. Ek olarak, sayısal örnekler kuramsal sonuçları desteklemek amacıyla sunulmaktadır. Stokastik kodlamanın kestirim kuramsal gizlilik problemler için önemli performans kazancı sağladığı gözlenmektedir.

Son olarak, rastgele bir parametrenin kestirim kuramsal güvenli olarak yayınlanması araştırılmaktadır. Çalışılan düzende her alıcı cihaz, sabit bir kestirici kullanmakta, ayrıca belli bir güvenlik riski taşımaktadır. Yani kestiricilerin kararları belli bir olasılıkla zararlı üçüncü taraf kullanıcılar tarafından ele geçirilmiş olabilmektedir. Göndericideki kodlayıcı, güvenlik ve ortalama güç kısıtlamaları altında, kestiricilerin ağırlıklandırılmış koşullu Bayes risk toplamlarını minimize edecek şekilde rastgele bir eşleme kullanabilmektedir. Optimal parametre tasarımı problemi formüle edildikten sonra, optimizasyon probleminin her bir parametre değeri için ayrı ayrı çözülebileceği gösterilmektedir. Ayrıca, göndericide kullanılacak optimal eşlemede, en fazla üç farklı sinyal seviyesi arasında rastgeleleştirme yapılabileceği gösterilmektedir. Stokastik kodlama vasıtasıyla deterministik kodlamanın geliştirilebilirliği ve geliştirilemezliği üzerine yeter koşullar elde edilmektedir. Sayısal örnekler kuramsal sonuçları desteklemek amacıyla verilmektedir.

Anahtar sözcükler: Parametre kestirimi, fiziksel katman güvenliği, gizlilik, Cramér-Rao sınırı, eniyileme, karesel ortalama hata, Fisher bilgi matrisi, Gauss hat dinleme kanalı, yayın kanalı.

Acknowledgement

I would like to express my deepest gratitude to my supervisor Prof. Sinan Gezici. His knowledge, experience and vision have provided an invaluable guidance throughout my years in Bilkent. His positive attitude and support have motivated me towards completing my PhD degree. He has also provided the necessary encouragement to begin my graduate studies. It was a great pleasure and honor for me to work with him.

I would like to thank my thesis monitoring committee members Prof. Orhan Arıkan and Assoc. Prof. Berkan Dülek for their support and invaluable suggestions during my studies. I would also like to extend my thanks to Prof. Tolga Mete Duman and Assoc. Prof. Ayşe Melda Yüksel Turgut for agreeing to serve in my dissertation committee.

I owe my deepest gratitude to my family for their love, continuous support and much needed motivation to begin my PhD studies. This thesis would not have been possible without them.

I am grateful to my friends and colleagues in Bilkent EEE Department and ASELSAN for the valuable technical discussions and their encouragement. I would also like to thank the administration of ASELSAN for the support on my graduate studies.

I appreciate and acknowledge the financial support from the Scientific and Technological Research Council of Turkey (TÜBİTAK) through 2211-A Scholarship Program of Directorate of Science Fellowships and Grant Programmes (BİDEB) during my PhD studies.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Security has been a crucial issue for communications. In a secure communication system, the main goal is to secretly transmit data to an intended receiver in the presence of a malicious third party such as an eavesdropper. As the age of Internet of Things (IoT), smart homes and cities, self-driving cars, and wireless sensor networks with a vast number of nodes has already arrived, it is necessary to find ways to ensure secure communication of data in such systems. Massive deployments of sensors, the nature of wireless links across a network, and the sensitivity of data collected by sensors present serious security challenges. Traditionally, key-based cryptographic approaches have been employed in many applications for secure communication [1], [2]. In [3], Shannon proved that the cryptographic approach known as one-time-pad can achieve the perfect secrecy; that is, the original message and the cypher text become independent, if the number of different keys is at least as high as the number of messages. However, the management of key generation and distribution can be very challenging in heterogenous and dynamic networks with vast numbers of connections [4], [5]. Furthermore, as many nodes in sensor networks are low-cost with limited battery power and bandwidth and have strict latency requirements, it may not be suitable to consider cryptographic solutions as the only layer of security in such systems [6].

Based on these motivations, there has been a renewed interest in physical layer

secrecy to develop alternative or complementary layers of security technologies. Physical layer secrecy is based on the idea of exploiting the randomness in wireless channel conditions to ensure secure communication [7]. In [8], Wyner proved that when the channel between the transmitter and the eavesdropper is a degraded version of the channel between the transmitter and the intended receiver, then reliable communication can be achieved without information leakage to the eavesdropper. One common approach to measure the amount of achieved secrecy is to use information theoretic metrics and tools, such as capacity, and to examine the highest rates at which the transmitter can encode a message while maintaining a certain equivocation level at the eavesdropper. Following Wyner's work, a multitude of studies have been performed based on this approach for various channel models such as fading channels [9]-[11], Gaussian wiretap, broadcast and interference channels [12]-[18] and transmission scenarios such as multiantenna systems [19], cooperative communications with user or jammer cooperation [20]-[23]. In the literature, alternative metrics and frameworks have also been utilized to quantify secrecy levels. For example, secure communication problem is investigated based on the signal-to-noise ratio (SNR) metric in the quality-of-service (QoS) framework in [24]-[26]. In [27], the secrecy constrained distributed detection problem is studied under Bayesian and Neyman-Pearson frameworks. Alternatively, estimation theoretic tools such as mean-squared error (MSE) and Fisher information have recently been used to measure security performance in parameter estimation problems and to design low-complexity, practical and secure communication systems [28]-[46]. In this approach the aim is to optimize the estimation accuracy performance of the estimator at the intended receiver, while keeping estimation error at the eavesdropper above a certain target.

In this dissertation, optimal parameter encoding strategies are investigated to ensure estimation theoretic secure communications in the presence of an eavesdropper. In Chapter 2, we investigate the optimal deterministic encoding of a scalar random parameter under secrecy constraints, where the objective is to optimize the estimation accuracy based on the expectation of the conditional Cramér-Rao bound (ECRB) and alternatively worst-case Fisher information at

the intended receiver while keeping the mean-squared error (MSE) at the eavesdropper above a certain level [42, 43]. In Chapter 3, we focus on the optimal deterministic encoding of a random vector parameter in the presence of an eavesdropper and develop practical solution strategies to minimize the ECRB at the intended receiver while satisfying an individual secrecy constraint on the MSE of estimating each parameter at the eavesdropper [44]. In both chapters, the common assumption is that the eavesdropper is not aware of the encoding operation at the transmitter. In Chapter 4, we investigate optimal encoding of a scalar random parameter under the assumption that the encoding strategy is fully available to the eavesdropper, and the transmitter can utilize a randomized mapping between two one-to-one and continuous functions to enhance security [45]. Finally, in Chapter 5, we work on the optimal stochastic parameter design for secure broadcast problem, where each receiver device employs a fixed estimator that can be compromised by a malicious third party with a certain probability [46]. In the following, we present a literature review and summarize the contributions of the thesis.

1.1 Optimal Deterministic Encoding for Secure Communications

As a common alternative approach to the information theoretic secrecy, estimation theoretic secrecy has been employed in a wide variety of problems in the literature [28]–[36]. In [28], the output Y of a channel for a given input X is encoded by a random mapping $P_{Z|Y}$ in order to ensure that the MMSE for estimating Y based on Z is minimized while the MMSE for estimating X based on Z is above $(1-\epsilon)Var(X)$ for a given $\epsilon \geq 0$, where $Var(X)$ denotes the variance of X. In [29], the secret communication problem is considered for Gaussian interference channels in the presence of eavesdroppers. The problem is formulated to minimize the total MMSE at the intended receivers while keeping the MMSE at the eavesdroppers above a certain threshold, where joint artificial noise and linear precoding schemes are used to satisfy the secrecy requirements.

Another application area of the estimation theoretic secrecy is distributed inference networks, where the information coming to a fusion center (FC) from various sensor nodes can also be observed by eavesdroppers. The secrecy for distributed detection and estimation can be ensured via various techniques such as design of sensor quantizers and decision rules, stochastic encoding, artificial noise to confuse eavesdroppers, and MIMO beamforming [30]. In [31]-[33] the secrecy problem in a distributed inference framework is investigated, where the information coming to a fusion center from various sensor nodes can also be observed by eavesdroppers. In [31], the estimation problem of a single point Gaussian source in the presence of an eavesdropper is analyzed for the cases of multiple transmit sensors with a single antenna and a single sensor with multiple transmit antennas. Optimal transmit power allocation policies are derived to minimize the average MSE for the parameter of interest while guaranteeing a target MSE at the eavesdropper. In [32], the asymptotic secrecy and estimation problem is studied when the sensor measurements are quantized and the channel between sensors and receivers are assumed to be binary symmetric channels. Furthermore, in [33], the secrecy is investigated in terms of distortion (and secrecy) outage, which is the probability that the MMSE at the FC (eavesdropper) is above (below) certain distortion levels. The optimal transmit power allocation policies are derived to minimize the distortion outage at the FC under an average transmit power and a secrecy outage constraint at the eavesdropper. In [34], the secure inference problem is investigated for deterministic parameters in IoT systems under spoofing and man-in-the-middle-attack (MIMA). For MIMAs, necessary and sufficient conditions are derived to decide when the attacked data can or cannot improve the estimation performance in terms of the Cramér-Rao bound. For spoofing attacks, effective attack strategies are described with a guaranteed performance in terms of Cramér-Rao bound (CRB) degradation and it is shown that quantization imposes a limit on the robustness of the system against such attacks. In [35], privacy of households using smart meters is considered in the presence of adversary parties who estimate energy consumption based on data gathered in smart meters. The house utilizes the batteries to mask the real energy consumption. The Fisher information is employed as a metric for both scalar and multivariable case and the optimal policies for the utilization of batteries are derived to minimize

the Fisher information to achieve privacy.

For estimation theoretic approaches, the Cramér-Rao bounds provide useful theoretical limits for assessing performance of estimators. It is known that when the parameter to be estimated is non-random, the conditional CRB states that, under some regularity conditions, the MSE of any unbiased estimator is bounded by the inverse of the Fisher information for each given value of the parameter [47]. On the other hand, if the parameter to be estimated is random with a known prior distribution, then the extended versions of the CRB, such as the Bayesian Cramér-Rao bound (BCRB) and the expectation of the conditional Cramér-Rao bound (ECRB), can be employed [48]. Even though the BCRB effectively takes the prior information into account and can provide a useful lower bound for the maximum a-posterior probability (MAP) estimator in the low signal-to-noise ratio (SNR) regime, it does not exist for some prior distributions due to the violation of an assumption in its derivation. For example, the BCRB does not exist when the parameter has a uniform prior distribution over a closed set [48]–[49]. More importantly, when the conditional CRB is a function of the unknown parameter, which is commonly the case, the BCRB does not present a tight bound in the high SNR regime.[1]

Therefore, for the parameter encoding problem, the use of the BCRB as the objective function may be misleading and can result in trivial bounds in some cases. For these reasons, ECRB can be employed instead of BCRB, as it has widely been utilized in a variety of applications in the literature; e.g., [50]-[53]. The ECRB is known to provide a tight limit for the MAP estimator asymptotically, and converges to the Ziv-Zakai bound (ZZB) in the high SNR regime [48]. Therefore, the optimization of parameter encoding according to the ECRB metric leads to close-to-optimal performance for practical MAP estimators in the high SNR regime. Although the ZZB can provide a tight limit for all SNRs, it has high computational complexity compared to the ECRB [48, 54] and does not allow theoretical investigations for achieving an intuitive understanding of the

[1]This is also a problem for the weighted Cramér-Rao bound (WCRB), which is a generalized version of the BCRB using a weighting function, and can be employed for the cases in which the BCRB does not exist [48, 49].

parameter encoding problem.

In the first part of Chapter 2, we consider the transmission of a scalar parameter to an intended receiver in the presence of an eavesdropper. In order to ensure secret communications, we utilize an encoding function (continuous and one-to-one) applied on the original parameter. The aim is to minimize the ECRB at the intended receiver while ensuring a certain MSE target at the eavesdropper. It is assumed that the eavesdropper uses a linear MMSE (LMMSE) estimator without being aware of the encoding. An optimization problem is formulated to obtain the optimal encoding function for given target MSE levels. At the first step, the secrecy requirements are omitted and the optimization problem is solved under no constraints. In that case, a closed-form analytical solution is provided for the optimal encoding function for any given prior distribution. Next, the MSE constraint for the eavesdropper is included and various solution approaches, such as polynomial approximation, piecewise linear approximation, and linear encoding are proposed. Also, theoretical results are derived related to the structure of the optimal encoding function under some assumptions.

In the second part of Chapter 2, we focus on the worst-case CRB (equivalently, the worst-case Fisher information) in order to develop a robust parameter encoding approach that guarantees a certain level of estimation accuracy at the intended receiver. The proposed problem requires different solution approaches than that of the problem based on ECRB due to the minimax nature of the worst-case optimization. In particular, we investigate the transmission of a uniformly distributed scalar parameter to an intended receiver in the presence of an eavesdropper. Similarly to the first part of the chapter, we utilize an encoding function (which is one-to-one and continuous except at a finite number of points) applied on the original parameter to facilitate secret communications, and the eavesdropper is modeled to employ the LMMSE estimator based on the noisy observation of the encoded parameter without being aware of encoding. The objective is to minimize the maximum CRB (equivalently, to maximize the minimum Fisher information) at the intended receiver while ensuring a certain MSE target at the eavesdropper. An optimization problem is formulated to obtain the optimal encoding function for a given target MSE level at the eavesdropper.

First, the secrecy constraint is omitted and the optimization problem is solved under no constraints, which yields a closed-form analytical solution. Then, to solve the optimal encoding problem in the presence of the MSE constraint on the eavesdropper, the optimal encoding function that maximizes the MSE at the eavesdropper is derived analytically for any given level of minimum Fisher information at the intended receiver. Based on this analytical result, a low-complexity algorithm is proposed to obtain the solution of the proposed problem.

Even though the optimal parameter encoding problem has been investigated for scalar parameters in Chapter 2 from a CRB-based optimization perspective, it is possible that the channel input can contain multiple parameters in many practical scenarios such as [29], [35], [36]–[38]. Estimation of multiple parameters is required in many applications such as in localization [47] and joint frequency and phase estimation [48]. Secure transmission of multiple parameters has also been investigated in the literature for different applications and scenarios. In [36], the filter design with secrecy constraints is studied for a multiple-input multiple-output (MIMO) Gaussian wiretap channel, where the parameter of interest is a vector, each component of which is zero mean with a unit variance and is independent of others. In [37], a beamforming scheme is proposed for a downlink multiuser MIMO system for secure communication, where the vector parameter carries the unit-energy data symbols of each user. Another important use-case for the secure multiple parameter estimation problem occurs in smart grids/homes and internet of things (IoT) systems [38]. For example, the vector parameter carries the state of the grid, i.e., the voltage angles and magnitudes at each of the buses, in the scenario of state estimation problem in a smart-grid system. In another example, the parameter is the state of the position and velocity of an autonomous vehicle. In a further example, the parameter represents the pollutant concentration over an entire city in an air monitoring system in a smart city, where each individual component of the vector can represent the pollutant concentration in a certain neighborhood [38].

In Chapter 3, we focus on a secure multi-parameter transmission scenario based on the preceding motivations. Similarly to Chapter 2, the parameter is encoded using an encoding function prior to transmission. It is important to emphasize

that the difference of the multiparameter scenario from the single parameter case is not only based on the number of parameters. In the encoding of a scalar parameter, a single scalar valued function is utilized as an encoder. In the multiparameter case, as the parameter of interest is a random vector, the encoding function becomes a vector valued function, which generates different opportunities compared to the scalar case during the encoding operation such as joint encoding of parameters using a nonlinear function. As a simple example, consider a scenario in which the parameter involves the coordinates of the location of a target. Then, before sending the true coordinate, a simple shuffle of the coordinates can create a considerable amount of localization error at the eavesdropper as the eavesdropper is not aware that such a secret-key is employed. This means that the problem of optimal encoding of multiple parameters requires new analyses and theoretical investigations as the theoretical analysis and tools employed in Chapter 2 are not able to cover it directly in general. When the encoding function is assumed to be an affine function as a special case, it corresponds to employing a linear precoding matrix strategy, which has been employed in various studies to ensure security [29], [30]. In Chapter 3, the objective of encoding design is to minimize the ECRB, which is defined as the average of the trace of the inverse Fisher Information Matrix (FIM). The eavesdropper is modeled to employ the linear MMSE (LMMSE) estimator based on the noisy observation of the encoded parameter without being aware of encoding. Compared to other studies in the estimation theoretic security literature, the proposed formulation is a novel approach for problems involving multiple parameters. Also, the possible correlations among the parameters and the correlations in the noise components of intended receiver/eavesdropper are taken into account, which is not applicable in the scalar case. First, the optimization problem is formulated to obtain the optimal encoding function for a given target MSE level based on the assumption that the joint encoding approach is applied via a nonlinear encoding function. Based on this formulation, two special cases of the generic form of the encoding function is studied to develop practical encoders. In the first approach, each element of the vector parameter is encoded individually by a nonlinear scalar function. For this strategy, it is shown that when the transmitted parameters are independent and the channel noise for the eavesdropper is white, the optimization problem

decouples into individual scalar problems, which are investigated in the first part of Chapter 2. Then, the case for colored Gaussian noise for the eavesdropper is investigated, where the optimization problem cannot be decoupled. For the two-parameters case, fundamental insights are provided about the optimal solution of the multiple parameter case by considering the correlation in the noise components, which cannot be obtained by studying the single parameter case. In the second approach, the encoding function is assumed to be an affine function. This method allows for joint encoding, or simple shuffle and scale of the parameters, which cannot be utilized in the single parameter case. Therefore, all the theoretical analyses related to this approach are new contributions. For this strategy, first the secrecy requirements are omitted, and an optimal solution is derived theoretically when the channel noise for the intended receiver is white. Next, the MSE constraint for the eavesdropper is considered and several theoretical results are provided regarding the form of the optimal affine joint encoder.

1.2 Encoder Randomization for Secure Communications

Stochastic encryption has been used as a defense mechanism against eavesdropper attacks in the estimation theoretic security framework [30],[39]-[41]. In [39], stochastic encryption is performed based on the 1-bit quantized version of a noisy sensor measurement to achieve secret communication, where both symmetric and asymmetric bit flipping strategies are considered under the assumptions that the intended receiver is aware of the flipping probabilities and the eavesdropper is unaware of the encryption. The effects of the flipping probabilities on the Cramér-Rao bound (CRB) and the maximum likelihood (ML) estimator at the fusion center, and on the bias and the MSE at the eavesdropper are investigated. It is shown that it is possible to create biased estimation and large errors at the eavesdropper via this simple scheme. In [40], the binary stochastic encryption (BSE) approach proposed in [39] is extended to non-binary stochastic encryption (NBSE) to facilitate vector parameter estimation. In [41], secrecy provided by

stochastic encryption is studied under the assumptions that the eavesdropper is aware of the particular technique, e.g., BSE, NBSE, employed in the transmitter, uses an unbiased estimator, and does not know the encryption key and quantizer regions. It is shown that such a scheme is secure in the domain of unbiased estimators.

While the aforementioned studies focus on the stochastic encryption of a quantized measurement of a deterministic parameter, we focus on the secrecy problem for a random parameter in the Bayesian estimation setting in this dissertation. The common assumption in both Chapter 2 and 3 is that the encoding function is not available to the eavesdropper; hence, it acts like a secret key similarly to the assumption of flipping probabilities not being available to the eavesdropper in [39] and [41]. On the other hand, for determining fundamental security limits of many systems (such as those investigated in the classical information theoretical framework), it is a common practice to assume that the eavesdropper has the full knowledge of the encoding strategy at the transmitter. For example, in a Gaussian wiretap channel, the positive secrecy capacity is possible even though the eavesdropper knows the encoding scheme [18]. In particular, data is kept private as a result of the condition that the noise present in eavesdropper's received signal is stronger than the noise at the intended receiver. In that setting, the key ingredient is to apply stochastic encoding at the transmitter to achieve a positive rate with no data leakage to the eavesdropper. The encoder is used to confuse the eavesdropper with the cost of a reduced communication rate.

In Chapter 4, inspired from this classical setting, estimation theoretic secure transmission of a scalar random parameter is investigated in a Gaussian wiretap channel under the Bayesian framework, which has not been investigated in the literature. As the encoding strategy is available to the eavesdropper, the encoder randomization is allowed to increase ambiguity to possibly enhance security. The work in Chapter 4 is distinguished from that of Chapter 2 and 3 as it assumes that the mapping strategy is available to both the eavesdropper and the receiver (i.e., not secret), allows stochastic encoding in the transmitter, considers multiple observations rather than a single one, and employs different performance metrics leading to a distinct optimization problem. It is also different from those

studies (such as [39, 40]) that allow stochastic encryption as it considers direct encoding of a random parameter rather than a measured deterministic one. In Chapter 4, estimation theoretic secure transmission of a scalar random parameter is investigated in the presence of an eavesdropper in a Gaussian wiretap channel. The aim is to achieve accurate estimation of the parameter at the intended receiver while keeping the estimation error at the eavesdropper above a certain level; or, alternatively, to ensure that the estimation error at the eavesdropper is as large as possible while satisfying an estimation accuracy constraint at the intended receiver. To enhance security, stochastic encoding is employed at the transmitter, and the encoder is modeled to perform randomization between two one-to-one, continuous encoding functions, which should be designed. It is assumed that the mapping at the encoder is fully available to the eavesdropper and the receiver. For small numbers of channel observations, both the eavesdropper and the receiver are modeled to employ linear MMSE (LMMSE) estimators, and for large numbers of observations, the ECRB metric is employed both in the receiver and the eavesdropper [48]. This is because of the fact that even though the optimal estimator in terms of the MSE metric is the MMSE estimator, the calculations for its MSE have high computational complexity and do not yield closed-form expressions in general. LMMSE and ECRB tightly approximate the optimal metric for small and large numbers of observations, respectively, in our setting, and they facilitate theoretical analyses with intuitive explanations based on closed-form expressions. Therefore, based on these metrics, the optimization problems are formulated to perform optimal encoding for small and large numbers of observations separately. Both generic and affine functions are considered in the proposed encoding scheme, and a number of theoretical results on the solutions of the problems are provided.

1.3 Optimal Parameter Design for Secure Broadcast

Secure broadcast of data to multiple users is a critical issue in the secrecy literature [37],[55]–[57]. In [37], beamforming schemes are developed to ensure that legitimate users meet individual estimation error targets whereas the eavesdropper is deliberately jammed by an artificial noise component. In [57], security via regularized channel inversion precoding is investigated in a broadcast channel with confidential messages, where the transmitter broadcasts data to multiple users including potentially malicious ones and external eavesdroppers.

In Chapter 5, we consider the broadcast of a parameter to a number of low-complexity receivers with fixed estimators, where each receiver carries a certain risk of being compromised. This is because of the fact that malicious third parties can directly hijack the devices in the system or can access decoded/estimated data in certain scenarios. Our goal is to obtain an optimal parameter encoding strategy to minimize the average estimation performance at the receivers under secrecy and power constraints. To this end, each parameter is mapped using a stochastic function. In the literature, stochastic encoding of random parameters is studied for estimation problems [58], [59]; however, secrecy constraints are not considered, which become highly critical in modern systems. We show that an optimal signal design involves randomization among at most three different signal levels for each parameter value. We also provide sufficient conditions to specify when randomization can or cannot improve the optimal deterministic signaling approach.

1.4 Organization of the Dissertation

The organization of this thesis is as follows. In Chapter 2, the optimal deterministic encoding of a random scalar parameter is investigated under security constraints. In Chapter 3, secure transmission of a random vector parameter is

studied and practical deterministic encoding strategies are introduced. In Chapter 4, estimation theoretic security is investigated when the encoder at the transmitter is allowed to use a randomized mapping and the eavesdropper is fully aware of the encoding strategy. In Chapter 5, the optimal parameter design problem is studied for secure broadcast to multiple receivers with fixed estimators. Finally, the concluding remarks and possible future research directions are provided in Chapter 6.

Chapter 2

Optimal Parameter Encoding under Secrecy Constraints

In this chapter, optimal deterministic encoding of a scalar parameter is investigated in the presence of an eavesdropper [42, 43]. The main contributions of this chapter can be summarized as follows:

- First, the problem of optimal parameter encoding is proposed by considering an ECRB metric at the intended receiver and an MSE target level at the eavesdropper.
- Considering a generic prior distribution, a closed-form expression is derived for the optimal encoding function under no secrecy constraints.
- A closed form expression for $E(|\hat{\beta}(Z) - \theta|^2)$ is provided when the eavesdropper employs the linear MMSE estimator without being aware of the encoding, where $\hat{\beta}(Z)$ is the estimator of the eavesdropper and θ is the true value of the parameter. It is shown that the corresponding ECRB and MSE value do not change if the domain of the function is shifted. It is also proved that if the prior distribution is symmetric on the domain, the search for optimal encoding functions can be limited to decreasing functions. In addition,

a closed-form expression is derived for the supremum of $E(|\hat{\beta}(Z)-\theta|^2)$ over all feasible encoding functions when the prior distribution is uniform.

- Three solution approaches are proposed to find the optimal encoding function. The polynomial and piecewise linear approximations are used to calculate the optimal encoding functions numerically, and linear functions are employed to develop a suboptimal encoding scheme. It is shown that the optimal linear encoding function can be obtained simply by finding the roots of a polynomial equation. In addition, solutions are provided based on power functions in the numerical examples.

- A robust parameter encoding approach is developed. To that end, the optimization is based on the worst-case Fisher information of the uniformly distributed scalar parameter in order to guarantee a certain level of estimation accuracy at the intended receiver and an MSE target level at the eavesdropper.

- A closed-form analytical solution for robust design is obtained when the optimization problem is solved under no secrecy constraints.

- The optimal encoding function that maximizes the MSE at the eavesdropper is derived analytically for any given level of minimum Fisher information at the intended receiver. Based on this analytical result, a low-complexity algorithm is proposed to obtain the solution of the proposed optimal robust encoding problem in the presence of the MSE constraint on the eavesdropper.

- Via numerical examples, the optimal ECRB values and encoding functions are obtained based on the proposed approaches for the case of a varying target MSE level when eavesdropper's channel quality is fixed, and for the case of a varying eavesdropper's channel quality when the target MSE level is fixed. Also, a numerical example for robust encoding based on worst-case Fisher information is provided to illustrate the theoretical results and the proposed algorithm.

This chapter is organized as follows: The system model is introduced in Section 2.1. The optimal parameter encoding problem based on ECRB and worst-case Fisher information is investigated in Section 2.2 and Section 2.3, respectively. The numerical results are presented in Section 2.4, and the concluding remarks are given in Section 2.5.

2.1 System Model

Consider the transmission of a scalar parameter $\theta \in \Lambda$ to an intended receiver over a noisy and fading channel, where the noise is denoted by N_r and the instantaneous fading coefficient of the channel is denoted by the constant h_r. It is also assumed that there exists an eavesdropper trying to estimate parameter θ. The aim is to achieve accurate estimation of the parameter at the intended receiver while keeping the estimation error at the eavesdropper above a certain level. To that aim, the parameter is encoded by a continuous, real valued, and one-to-one function $f : \Lambda \to \Gamma$. Hence, the received signal at the intended receiver can be written as

$$Y = h_r f(\theta) + N_r \tag{2.1}$$

where N_r is modeled as a zero-mean Gaussian random variable with variance σ_r^2, and N_r and θ are assumed to be independent. On the other hand, the eavesdropper observes

$$Z = h_e f(\theta) + N_e \tag{2.2}$$

where h_e is the fading coefficient for the eavesdropper, and N_e is zero-mean Gaussian noise with variance σ_e^2, which is independent of θ and N_r. Also, the prior information on parameter θ is represented by a probability density function (PDF) denoted by $w(\theta)$ for $\theta \in \Lambda$. The intended receiver tries to estimate parameter θ based on observation Y whereas the eavesdropper uses observation Z for estimating θ. The system model is illustrated in Fig. 2.1. It is assumed that the

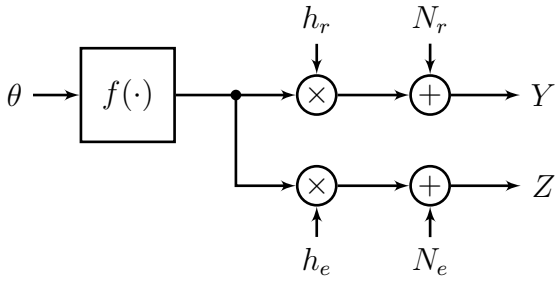

Figure 2.1: System model for the parameter encoding problem.

channels are slowly fading; that is, the channel coefficients are constant during the transmission of the parameter.[1]

The following assumptions are made about the eavesdropper's strategy:

- f acts like a secret key between the transmitter and the intended receiver and is not known by the eavesdropper. Hence, the estimator at the eavesdropper actually tries to estimate $f(\theta) \triangleq \beta$ without the knowledge of f based on observation $Z = h_e f(\theta) + N_e$.
- The eavesdropper observes a scaled and noise corrupted version of $f(\theta)$ (not θ) and it can only obtain prior information related to $f(\theta)$ (e.g., based on previous observations). It is assumed that the eavesdropper knows only the mean and the variance of $f(\theta)$, which are quite easy to obtain compared to the PDF of $f(\theta)$.
- Based on the previous assumption, the eavesdropper employs the linear MMSE estimator, which requires the prior knowledge of the mean and variance of $f(\theta)$ due to the independence of θ and N_e (see (2.24) and (2.25)).

According to this strategy, the MSE at the eavesdropper can be written as $E(|\hat{\beta}(Z) - \theta|^2)$, where $\hat{\beta}(Z)$ is the estimator of the eavesdropper and θ is the true value of the parameter. Optimal encoder design is performed based on

[1]Considering a block fading scenario in which the channel coefficients are constant for a block of transmissions [10, 11, 60, 61], the parameter encoding function should be designed for each block.

ECRB and alternatively, worst-case Fisher information using the system model described in Fig. 2.1.

2.2 ECRB Based Encoder Design

For quantifying the estimation accuracy at the intended receiver, first the ECRB will be used, as motivated in Section 1.1. The ECRB is defined as the expectation of the conditional CRB with respect to the unknown parameter [48], which is expressed as

$$E_\theta\big(I(\theta)^{-1}\big) = \int_\Lambda w(\theta)\frac{1}{I(\theta)}d\theta = ECRB \tag{2.3}$$

where $w(\theta)$ is the prior PDF of θ, $I(\theta)^{-1}$ corresponds to the conditional CRB for estimating θ,[2] and $I(\theta)$ denotes the Fisher information, i.e.,

$$I(\theta) = \int \left(\frac{\partial \log p_{Y|\theta}(y)}{\partial\theta}\right)^2 p_{Y|\theta}(y)dy \tag{2.4}$$

with $p_{Y|\theta}(y)$ representing the conditional PDF of Y for a given value of θ [47].

The aim is to minimize the ECRB at the intended receiver over the encoding function $f(\cdot)$. However, the estimation performance at the eavesdropper, which tries to estimate the parameter by using its observation Z, should also be considered. Therefore, the aim becomes the minimization of the ECRB for θ at the intended receiver while keeping the estimation error at the eavesdropper above a certain limit. Therefore, when deciding on the encoding scheme by using a one-to-one and continuous function in the presence of an eavesdropper, the average error at the eavesdropper should be considered, as well. Hence, the overall optimization problem is proposed as follows:

$$f_{opt} = \arg\min_f \int_\Lambda w(\theta)\frac{1}{I(\theta)}d\theta \quad s.t. \quad E\left(\left|\hat{\beta}(Z)-\theta\right|^2\right) \geq \alpha \tag{2.5}$$

[2]The conditional CRB presents a lower limit on the MSE of any unbiased estimator of θ based on Y for every $\theta \in \Lambda$.

where α is the MSE target at the eavesdropper and the expectation is over the joint distribution of θ and Z. In addition, the parameter space and the intrinsic constraints on the encoding function f are specified as follows:

- $\theta \in \Lambda = [a, b]$.
- $f(\theta) \in [a, b]$.
- f is a continuous and one-to-one function.

Namely, it is assumed that the parameter space is a closed set in $\mathbb{R}$ and the encoder function is an endofunction; that is, the domain and the codomain of the encoder function are the same. This is due to the practical concern that the transmitter should use the same hardware structure in the presence and absence of encoding. Furthermore, the endofunction assumption implies the peak power constraint on the encoder and it guarantees that the identity mapping $f(\theta) = \theta$ (i.e., no encoding) is a legal encoding function. It also preserves the maximum range of the parameter, $b-a$. Note that it is actually possible to impose different constraints (e.g., average power constraint, boundedness) or assumptions (e.g., stochastic encoding) on the encoding function depending on the design choice and application.

The use of the ECRB as the performance metric for the design of optimal encoding functions can be justified as follows: (i) For sufficiently high SNRs, the MSE of the MAP estimator converges to the ECRB [48]. (For low SNRs, the MAP estimator depends mainly on the prior information; hence, parameter encoding becomes ineffective.) (ii) Unlike the MSE metric, the ECRB metric does not depend on a specific estimator structure. (iii) The use of the ECRB facilitates theoretical investigations for achieving intuitive understanding of the parameter encoding problem.

2.2.1 Optimal Encoding Function

In this section, the optimization problem in (2.5) is investigated in detail. To that aim, the MSE of the eavesdropper in the constraint of (2.5) is analyzed first.

$$E\left(\left|\hat{\beta}(Z)-\theta\right|^2\right) = E\left(\left|\hat{\beta}(Z)-f(\theta)+f(\theta)-\theta\right|^2\right) \tag{2.6}$$

$$= E\left(\left|\hat{\beta}(Z)-f(\theta)\right|^2\right) + E\left(|f(\theta)-\theta|^2\right) + 2E\left(\left(\hat{\beta}(Z)-f(\theta)\right)(f(\theta)-\theta)\right). \tag{2.7}$$

It is noted from (2.7) that the MSE of the eavesdropper is determined by both the estimation error for estimating $f(\theta)$ (that is, $\hat{\beta}(Z)-f(\theta)$) and the distortion due to the encoding function (that is, $f(\theta)-\theta$). The last term in (2.7) can be written as

$$E\left(\left(\hat{\beta}(Z)-f(\theta)\right)(f(\theta)-\theta)\right)$$

$$= E_\theta E_{Z|\theta}\left(\left(\hat{\beta}(Z)-f(\theta)\right)(f(\theta)-\theta)\,|\,\theta\right) \tag{2.8}$$

$$= E_\theta\left((f(\theta)-\theta)E_{Z|\theta}\left(\hat{\beta}(Z)-f(\theta)\right)\right) \tag{2.9}$$

where E_θ denotes the expectation with respect to θ and $E_{Z|\theta}$ represents the conditional expectation with respect to Z given θ. As a special case, if the estimator of the eavesdropper, $\hat{\beta}(Z)$, satisfies $E_{Z|\theta}\left(\hat{\beta}(Z)-f(\theta)\right)=0$, $\forall\theta$, then the term in (2.9) becomes zero. This condition actually corresponds to the definition of an unbiased estimator for estimating $f(\theta)$ based on Z; i.e., $E_{Z|\theta}\left(\hat{\beta}(Z)\right)=f(\theta)$, $\forall\theta$. In other words, when the estimator of the eavesdropper is unbiased, its MSE in (2.6) simply becomes the sum of the MSE for estimating $f(\theta)$ (the first term in (2.7)) and the mean-squared distortion to θ due to the encoding function f (the second term in (2.7)).

The observations in the previous paragraph lead to an intuitive explanation of the proposed problem formulation. For example, suppose that the transmitter is to send parameter θ which is either 0 or 1 with equal probabilities, where $h_e=h_r=\sigma_e^2=\sigma_r^2=1$. In addition, the estimator at the eavesdropper is given

by

$$\hat{\beta}(Z) = \begin{cases} 1, & \text{if } Z \geq 0.5 \\ 0, & \text{otherwise} \end{cases}. \tag{2.10}$$

If the transmitter sends the parameter without any encoding; that is, if $f(\theta) = \theta$, then the MSE of the estimator at the eavesdropper can be calculated from (2.7) and (2.10) as $Q(0.5) = 0.309$ (the second and the third terms in (2.7) are zero), where $Q(x) = (1/\sqrt{2\pi}) \int_x^\infty e^{-u^2/2} du$ represents the Q-function. On the other hand, if the transmitter employs an encoding function specified by $f(\theta) = 1 - \theta$, then the MSE at the eavesdropper becomes $1 - Q(0.5) = 0.691$ (the first term in (2.7) is the same as in the previous case, but the second term is 1 and the third term is $-2\,Q(0.5)$). Hence, the eavesdropper has a higher MSE as a result of secret encoding, which is not known by the eavesdropper (i.e., the eavesdropper thinks that the transmitted value is the original parameter θ). The encoding function is known by the intended receiver, which can use this information to design its estimator accordingly. However, for a generic encoding function, there can occur a penalty at the intended receiver in terms of the estimation performance. Hence, in the design of the encoding function, the trade-off between the MSE at the eavesdropper and the estimation accuracy at the intended receiver should be considered.

To specify the Fisher information in (2.5), the conditional PDF of Y given θ is expressed from (2.1) as

$$p_{Y|\theta}(y) = \frac{1}{\sqrt{2\pi\sigma_r^2}}\, e^{-\frac{(y - h_r f(\theta))^2}{2\sigma_r^2}}. \tag{2.11}$$

Then, the Fisher information for parameter θ can be calculated via (2.4) and (2.11) as follows:

$$I(\theta) = \frac{h_r^2\, f'(\theta)^2}{\sigma_r^2} \tag{2.12}$$

where $f'(\theta)$ denotes the derivative of $f(\theta)$.

Based on (2.7) and (2.12), the optimization problem in (2.5) can be analyzed. However, before tackling the problem in (2.5), the unconstrained version of it is investigated in the next section to provide initial theoretical steps towards the analysis of the generic case.

2.2.1.1 Optimization without Secrecy Constraints

Consider the optimization problem in (2.5) without the secrecy constraint; that is, by omitting the presence of the eavesdropper. Then, the optimization problem is formulated as

$$f_{opt} = \arg\min_{f} \int_a^b w(\theta) \frac{1}{I(\theta)} d\theta \tag{2.13}$$

where $\Lambda = [a, b]$ is employed as specified in Section 2.2. Based on (2.12), the problem in (2.13) can be rewritten, by removing the constant terms, as

$$f_{opt} = \arg\min_{f} \int_a^b w(\theta) \frac{1}{f'(\theta)^2} d\theta \,. \tag{2.14}$$

The solutions of (2.14) are specified by the following proposition.

Proposition 1: *The optimal encoding functions in the absence of an eavesdropper are given by*

$$f(\theta) = a + \int_a^\theta g(\theta) d\theta \quad \textit{and} \quad f(\theta) = b - \int_a^\theta g(\theta) d\theta \tag{2.15}$$

where

$$g(\theta) \triangleq \frac{(b-a) w(\theta)^{1/3}}{\int_a^b w(\theta)^{1/3} d\theta} \,. \tag{2.16}$$

Proof: Since f is one-to-one and continuous, consider a monotonically increasing (decreasing) function with $f'(\theta) \geq 0$ ($f'(\theta) \leq 0$), $\forall \theta \in [a, b]$.[3] Also, due to the facts that $f(\theta)$ is monotone and $f(\theta) \in [a, b]$, the following relation can be obtained: $\int_a^b \frac{df}{d\theta} d\theta = f(b) - f(a) \leq b - a$ ($f(b) - f(a) \geq a - b$). Then, defining $g(\theta) \triangleq f'(\theta)$ ($g(\theta) \triangleq -f'(\theta)$), the problem in (2.14) becomes

$$\min_g \int_a^b w(\theta) \frac{1}{g(\theta)^2} \, d\theta \tag{2.17}$$

$$\text{s.t.} \int_a^b g(\theta) d\theta \leq b - a \tag{2.18}$$

$$g(\theta) \geq 0, \ \forall \theta \in [a, b] \tag{2.19}$$

Note that for all $\theta \in [a, b]$, increasing the value of $g(\theta)$ does not increase the value of the objective function; hence, the constraint in (2.18) is satisfied with equality. Now, in order to solve the optimization problem in (2.17)–(2.19), the calculus of variations is employed, and the problem is expressed in the form of

$$\min_{g \geq 0} \left\langle w, \frac{1}{g^2} \right\rangle \quad \text{s.t.} \quad \langle g, 1 \rangle = b - a \, . \tag{2.20}$$

Then, the Lagrangian is obtained as

$$L(g, \epsilon, t, \lambda) = \left\langle w, \frac{1}{(g + \epsilon t)^2} \right\rangle + \lambda \langle g + \epsilon t, 1 \rangle \tag{2.21}$$

where ϵ, t, and λ represent the perturbation, the test function and the Lagrange multiplier, respectively. The optimal solution must satisfy $\frac{\partial L}{\partial \epsilon}\big|_{\epsilon=0} = 0 \ \forall t$ [62], [63]. Hence, the following optimality condition is obtained:

$$\left\langle w, \frac{-2t}{(g + \epsilon t)^3} \right\rangle + \lambda \langle t, 1 \rangle \Big|_{\epsilon=0} = 0 \tag{2.22}$$

which leads to $\langle t, \lambda + \frac{-2w}{g^3} \rangle = 0$. In order for this to hold for all t, $g = k w^{1/3}$ must be satisfied for some constant $k \geq 0$. From the equality constraint, the constant can be calculated as $k = (b - a) \big/ \int_a^b w(\theta)^{1/3} d\theta$. Note that this $g(\theta)$ is valid, as θ

[3]Note that $f'(\theta)$ can be zero at certain points; however, it is not 0 for a closed interval in $[a, b]$ due to the one-to-one property.

takes values in $[a, b]$; hence, $w(\theta)$ is not 0 over a closed interval in $[a, b]$. Since $g(\theta) = f'(\theta)$ and $g(\theta) = -f'(\theta)$ for the monotone increasing and the monotone decreasing scenarios, respectively, the solutions can be obtained as in (2.15) and (2.16). ■

Proposition 1 states that either of the two functions given in (2.15) is an optimal solution for the minimization problem in (2.14). As a corollary to Proposition 1, if the prior distribution of the parameter is uniform over $[a, b]$, the optimal encoding functions can be found via (2.15) and (2.16) as $f(\theta) = \theta$ and $f(\theta) = a + b - \theta$. In other words, for the uniform prior, parameter encoding is not needed for reducing the ECRB at the intended receiver.

2.2.1.2 Optimization with Secrecy Constraints

In this part, the optimization problem in (2.5) is considered without omitting the secrecy constraint, where the parameter space is specified by $\Lambda = [a, b]$ as before. Although the linear MMSE estimator is assumed to be employed at the eavesdropper (see Section 2.1), a corollary to Proposition 1 is presented first for the case in which the eavesdropper employs the MMSE estimator, defined as $\hat{\beta}(z) = E(\beta|Z = z)$ with $\beta = f(\theta)$.

Corollary 1: *Suppose that the eavesdropper employs the MMSE estimator for a given encoding function $f(\theta)$. Denote the corresponding MSE at the eavesdropper as $R(f^+)$ when the encoding function is $f(\theta) = a + \int_a^\theta g(\theta)d\theta \triangleq f^+$, and as $R(f^-)$ when the encoding function is $f(\theta) = b - \int_a^\theta g(\theta)d\theta \triangleq f^-$, where $g(\theta)$ is as defined in Proposition 1. Then, the following statements hold:*

a) *If the target MSE of the eavesdropper, α in (2.5), satisfies $\alpha \leq \min\{R(f^+), R(f^-)\}$, then both f^+ and f^- are optimal encoding functions.*

b) *If $\min\{R(f^+), R(f^-)\} \leq \alpha \leq \max\{R(f^+), R(f^-)\}$, then the optimal encoding function is f^+ if $R(f^+) > R(f^-)$ and it is f^- otherwise.*

Proof: Proposition 1 implies that if f^+ or f^- is admissible by the constraint,

it becomes the minimizer of the objective function. When the eavesdropper employs the MMSE estimator, $\hat{\beta}(z) = E(\beta|Z = z)$, the MSE at the eavesdropper can be calculated from (2.7) for a given encoding function. For the special cases of encoding functions f^+ and f^-, the corresponding MSE values are denoted by $R(f^+)$ and $R(f^-)$, respectively. If α is less than both of $R(f^+)$ and $R(f^-)$, then f^+ and f^- do not violate the constraints and solve (2.5). If α is less than only one of $R(f^+)$ or $R(f^-)$, then still one of f^+ and f^- is admissible; hence, the optimal encoding function. ■

It is noted that when $\alpha \geq \max\{R(f^+), R(f^-)\}$, the shortcut provided in Corollary 1 cannot be used, and it is required to design another encoding function to satisfy the secrecy constraint.

Remark 1: The statement in Corollary 1 in fact holds for any estimator at the eavesdropper since the proof is not specific to the MMSE estimator. In other words, as long as any of the encoding functions in Proposition 1 results in an MSE at the eavesdropper that is higher than the target MSE α, that encoding function is also optimal for the problem in (2.5). Since the MMSE estimator achieves the minimum MSE among all estimators, it is concluded that if one of the encoding functions in Proposition 1 is optimal when the eavesdropper employs the MMSE estimator, then that encoding function is in fact optimal for any other estimator at the eavesdropper.

Even though the MMSE estimator is the optimal estimator according to the MSE metric, for implementing the MMSE estimator, the eavesdropper must know the prior PDF of $f(\theta)$, which can be difficult to obtain (learn). In this study, it is assumed that the eavesdropper has the knowledge of the mean and variance of $f(\theta)$. Therefore, the eavesdropper is assumed to employ the linear MMSE estimator to estimate $\beta = f(\theta)$ based on Z, as noted in Section 2.2. It is known that the linear MMSE estimator is the optimal linear estimator according to the MSE metric [64]. Furthermore, it would actually be the optimal MMSE estimator to estimate β based on Z, $E(\beta|Z = z)$, if β and Z were jointly Gaussian random variables [47]. For the system model in this chapter, the MMSE estimator and the linear MMSE estimator will have similar performance at low SNRs if the prior

is uniformly distributed.

When the linear MMSE estimator is employed at the eavesdropper, $\hat{\beta}(z)$ can be expressed as

$$\hat{\beta}(z) = k_0 + k_1 z \tag{2.23}$$

where k_0 and k_1 are chosen to minimize $E\left(\left|\hat{\beta}(Z) - \beta\right|^2\right) = E\left(\left|k_0 + k_1 Z - \beta\right|^2\right)$ as the eavesdropper does not know the encoding. The resulting coefficients for the eavesdropper's estimator are given as (see Appendix 2.6.1 for the derivation)

$$k_1 = \frac{h_e Var(\beta)}{h_e^2 Var(\beta) + \sigma_e^2} \tag{2.24}$$

$$k_0 = (1 - k_1 h_e) E(\beta). \tag{2.25}$$

Then, the resulting MSE between the estimate of the eavesdropper and the true value of parameter θ can be derived from (2.23)–(2.25) and (2.7) as (see Appendix 2.6.2 for the derivation)

$$E\left(\left|\hat{\beta}(Z) - \theta\right|^2\right) = \frac{h^2 V(V - 2C)}{h^2 V + 1} + (E(\beta) - E(\theta))^2 + Var(\theta) \tag{2.26}$$

where $\beta = f(\theta)$, $V = Var(\beta)$, $C = Cov(\beta, \theta)$, and $h = h_e/\sigma_e$.

It is observed that the MSE value at the eavesdropper corresponding to the linear MMSE estimator depends on both the encoding function and the channel quality h at the eavesdropper. It is noted that for a given encoding function with $V - 2C > 0$, the first term in (2.26) is positive, and the MSE at the eavesdropper becomes an increasing function of h^2. This means that as the channel quality for the eavesdropper improves, the resulting MSE at the eavesdropper increases in that scenario. This seemingly counterintuitive result is simply due to the fact that the estimator of the eavesdropper is based on the noisy observation of the distorted version of the original parameter. Hence, one can transmit the inflicted distortion more efficiently to the eavesdropper under good channel conditions

leading to a higher MSE. If the eavesdropper knew the prior distribution of the original parameter and realized that the transmitter sends the encoded version, it would simply stop using the observation and set $\hat{\beta}(Z) = E(\theta)$, resulting in an MSE of $Var(\theta)$, which is lower than the value in (2.26) for the case of $V - 2C > 0$. However, the eavesdropper does not have that knowledge and the channel observation is the only information it can use to estimate the parameter, which is utilized by the transmitter.

Remark 2: In the considered setting, the eavesdropper employs the linear MMSE estimator and the transmitter is aware of this situation. Then, to obtain the optimal encoding function based on (2.5), (2.12), and (2.26), the transmitter should have the knowledge of the prior PDF of the parameter and the channel quality parameter h_e^2/σ_e^2 for the eavesdropper. In practice, it can be challenging for the transmitter to have an accurate knowledge of the channel quality for the eavesdropper. In such cases, a conservative approach can be taken by either increasing the MSE target α in (2.5) or considering the worst-case (minimum) value of the MSE at the eavesdropper according to the uncertainty in the channel quality parameter.

The following proposition presents a shift invariance property for the considered problem.

Proposition 2: *Suppose that the unknown parameter θ resides in $[a, b]$ with a prior distribution specified by $w(\theta)$, and the encoding function $f(\theta) : [a, b] \to [a, b]$ results in a certain ECRB at the intended receiver and a certain MSE at the eavesdropper, which employs the linear MMSE estimator. If the parameter θ were defined in $[0, b-a]$ with the prior distribution $\hat{w}(\theta) = w(\theta + a)$, then the use of the encoding function $\hat{f}(\theta) : [0, b-a] \to [0, b-a]$ such that $\hat{f}(\theta) = f(\theta + a) - a$ would result in the same MSE at the eavesdropper and the same ECRB at the intended receiver as in the original scenario.*

Proof: The ECRB in the original scenario can be expressed from (2.12) and

(2.13) as

$$\frac{\sigma_r^2}{h_r^2}\int_a^b w(\theta)\frac{1}{f'(\theta)^2}d\theta \tag{2.27}$$

which is equivalent to

$$\frac{\sigma_r^2}{h_r^2}\int_0^{b-a} w(\theta+a)\frac{1}{((f(\theta+a)-a)')^2}d\theta \tag{2.28}$$

since $(f(\theta+a)-a)' = f'(\theta+a)$. As the expression in (2.28) corresponds to the ECRB in the second scenario, the equivalence of the ECRBs is established. To prove that the MSE at the eavesdropper does not change, it is noted that the parameter defined in $[0, b-a]$ with the prior distribution $\hat{w}(\theta) = w(\theta + a)$ corresponds to shifting the original parameter as $\theta - a$. Also, let $\bar{\beta}$ and β denote the random variables for the encoded versions of the shifted and original parameters via encoding functions $\bar{f}(\theta)$ and $f(\theta)$, respectively. Then, $\bar{\beta} = \beta - a$ holds. Furthermore, it is noted that shifting the specified random variables (θ and $\beta = f(\theta)$) just changes their means by the amount of the shift without modifying the second order statistics V and C. Hence, (2.26) reveals that the MSE at the eavesdropper stays the same as in the original scenario after the shift operations. ■

Based on Proposition 2, the estimation of a parameter in $\theta \in [0, b-a]$ can be considered without loss of generality for the case of the linear MMSE estimator at the eavesdropper (see Proposition 4).

The next proposition states that when the prior PDF of $\theta \in [a, b]$ is symmetric around $(a+b)/2$, parameter encoding via a strictly decreasing function is more desirable than that via a strictly increasing one.

Proposition 3: *Suppose that the eavesdropper employs the linear MMSE estimator and $w(\theta)$ is symmetric around $(a+b)/2$. Then, for any given continuous and strictly increasing encoding function, there exists a corresponding continuous and strictly decreasing encoding function that yields the same ECRB at the intended receiver with a higher MSE at the eavesdropper.*

Proof: Consider two encoding functions $f(\theta)$ and $s(\theta) = f(a+b-\theta)$, where $\theta \in [a, b]$ and $f(\theta)$ is a continuous and monotonically increasing function. Since $w(\theta) = w(a+b-\theta)$ due to the symmetry assumption and $s'(\theta) = -f'(a+b-\theta)$ by definition, both encoding functions result in the same ECRB, which can be proved via (2.14) as follows:

$$\begin{aligned}\int_a^b w(\theta)\frac{1}{s'(\theta)^2}d\theta &= \int_a^b w(a+b-\theta)\frac{1}{f'(a+b-\theta)^2}d\theta \\ &= \int_a^b w(\theta)\frac{1}{f'(\theta)^2}d\theta \end{aligned} \tag{2.29}$$

where the final expression is obtained via a change of variables. To compare the MSEs corresponding to the two encoding functions, define $\beta_f \triangleq f(\theta)$ and $\beta_s \triangleq s(\theta)$, and let $p_{\beta_f}(x)$ and $p_{\beta_s}(x)$ represent the PDFs of β_f and β_s, respectively. Then, it is noted that $p_{\beta_f}(x) = p_{\beta_s}(x)$ for $x \in [a, b]$ since $w(\theta) = w(a+b-\theta)$ due to symmetry. Hence, both β_f and β_s have the same expectation and the variance. For the covariance, $Cov(\beta, \theta) = E(\beta\theta) - E(\beta)E(\theta)$, the following expression can be obtained:

$$\begin{aligned}&E(\beta_f\theta) - E(\beta_s\theta) \\ &= \int_a^b w(\theta)f(\theta)\theta d\theta - \int_a^b w(\theta)f(a+b-\theta)\theta d\theta \qquad (2.30)\\ &= \int_a^b w(\theta)f(\theta)(2\theta - a - b)d\theta. \qquad (2.31)\end{aligned}$$

where (2.30) follows from the definitions of β_f and β_s, and (2.31) is obtained from the symmetry of $w(\theta)$. Since $f(\frac{a+b}{2} - x) < f(\frac{a+b}{2} + x)$ for $x \in (0, \frac{a+b}{2}]$, $E(\beta_f\theta) - E(\beta_s\theta) > 0$. Then, $Cov(\beta_f, \theta) > Cov(\beta_s, \theta)$ and $E(|\hat{\beta}_f - \theta|^2) < E(|\hat{\beta}_s - \theta|^2)$ according to (2.26). Therefore, it is always possible to achieve a higher MSE by employing $s(\theta)$ instead of $f(\theta)$ while keeping the ECRB the same. ■

Proposition 3 implies that it is sufficient to search for the optimal encoding function among strictly decreasing functions if the prior distribution of the parameter satisfies the symmetry condition (e.g., the uniform distribution). This is based on the idea that for any given increasing encoding function that solves

(2.5), there exists a legitimate decreasing function obtained by a simple transformation, which yields the same optimal ECRB value with an increased MSE at the eavesdropper. Hence, from a practical point of view, the search space for the optimal encoding function can be confined to strictly decreasing functions under the conditions in the proposition.

2.2.1.2.1 Special Case: Uniform Prior Distribution For the special case of a uniform prior distribution, the following result characterizes the optimal encoding function when the eavesdropper employs the linear MMSE estimator.

Corollary 2: *Suppose that the parameter has uniform prior distribution over* $[a, b]$ *and the eavesdropper employs the linear MMSE estimator. Then, if the target MSE* α *satisfies* $\alpha \leq \frac{V_u}{h^2 V_u + 1}$, *then* $f(\theta) = \theta$ *is an optimal encoding function, where* $V_u \triangleq (b-a)^2/12$. *On the other hand, if* $\alpha \leq \frac{4h^2 V_u^2 + V_u}{h^2 V_u + 1} + (a + b - 2E(\theta))^2$, *then* $f(\theta) = a + b - \theta$ *is an optimal encoding function.*

Proof: From the expressions in Proposition 1, it can be shown, for the uniform prior distribution, that either of $f(\theta) = \theta$ or $f(\theta) = a + b - \theta$ is an optimal encoding function in the absence of the constraint (i.e., in the absence of the eavesdropper). When the eavesdropper employs the linear MMSE estimator, the use of $f(\theta) = \theta$ leads to an MSE of $\frac{V_u}{h^2 V_u + 1}$ and the use of $f(\theta) = a + b - \theta$ results in an MSE of $\frac{4h^2 V_u^2 + V_u}{h^2 V_u + 1} + (a + b - 2E(\theta))^2$, which can be derived based on (2.26). Then, based on similar arguments to those in Corollary 1, it is deduced that if the MSE corresponding to $f(\theta) = \theta$ is larger than or equal to α, $f(\theta) = \theta$ is an optimal encoding function. Similarly, if the MSE for $f(\theta) = a + b - \theta$ is larger than or equal to α, $f(\theta) = a + b - \theta$ is an optimal encoding function. ■

The following proposition provides an upper bound on the MSE at the eavesdropper, which employs the linear MMSE estimator, when the parameter has uniform prior distribution.

Proposition 4: *If the eavesdropper employs the linear MMSE estimator and*

θ has uniform distribution over $[0, \gamma]$, then

$$\sup_f E\left(\left|\hat{\beta}(Z) - \theta\right|^2\right) = \begin{cases} \dfrac{\gamma^2}{3}, & \gamma \leq \dfrac{2}{h} \\ \dfrac{h^2\gamma^4}{2h^2\gamma^2 + 8} + \dfrac{\gamma^2}{12}, & \gamma > \dfrac{2}{h} \end{cases} \tag{2.32}$$

where $f(\theta) : [0, \gamma] \to [0, \gamma]$ is a continuous and one-to-one function.

Proof: For an encoding function $f(\theta) = \beta$, let $V = Var(\beta)$, $C = Cov(\beta, \theta)$, and $\mu = E(\beta)$. It can be shown that for a random variable defined on the bounded interval of $[0, \gamma]$, the following relations hold for $\mu \in [0, \gamma]$:

$$0 \leq V \leq \mu(\gamma - \mu) \leq \frac{\gamma^2}{4}\cdot \tag{2.33}$$

In addition, C can be expressed as

$$C = \frac{1}{\gamma}\int_0^\gamma f(\theta)\left(\theta - \frac{\gamma}{2}\right)d\theta.$$

For a given continuous endofunction on $[0, \gamma]$, it can be shown that C is in $(-\gamma^2/8, \gamma^2/8)$. Also, from (2.26), the MSE at the eavesdropper can be stated as

$$\begin{aligned} E\left(\left|\hat{\beta}(Z) - \theta\right|^2\right) &= \frac{h^2V(V - 2C)}{h^2V + 1} + \left(\mu - \frac{\gamma}{2}\right)^2 + \frac{\gamma^2}{12} \\ &\leq \frac{h^2V(V - 2C)}{h^2V + 1} - V + \frac{\gamma^2}{3} \\ &= \frac{-V(1 + 2h^2C)}{h^2V + 1} + \frac{\gamma^2}{3} \end{aligned} \tag{2.34}$$

where the inequality holds for any continuous encoding function defined on $[0, \gamma]$. Therefore, the upper bound on $E(|\hat{\beta}(Z) - \theta|^2)$, specified in (2.34), holds for all possible encoding functions. Next, the maximum of this generic upper bound is to be found over the PDF of β, denoted by p_β, where $\beta = f(\theta)$. It is observed that if $(1+2h^2C) > 0$ for any given p_β, the first term in (2.34) is nonpositive as $V \geq 0$ and $h > 0$; hence, the maximum of the upper bound is $\gamma^2/3$. When $(1 + 2h^2C) \leq 0$, then the first term in (2.34) is maximized by increasing V and decreasing C at the same time. Therefore, if $V = \gamma^2/4$ and $C = -\gamma^2/8$, the maximum of the

upper bound is achieved. Thus, for $\gamma \leq 2/h$, $E(|\hat{\beta}(Z) - \theta|^2) \leq \gamma^2/3$ holds, and for $\gamma > 2/h$,

$$E(|\hat{\beta}(Z) - \theta|^2) \leq \frac{h^2\gamma^4}{2h^2\gamma^2 + 8} + \frac{\gamma^2}{12} \tag{2.35}$$

is obtained. Furthermore, in the first case (i.e., $\gamma \leq 2/h$), $\tilde{f}(\theta) = 0$ (or, $\tilde{f}(\theta) = \gamma$) for $\theta \in [0, \gamma]$ attains the upper bound on the MSE, that is, $\gamma^2/3$. In the second case (i.e., $\gamma > 2/h$), it is possible to achieve the upper bound in (2.35) by using $\tilde{f}(\theta)$ defined as

$$\tilde{f}(\theta) = \begin{cases} \gamma, & 0 \leq \theta \leq \gamma/2 \\ 0, & \gamma/2 < \theta \leq \gamma \end{cases}. \tag{2.36}$$

Even though the maximum values for the upper bounds are obtained and it is argued that they are exactly attained by using $\tilde{f}(\theta)$, it should be noted that $\tilde{f}(\theta)$'s are not in the feasible function set as they do not satisfy the one-to-one and continuity properties. However, it is possible to approach arbitrarily close to $\tilde{f}(\theta)$ while staying in the feasible function set (e.g., take $\delta > 0$, set $f(\theta) = \gamma - \delta\theta$, and let $\delta \to 0$ for the first case). Furthermore, the objective is continuous functional acting on the encoding function. Hence, the upper bound values for the MSE cannot exactly be achieved; however, one can get arbitrarily close to them by employing one-to-one and continuous functions, which yield them as the supremum values for the MSE at the eavesdropper, resulting in the expression in (2.32). ■

In addition to providing a closed form upper bound on the distortion at the eavesdropper, Proposition 4 plays another important role by helping us gain practical intuition about the behavior of the optimal encoding function (*variance minimizing mode* or *variance maximizing mode*). As argued in Proposition 3, if there are two alternative encoding functions with the same ECRB, then it is better to choose the one which yields the higher MSE. Note that given an encoding function $f(\theta)$, one can shuffle the increments of the given function and end up with an alternative encoding function with the same ECRB. The alternative

encoding function will possibly have different (μ, V, C). Therefore, it is important to understand how the MSE behaves as μ, V, and C change. Note that in (2.32), the supremum changes depending on the value of the channel quality of the eavesdropper h for a given γ ($h \leq 2/\gamma$ or $h \geq 2/\gamma$). Let us investigate those two cases:

- If the channel quality h is small enough, one can let $\beta \to 0$ (or γ) to maximize its MSE. This strategy is equivalent to minimizing the variance and letting $E(\beta) \to 0$ (or γ).
- If the channel quality h is large enough, then one can increase the variance V and decrease C at the same time to maximize its MSE (effectively maximize $E(|\beta - \theta|^2)$).

This discussion becomes clearer at the extreme cases of the value of h^2. For example, suppose that h^2 is very small. Then, (2.26) reveals that $E(|\hat{\beta}(Z) - \theta|^2) \approx (E(\beta) - E(\theta))^2 + Var(\theta)$; hence, it is possible to generate a larger MSE by making $E(\beta)$ as close to the boundaries 0 and γ as possible. This behavior can be regarded as the *variance minimizing mode*. If h^2 is very large, then $E(|\hat{\beta}(Z) - \theta|^2) \approx (V - 2C) + (E(\beta) - E(\theta))^2 + Var(\theta) = E((\beta - \theta)^2) + E(\theta)^2$; hence, it is possible to generate a higher MSE by maximizing $E((\beta - \theta)^2)$. This behavior can be regarded as the *variance maximizing mode*. Of course, if the resulting MSE is higher than the target MSE α for a given h, then one can use linear encoding $f(\theta) = \gamma - \theta$ for minimizing the ECRB.

It is important to note that Proposition 4 does not have any constraints on the ECRB. The original problem tries to minimize the ECRB for a target MSE. Among two candidates with the same ECRB, the one yielding the larger MSE at the eavesdropper is preferred in the search of the optimal encoder. The feasible set for (μ, V, C) is specific to that ECRB value. For example, one might not be able to let $\mu \to 0$ anymore. However, one can generate a larger MSE by making μ very close to its limit in the feasible set for sufficiently small h values (e.g., by using a decreasing concave function) or by making $E(|\beta - \theta|^2)$ as large as possible for sufficiently high h values (e.g., by using a decreasing concave

function for $\beta < \gamma/2$ and a decreasing convex function for $\beta > \gamma/2$). Hence, the optimal encoding function will be in either of the modes (variance minimizing or maximizing) described above.

Remark 3: The optimal value of the optimization problem in (2.5) can be named as $G(\alpha)$ since the optimal ECRB value depends on the target MSE α. That is,

$$G(\alpha) \triangleq \frac{\sigma_r^2}{h_r^2} \int_a^b w(\theta) \frac{1}{f'_{opt}(\theta)^2} d\theta \tag{2.37}$$

where $f_{opt}(\theta)$ is a solution to (2.5). Note that the optimal value of the ECRB in the case of optimization without secrecy constraints can be denoted by $G(0)$. Then, $G(\alpha)$ has the following properties:

- $G(\alpha)$ is constant between $0 \leq \alpha \leq \alpha_{th}$ with $\alpha_{th} = \max\{R(f^+), R(f^-)\}$, where $R(f^+)$ and $R(f^-)$ can similarly be defined as in Corollary 1 except that the linear MMSE is employed at the eavesdropper.
- $G(\alpha)$ is a non-decreasing function between $\alpha_{th} \leq \alpha < \alpha_{max}$, where $\alpha_{max} = \sup_f E\left(\left|\hat{\beta}(Z) - \theta\right|^2\right)$.
- $G(\alpha) \to \infty$ as $\alpha \to \alpha_{max}$.

The second property follows from the following argument: Let S_{α_1} and S_{α_2} be the feasible sets for α_1 and α_2, respectively. If $\alpha_1 \geq \alpha_2$, then $S_{\alpha_1} \subseteq S_{\alpha_2}$; hence, $G(\alpha_1) \leq G(\alpha_2)$. Note that a closed form expression for α_{max} is provided in Proposition 4 for the special case of uniform prior distribution.

2.2.2 Solution Approaches

In general, the optimal parameter encoding problem formulated by (2.5), (2.12), and (2.26) is a difficult optimization problem as it requires a search over functions. Although the theoretical results in the previous section can lead to closed-form

solutions or reductions in the search space in certain scenarios, it may still be necessary to solve the problem directly in some cases. Therefore, various solution approaches are developed in this section for obtaining suboptimal solutions of (2.5). In the proposed approaches, it is assumed that the encoding function f is picked among a family of functions characterized by a certain number of parameters. Then, the optimization problem becomes easier to solve as it involves minimization over a few variables (instead of functions), which also leads to some analytical solutions, as discussed below. However, the obtained encoding function will be suboptimal in general since the actual solution of (2.5) may not be a function from the assumed family of functions.

2.2.2.1 Linear Encoding Functions

One suboptimal encoding scheme is to employ a linear encoding function to minimize the ECRB at the intended receiver while satisfying the MSE constraint at the eavesdropper. To obtain analytical results for generic prior PDFs, the eavesdropper is modeled to employ the linear MMSE estimator as before, and the encoding function is assumed to be a decreasing linear function. However, the analysis can also be performed easily for increasing linear functions in a similar fashion, which yields similar analytical results to those in Proposition 5 and afterwards. (In practice, it is advised to solve the encoding problem restricted to decreasing linear functions and to increasing linear functions separately, and select the one with the lower objective value. However, when the prior PDF of the parameter, $w(\theta)$, is symmetric around $(a+b)/2$, where $\theta \in [a,b]$, it is sufficient to consider decreasing functions only, as shown in Proposition 3.)

For the considered model, the linear encoding function can be expressed as

$$f(\theta) = c_0 + m(b-\theta) \tag{2.38}$$

where $m \in (0,1]$, $c_0 \geq a$, and $c_0 + m(b-a) \leq b$. In other words, for a fixed m, c_0 can be any real number in $[a, b-m(b-a)]$. In addition, the random variable $\beta = f(\theta)$ has the following PDF: $p_\beta(x) = \frac{1}{m} w\big(\frac{c_0+mb-x}{m}\big)$ for $x \in [c_0, c_0+m(b-a)]$.

For example, if $w(\theta)$ is the uniform PDF over $[a, b]$, then β will have uniform distribution over $[c_0, c_0 + m(b-a)]$; hence, its amplitude is $\frac{1}{m(b-a)}$ inside that interval and 0 elsewhere. Also, the value of c_0 does not change this amplitude but only causes a shift in the domain of β. First, the following proposition is presented about c_0 for any given input distribution $w(\theta)$.

Proposition 5: *When the eavesdropper employs the linear MMSE estimator, the MSE at the eavesdropper for the linear encoding function $f(\theta) = c_0 + m(b-\theta)$ is a convex function of c_0 for a fixed $m > 0$. Hence, the MSE is maximized either at $c_0 = a$ (if $E(\theta) > (a+b)/2$) or at $c_0 = b - m(b-a)$ (if $E(\theta) < (a+b)/2$).*

Proof: The variance and the mean of $\beta = f(\theta)$ can be calculated as $Var(\beta) = m^2 Var(\theta)$ and $E(\beta) = c_0 + mb - mE(\theta)$. Also, the covariance of β and θ can be obtained as $Cov(\beta, \theta) = -mVar(\theta)$. In (2.26), only the second term depends on c_0. In addition, $(E(\beta) - E(\theta))^2 = (c_0 - (E(\theta)(1+m) - mb))^2$ is a convex function of c_0 for a fixed m, and it is equal to $(a - E(\theta) - m(E(\theta) - b))^2$ at $c_0 = a$ and $(b - E(\theta) - m(E(\theta) - a))^2$ at $c_0 = b - m(b-a)$. Hence, for a given $m \in (0, 1)$, the MSE is maximized either at $c_0 = a$ if $E(\theta) > (a+b)/2$ or at $c_0 = b - m(b-a)$ if $E(\theta) < (a+b)/2$. (If $m = 1$ or $E(\theta) = (a+b)/2$, it has the same value at both of the boundaries, hence, there exist two maximizers in that case.) ■

Proposition 5 leads to the closed-form solution for the optimal linear encoding function as follows: Since the ECRB expression depends only on the derivative of the encoding function (see (2.5) and (2.12)), it is proportional to $1/m^2$ for the linear encoding function in (2.38); hence, it does not depend on c_0. Therefore, c_0 can be chosen to maximize the MSE at the eavesdropper based on Proposition 5, which implies that c_0 is equal to either a or $b - m(b-a)$ (which corresponds to either $f(b) = a$ or $f(a) = b$). Based on these observations, it is sufficient to perform a search only over parameter m in order to determine the optimal linear encoding function. Suppose that $E(\theta) > (a+b)/2$ and model the linear encoding function as $f(\theta) = a + m(b-\theta)$ (see Proposition 5). (The case of $E(\theta) < (a+b)/2$ and $f(\theta) = b + m(a-\theta)$ can be treated similarly.) Then, the optimization problem

specified by (2.5) and (2.12) can be rewritten to find the optimal m as follows:

$$m_{opt} = \arg\min_{m} \frac{1}{m^2} \text{ s.t. } E\Big(\big|\hat{\beta}(Z) - \theta\big|^2\Big) \geq \alpha,\ 0 < m \leq 1 \tag{2.39}$$

where $E(|\hat{\beta}(Z) - \theta|^2) = \frac{h^2 m^2 V(m^2 V + 2mV)}{h^2 m^2 V + 1} + V + (a - E(\theta) - m(E(\theta) - b))^2$ with $V = Var(\theta)$ due to (2.26). Obviously, the optimal m is the largest m that satisfies the constraints. After some algebra, the first constraint can be expressed as

$$\begin{aligned} &\left(tV^2 + \kappa_1^2 tV\right) m^4 + \left(2tV^2 + 2tV\kappa_1\kappa_2\right) m^3 \\ &+ \left(tV^2 + (\kappa_2^2 - \alpha)tV + \kappa_1^2\right) m^2 + (2\kappa_1\kappa_2)m \\ &+ \left(\kappa_2^2 + V - \alpha\right) \geq 0 \end{aligned} \tag{2.40}$$

where $t \triangleq h^2$, $\kappa_1 \triangleq b - E(\theta)$, and $\kappa_2 \triangleq a - E(\theta)$. Hence, the optimal m is the largest m in $(0, 1]$ satisfying (2.40). This optimal value can be obtained algebraically by finding the roots of the fourth degree polynomial in (2.40). For example, when $h = 1$, $a = 0$, $b = 1$, $w(\theta)$ is uniform, and $\alpha = 0.15$, (2.40) becomes $m^4 - m^3 + 9.55m^2 - 18m + 6.6 \geq 0$. This polynomial has roots at 1.3001, 0.4915, and $-0.3958 \pm 3.1895i$, implying that the constraint holds when $m \geq 1.3001$ or $m \leq 0.4915$; thus, the optimal m is given by $m = 0.4915$. Overall, it is concluded that considering an encoding function among the family of linear functions, the optimal solution can be obtained by finding the roots of a polynomial equation without performing any functional optimization.

Remark 4: One alternative approach could be to consider an encoding function in the form of $f(\theta) = a + p\left(\frac{b-\theta}{b-a}\right)^q$, where the function is parameterized by p and q. Hence, instead of trying to optimize over functions, one can try to use this family of *power functions*, and perform optimization over $p \in (0, b - a]$ and $q \in (0, 3/2)$. Even though this will lead to a suboptimal encoding function, it is still easier to perform optimization via a 2-dimensional search than optimizing over functions as in (2.5). On the other hand, this approach will have higher computational complexity than the one that employs (2.38).

2.2.2.2 Polynomial Approximation

The second approach for obtaining a suboptimal solution of (2.5) is to use a polynomial approximation method. Approximating a function via polynomials is a well-known numerical analysis method [65]–[66]. To apply this method to the parameter encoding problem, it is assumed that the encoding function is in the form of a polynomial. In fact, any continuous real-valued function defined on $[a, b]$ can be uniformly approximated by polynomials in that interval [67]. That is, for a given continuous and bounded function $f(x)$ and $\epsilon > 0$, there exists a polynomial $P(x)$ on $[a, b]$ such that $\sup_x |f(x) - P(x)| < \epsilon$. Motivated by this fact, the encoding function is expressed by Kth degree polynomials, i.e., $P(x) = \sum_{n=0}^{K} c_n x^n$, and the aim becomes the calculation of the optimal coefficients c_n for $n = 0, 1, \ldots, K$. Hence, by using $f(\theta) = \sum_{n=0}^{K} c_n \theta^n$, the optimization problem specified by (2.5) and (2.12) can be rewritten to find the optimal coefficients as follows:

$$\begin{aligned} \mathbf{c}^{\text{opt}} = & \underset{c_0, c_1, \ldots, c_K}{\arg\min} \int_{\Lambda} w(\theta) \left(\sum_{n=0}^{K} n c_n \theta^{n-1} \right)^{-2} d\theta \\ & \text{s.t.} \quad E\left(\left| \hat{\beta}(Z) - \theta \right|^2 \right) \geq \alpha \end{aligned} \tag{2.41}$$

After finding the optimal coefficients, the encoding function can be written as $f_{\text{opt}}(\theta) = \sum_{n=0}^{K} c_n^{\text{opt}} \theta^n$, where c_n^{opt} represents the nth element of $\mathbf{c}^{\text{opt}}$. Note that the resulting encoding function should also satisfy the implicit conditions, that is, $f(\theta) \in [a, b]$ and the monotonicity.

2.2.2.3 Piecewise Linear Approximation:

Finally, a third approach is proposed, which is based on the idea that any continuous bounded function can be uniformly approximated by piecewise linear functions. Therefore, the parameter space $[a, b]$ is partitioned into M intervals and the

optimal increment (or, decrement) is found in each interval, which results in an approximation of the encoding function f via a piecewise linear function. In particular, the increments/decrements are defined as $\Delta x_k = f(a+k\Delta\theta) - f(a+(k-1)\Delta\theta)$, and the optimization is performed over M variables, $\Delta x_1, \Delta x_2, \ldots, \Delta x_M$. As M increases, more accurate approximation is achieved; however, the computational complexity of solving the optimization problem increases, as well. Note that, for $M = 1$, this approach reduces to the linear encoding function case in Section 2.2.2.1. The optimization problem specified by (2.5) and (2.12) can be stated to find the optimal increments as follows:

$$\begin{aligned} \mathbf{\Delta x}_{\text{opt}} = & \underset{\Delta x_1, \Delta x_2, \ldots, \Delta x_M}{\arg\min} \sum_{k=1}^{M} \frac{1}{\Delta x_k^2} \int_{a+(k-1)\Delta\theta}^{a+k\Delta\theta} w(\theta) d\theta \\ & \text{s.t.} \quad E\left(\left|\hat{\beta}(Z) - \theta\right|^2\right) \geq \alpha \end{aligned} \tag{2.42}$$

Similar to the previous case, the resulting encoding function should also satisfy the implicit conditions, that is, $f(\theta) \in [a, b]$ and the monotonicity. For example, if a decreasing encoding function is used, then all the elements in $\mathbf{\Delta x}_{\text{opt}}$ should be negative. In order to solve the problems given in (2.41) and (2.42), we have used the Global Optimization Toolbox of MATLAB. As the initial point, the linear solution, which is calculated analytically, can be used. It is noted that the objective function given in (2.14) is a convex operation on f; however, the feasible set does not need to be convex. This discussion holds for both of the problems in (2.41) and (2.42).

Remark 5: Most of the theoretical results in this chapter can be extended, under certain conditions, to scenarios in which the eavesdropper employs an arbitrary affine estimator, $\hat{\beta}(z) = R_0 + R_1 z$, instead of the linear MMSE estimator. In this case, after some manipulation, the MSE of the eavesdropper can be obtained for given R_1 and R_0 as

$$\begin{aligned} E\left(\left|\hat{\beta}(Z) - \theta\right|^2\right) = & R_1^2(h_e^2 V + \sigma_e^2) - 2R_1 h_e C + Var(\theta) \\ & + (R_1 h_e E(\beta) - E(\theta) + R_0)^2 \end{aligned} \tag{2.43}$$

where $V = Var(\beta)$ and $C = Cov(\beta, \theta)$. Then, the results can be extended as

follows:

- Proposition 2 does not hold for general R_1 and R_0. However, for the special case of $R_1 = 1/h_e$, it holds for any R_0, and for $R_0 = E(\beta)(1-R_1h_e)$, it holds for any R_1. It is noted that the second case implies that $E(\hat{\beta}(z)) = E(\beta)$.

- Proposition 3 holds if $R_1h_e > 0$. If $R_1h_e < 0$, then the reverse of the argument holds; that is, for a given strictly decreasing function, one can find a simple transformation such that the resulting encoding function has a lower MSE. Corollary 2 can also be generalized in a similar fashion.

- Proposition 4 is particular to the assumption of the linear MMSE estimator; hence, it cannot be generalized directly for arbitrary R_1 and R_0. However, an upper limit can be found as follows by considering R_1 and R_0 as given constants:

$$\sup_f E\left(\left|\hat{\beta}(Z)-\theta\right|^2\right) = \sup_f \Big(E\left(\left|R_1h_e\beta-\theta\right|^2\right) + 2R_1R_0h_eE(\beta) + g(R_0,R_1)\Big)$$

  where $g(R_0,R_1) = R_1^2 - 2R_0E(\theta)$. Next, let $R_1h_e = k$ and $k > 0$. Then, for a fixed $E(\beta) = \alpha$ with $\alpha \in [0,\gamma]$, $E\left(\left|k\beta-\theta\right|^2\right)$ is maximized if $\beta = \gamma$ for $\theta < \alpha$ and 0 otherwise. Then, the analysis can be completed by finding the optimal α.

- Finally, if $R_1h_e > 0$, Proposition 5 can also be generalized. Namely, the MSE is a convex function of c_0 for a fixed $m > 0$ and is maximized either at $c_0 = a$ or $c_0 = b - m(b-a)$.

2.3 Worst-Case Fisher Information Based Encoder Design

In this part, a robust approach is proposed for the optimal parameter encoding design and the worst-case (maximum) CRB is used for quantifying the estimation accuracy at the intended receiver and the system model given in Fig. 2.1 is employed. It is assumed that the parameter θ has uniform distribution over Λ. The aim is to minimize the maximum CRB over the parameter set via an encoding function while keeping the MSE at the eavesdropper (which employs the LMMSE estimator) above a certain target value. Hence, the following problem formulation is proposed:

$$f_{opt} = \arg\min_{f} \max_{\theta} \big(I(\theta)\big)^{-1} \; s.t. \; E\big(\big|\hat{\beta}(Z) - \theta\big|^2\big) \geq \eta \tag{2.44}$$

where $\hat{\beta}(Z)$ is the LMMSE estimator employed at the eavesdropper, η is the MSE target for the eavesdropper, $(I(\theta))^{-1}$ represents the CRB, and $I(\theta)$ denotes the Fisher information given in (2.4). The problem in (2.44) can also be stated as

$$f_{opt} = \arg\max_{f} \min_{\theta} I(\theta) \; s.t. \; E\big(\big|\hat{\beta}(Z) - \theta\big|^2\big) \geq \eta \tag{2.45}$$

which means that the aim is to maximize the minimum (worst-case) Fisher information at the intended receiver. It is noted that the distribution of θ does not affect the objective function in (2.45) since the worst-case parameter value is the main concern.

As motivated in Section 2.2, the parameter space and the intrinsic constraints on the encoding function f are specified as follows:

- $\theta \in \Lambda = [a, b]$.
- $f(\theta) \in [a, b]$.
- f is a continuous (except at a finite number of points) and one-to-one function.

2.3.1 Optimal Encoding Function and Solution Algorithms

In this section, the solution of the proposed problem in (2.45) (equivalently, in (2.44)) is investigated in the absence and in the presence of the secrecy constraint similarly to Section 2.2.1.

2.3.1.1 Optimization without Secrecy Constraint

Consider the optimization problem in (2.45) without the secrecy constraint; i.e., in the absence of the eavesdropper. From (2.12), the problem in (2.45) can be expressed by removing the constant terms as

$$f_{opt}(\theta) = \arg\max_{f} \min_{\theta} f'(\theta)^2 \,. \tag{2.46}$$

The following proposition is presented related to the solutions of (2.46).

Proposition 6: *The optimal continuous encoding functions in the absence of an eavesdropper are*

$$f(\theta) = a + b - \theta \quad \textit{and} \quad f(\theta) = \theta \,. \tag{2.47}$$

Proof: Let T denote an operator on $f(\theta)$ such that $T(f) = \min_\theta f'(\theta)^2$. It is given that f is one-to-one but not necessarily a monotone function over $[a, b]$ due to the possibility of discontinuous points. However, f has to be monotone over the interval between any two consecutive discontinuous points as it is one-to-one. Thus, for any one-to-one function f, there exists a monotone function f_m such that $T(f) = T(f_m)$, which can be generated by adjusting the signs of the derivatives without changing their absolute values. Hence, it can be assumed without loss of generality that f is a monotone function. Furthermore, it is noted that since f is not differentiable at discontinuous points and $T(f)$ is the pointwise minimum of $f'(\theta)^2$, the points at which the jumps occur cannot be the optimal points. Therefore, one can remove the jumps at the discontinuities to obtain a

continuous version, denoted by f_c. Thus, for any one-to-one function f, there exists a continuous function f_c such that $T(f) = T(f_c)$; hence, it can also be assumed that f is a continuous function without any loss. First, consider the case of $f'(\theta) > 0$, $\forall \theta \in [a, b]$. Then, based on the properties of the encoding function f, $\int_a^b \frac{df}{d\theta} d\theta = f(b) - f(a) \leq b - a$. Let $g(\theta)$ be defined as $g(\theta) \triangleq f'(\theta)$. Then, the problem in (2.46) becomes $\max_g \min_\theta g(\theta)^2$ subject to $\int_a^b g(\theta) d\theta \leq b-a$ and $g(\theta) > 0$. Consider the function $g^*(\theta) = 1$, $\forall \theta \in [a, b]$, which satisfies both of the constraints. Next, suppose that there exists a function h with $\min_\theta h(\theta) > 1$. Then, $\int_a^b h(\theta) d\theta > b - a$, leading to a violation of the constraint. Hence, for any given function g, there is an upper bound specified as $\min_\theta g(\theta) \leq 1$. Since the constant function satisfies this upper bound, it is the maximizer over all possible functions. Since $g(\theta) = 1$ for $\theta \in [a, b]$, it is obtained that $f(\theta) = \theta$ is an optimal solution. For the case of $f'(\theta) < 0$, let $g(\theta) \triangleq -f'(\theta)$. Then, based on similar arguments, $g(\theta) = 1$ can be obtained, resulting in an optimal solution of $f(\theta) = a + b - \theta$.[4] ■

Proposition 6 reveals that if there exist no secrecy constraints, parameter encoding does not provide any benefits in terms of the worst-case Fisher information as $f(\theta) = \theta$ is an optimal solution.

2.3.1.2 Optimization with Secrecy Constraint

To obtain the optimal encoding function in the presence of the secrecy constraint, the problem in (2.45) can be rewritten, based on (2.12), as

$$f_{opt}(\theta) = \arg \max_f \min_\theta f'(\theta)^2 \ s.t. \ E\left(\left|\hat{\beta}(Z) - \theta\right|^2\right) \geq \eta \tag{2.48}$$

where the additional constraints on the parameter domain and the encoding function are as stated at the end of Section 2.3. Since the eavesdropper employs the LMMSE estimator, the MSE at the eavesdropper, that is, $E\left(\left|\hat{\beta}(Z) - \theta\right|^2\right)$ is as

[4]The solution set for (2.46) also contains the set of all one-to-one functions on $[a, b]$ with $f(\theta) \in [a, b]$ and with finitely many discontinuous points, where between any two consecutive discontinuities, $|f'(\theta)| = 1$. Hence, there exist infinitely many encoding functions that solve (2.46). The encoding functions in (2.47) correspond to the optimal continuous solutions.

given in (2.26).

From (2.48), it is noted that the optimal encoding function should satisfy the MSE constraint by making the smallest slope in $[a, b]$ as large as possible. It is known that when the secrecy constraint is not effective (or, removed), the linear encoding function is optimal according to Proposition 6, and $|f'_{opt}(\theta)| = 1$. Therefore, for a given target level η in (2.48), one strategy to find the optimal encoding function is to search among eligible encoding functions that satisfy $\min_{\theta\in[a,b]} |f'(\theta)| = k$ and to check if any of them satisfies the target secrecy level, where k is set to 1 initially. If there exist no solutions for a given k, then k is decreased and the procedure is repeated, until a feasible function satisfying the secrecy constraint is found. Let $\mathcal{F}^k$ denote the family of one-to-one and continuous (except at a finite number of points) functions with the domain and codomain being given by $[a, b]$, and $\min_\theta |f'(\theta)| = k$. Then, a sufficient condition for optimality of $f \in \mathcal{F}^k$ is that it should satisfy the secrecy constraint and there should be no elements in $\mathcal{F}^m$ that satisfy the secrecy constraint for $m > k$. To determine whether the secrecy constraint can be satisfied for a given k, the highest MMSE at the eavesdropper has to be calculated for that specific value of k. Hence, the solution of the following optimization problem should be performed in the first step:

$$\hat{f}_{opt} = \arg\max_{\hat{f}} E\left(\left|\hat{\beta}(Z) - \theta\right|^2\right) \; s.t. \; k \leq \left|\hat{f}'(\theta)\right|, \forall\theta \in [a, b] \tag{2.49}$$

where $0 \leq k \leq 1$ is a given parameter.

Remark 6: The domain of the parameter is taken to be $\Lambda = [a, b]$ in the general case. However, due to Proposition 2, it can be assumed that $\Lambda = [0, \gamma]$ and $\hat{f}(\theta) : [0, \gamma] \to [0, \gamma]$, where $\gamma = b - a$, without loss of generality. Hence, in the rest of the manuscript, θ is assumed to be distributed uniformly in $[0, \gamma]$.

The following result characterizes the solution of (2.49).

Proposition 7: *For a given k, the form of the solution of* (2.49) *is given by*

$$\hat{f}_{opt}(\theta) = \begin{cases} \gamma - \theta k, & \text{if } 0 \leq \theta \leq \alpha \\ \gamma k - \theta k, & \text{if } \alpha < \theta \leq \gamma \end{cases}. \tag{2.50}$$

Furthermore, if

$$2 - \frac{h^2\gamma^2}{12}(2k - k^2) \geq (k+1)(h^2 V_{min} + 1)(h^2 V_{max} + 1) \tag{2.51}$$

where h is the channel quality for the eavesdropper,

$$V_{min} = \frac{k^2\gamma^2}{12} \quad \textit{and} \quad V_{max} = \frac{k^2\gamma^2}{12} + \frac{(1-k)\gamma^2}{4}, \tag{2.52}$$

then, both $\alpha = 0$ and $\alpha = \gamma$ are optimal α values. Otherwise, $\alpha = \gamma/2$ is optimal.

Proof: The first step in the proof is to specify the characteristics of the encoding function that maximizes the LMMSE. Note that $f(\theta) = X$ results in a random variable with $V = \mathrm{Var}(X)$, $C = \mathrm{Cov}(X, \theta)$ and $\mu = E(X)$, and the value of $E(|\hat{\beta}(Z) - \theta|^2)$ depends on these values. Hence, the LMMSE value is to be maximized over the possible values of V, C, and μ. It is noted that the slope constraint induces limitations on the possible values of μ, V, and C. Let S^k denote the feasible set of μ, V, and C values in the presence of the constraint $k \leq |f'(\theta)|$. As parameter θ is distributed uniformly on the interval $[0, \gamma]$, $E(\theta) = \gamma/2$ and $\mathrm{Var}(\theta) = \gamma^2/12$. Then, the optimization problem in (2.49) can be expressed as

$$\max_{\mu,V,C} \ \frac{h^2 V(V - 2C)}{h^2 V + 1} + \left(\mu - \frac{\gamma}{2}\right)^2 + \frac{\gamma^2}{12}, \ \ (\mu, V, C) \in S^k \tag{2.53}$$

After some manipulation, the objective function in (2.53) can be stated as $\lambda(V)E(|X - \theta|^2) + (1 - \lambda(V))(\mu^2 - \gamma\mu + \gamma^2/3)$, where $\lambda(V) \triangleq h^2V/(h^2V + 1)$. Note that for a given μ, $E(|X - \theta|^2)$ can be maximized, which would yield an upper bound on the objective function. It can be found by inspection that when

the slope constraint is taken into account, $E(|X-\theta|^2)$ is maximized for

$$\hat{X}^{\alpha} = \begin{cases} \gamma - \theta k, & \text{if } 0 \le \theta \le \alpha \\ \gamma k - \theta k, & \text{if } \alpha < \theta \le \gamma \end{cases} \tag{2.54}$$

where $(1-k)\alpha = \mu - k\gamma/2$ and $k\gamma/2 \le \mu \le \gamma - k\gamma/2$. Hence, the following the relationship is obtained:

$$\begin{aligned} E\left(\left|\hat{\beta}(Z)-\theta\right|^2\right) &\le \lambda(V)\beta_1(\alpha,k) + (1-\lambda(V))\beta_2(\alpha,k) \\ &= \lambda(V)(\beta_1(\alpha,k)-\beta_2(\alpha,k)) + \beta_2(\alpha,k) \end{aligned} \tag{2.55}$$

with $\beta_1(\alpha,k) \triangleq (k^2-1)(\alpha^2-\gamma\alpha)+(k^2-k+1)\gamma^2/3$ and $\beta_2(\alpha,k) \triangleq (k-1)^2(\alpha^2-\gamma\alpha)+(3k^2/4-3k/2+1)\gamma^2/3$. Now, notice that for a fixed k, the following equality holds:

$$\beta_1(\alpha,k)-\beta_2(\alpha,k) = (\alpha^2-\gamma\alpha)(2k-2) + \left(\frac{k^2}{4}+\frac{k}{2}\right)\frac{\gamma^2}{3}$$

Since $\beta_1(\alpha,k)$ is a concave function of α and $\beta_2(\alpha,k)$ is a convex function of α for $0 \le k \le 1$, $\beta_1(\alpha,k)-\beta_2(\alpha,k)$ is a concave function of α; hence, it attains its minimum at $\alpha = 0$ and $\alpha = \gamma$. Therefore, the following inequality is obtained: $\beta_1(\alpha,k)-\beta_2(\alpha,k) \ge (k^2/4+k/2)\gamma^2/3 \ge 0$, which implies that for a given value of μ, the right-hand-side of (2.55) is an increasing function of $\lambda(V)$. Hence, a further upper bound can be obtained for (2.55) by using the same $\hat{X}^{\alpha}$ defined above since it maximizes the variance under the slope constraint. For this function, the variance is given by $V(\alpha,k) = (k-1)(\alpha^2-\alpha\gamma)+k^2\gamma^2/12$. It is noted that $\lambda(V(\alpha,k))$ and the resulting upper bound are functions of α for fixed k and h.

Hence, the upper bound can be maximized over α as follows:

$$\begin{aligned} & E\big(\big|\hat{\beta}(Z)-\theta\big|^2\big) \\ & \quad \le \lambda(V(\alpha,k))\beta_1(\alpha,k)+(1-\lambda(V(\alpha,k)))\beta_2(\alpha,k) \\ & \quad = \lambda(V(\alpha,k))(\beta_1(\alpha,k)-\beta_2(\alpha,k))+\beta_2(\alpha,k) \\ & \quad \triangleq g(\alpha,k) \le \max_{\alpha\in[0,\gamma]} g(\alpha,k) \end{aligned} \tag{2.56}$$

If $\hat{\alpha} = \arg\max_{\alpha\in[0,\gamma]} g(\alpha,k)$, then $E(|\hat{\beta}(Z)-\theta|^2)$ achieves this upper bound by employing $\hat{\alpha}$ at the encoding function. Therefore, the optimal encoding function is $\hat{X}^{\hat{\alpha}}$, where $\hat{\alpha} = \arg\max_{\alpha\in[0,\gamma]} g(\alpha,k)$.

To conclude the proof, $\hat{\alpha}$ should be characterized for given k and h. Overall, the optimization problem can be written as

$$\max_{\alpha\in[0,\gamma]} \frac{h^2V(\alpha,k)}{h^2V(\alpha,k)+1}(\beta_1(\alpha,k)-\beta_2(\alpha,k))+\beta_2(\alpha,k) \tag{2.57}$$

where $h,\gamma>0$ and $k\in[0,1]$. Instead of optimizing over α, the optimization can be performed over V based on a change of variables by noting that for $\alpha\in[0,\gamma]$, $V(\alpha,k)\in[V_{min},V_{max}]$, where $V_{min}=k^2\gamma^2/12$ and $V_{max}=k^2\gamma^2/12+(1-k)\gamma^2/4$. Then, (2.57) is rewritten as

$$\max_{V\in[V_{min},V_{max}]} z(V) = \frac{h^2(k+1)V^2+HV+F}{h^2V+1} \tag{2.58}$$

where $H=(h^2\gamma^2/12)(4-4k+3k^2-k^3)+k-1$ and $F=(\gamma^2/12)(4-6k+4k^2-k^3)$. Then, according to the Weierstrass theorem, the global maximum exists for (2.58), and the solution can be found by applying Fermat's rule. Namely, the optimal solution either satisfies $z'(V)=0$ or is at the boundary, i.e., $V=V_{min}$ or $V=V_{max}$. For $z'(V)=0$, $V^2+2V/h^2+d/h^4=0$, where $d=(H-Fh^2)/(k+1)$. Then, $\hat{V}=-h^{-2}+h^{-2}\sqrt{1-d}$ is a candidate solution. However, $\hat{V}$ should belong to $[V_{min},V_{max}]$. To guarantee this condition, $h^2V_{max}\ge\sqrt{1-d}-1\ge h^2V_{min}$ should be satisfied. Therefore, $h^2V_{min}+1\le\sqrt{1-d}$. If this holds, then

$sgn(\lim_{V \to V_{min}^{+}} z'(V) = sgn(V_{min}^2 + 2h^{-2}V_{min} + h^{-4}d) \leq 0$. In conclusion, it is possible that a candidate solution is inside the feasible interval $[V_{min}, V_{max}]$; however, there is only one such solution and V is decreasing at the beginning of the interval. Due to continuity, it is noted that if $\hat{V} \in (V_{min}, V_{max})$, then it is in fact the global minimum. Hence, it is concluded that the solution of (2.58) is either V_{min} or V_{max}, excluding the possibility of the other case. Finally, the regions in which a certain end point is optimal are characterized. The condition of $z(V_{min}) \geq z(V_{max})$ occurs if h and k satisfy

$$2 - \frac{h^2\gamma^2}{12}(2k - k^2) \geq (k+1)(h^2V_{min} + 1)(h^2V_{max} + 1)$$

and $z(V_{min}) < z(V_{max})$ holds otherwise. Note that if the optimal solution is V_{max}, then $\hat{\alpha} = \gamma/2$. If the optimal solution is V_{min}, both $\hat{\alpha} = 0$ and $\hat{\alpha} = \gamma$ are the optimal solutions. ■

As the form of the optimal encoding function that maximizes the LMMSE at the eavesdropper is derived for any value of the minimum slope constraint (k) via Proposition 7, the optimal encoding function based on the worst-case Fisher information metric can be obtained by finding the maximum of such constraints. Hence, the problem reduces to the determination of the best (maximum) value of $k \in (0, 1]$ such that $\exists f \in \mathcal{F}^k$ in the form specified by (2.50) that satisfies the secrecy constraint. This approach can be implemented by using the procedure shown in Algorithm 1. It is noted that $E(|\hat{\beta}(\hat{X}^{\alpha}) - \theta|^2$ in Algorithm 1 can be calculated explicitly via (2.26) and (2.50).

2.4 Numerical Results

In this section, numerical examples are provided for both ECRB and worst-case Fisher information based encoder designs. Before providing the results, we first provide a motivation about the operational significance of ECRB metric to design optimal encoding functions.

Algorithm 1: $f_{opt} = \text{ENCODER}(\eta)$

```
% Δ is the decrement of slope at each iteration.
k ← 1
while k > 0 do
    Pick α = 0 or α = γ, if (2.51) holds. Else, α = γ/2.
    X̂^α = f̂_opt(θ) as given in (2.50)
    MSE ← E(|β̂(X̂^α) − θ|^2
    if MSE ≥ η then
        f_opt = X̂^α
        break
    else
        k ← k − Δ
    end
end
if k < 0 then
    Problem is infeasible
else
    return f_opt
end
```

2.4.1 Operational Significance of ECRB

Even though the ECRB is defined as the average of the CRB with a certain intuition, there is actually an operational significance of the ECRB and there exists a relationship between the MSE of practical Bayesian estimators and the ECRB; namely, minimizing the ECRB can actually correspond to performance optimization of practical estimators. More specifically, the MSE of the MAP estimator in the asymptotic region is actually the expectation of the conditional CRB [48]. Here, the asymptotic region, which is also known as the small error region, can refer either to the high SNR region or to a large number of independent observations. In this chapter, the asymptotic region corresponds to the high SNR region, i.e., high values of $(h_r/\sigma_r)^2$. For the system model we use, in the high SNR region, the MSE of a MAP estimator will be approximately same as the MSE of MMSE estimator. Therefore, optimization based on the ECRB metric provides optimal performance for practical Bayesian estimators such as MAP or MMSE estimator in the small error region. To elaborate this discussion, we focus on the characteristics of the MSE in general nonlinear parameter estimation problems [48]:

- Small error (asymptotic) region: The estimate will most likely be on the correct peak of log-likelihood function (or, the posterior distribution), leading to small errors.
- Threshold region: As the SNR decreases, after one point, some of the estimates will be close to the correct peak but others will be randomly distributed due to noise.
- Prior (data irrelevant) region: The peaks happen randomly due to noise, hence the observation carries no useful information about the parameter.

The bounds such as the ECRB are useful especially in the small error region, as the observations carry much more information in that region.Therefore, while searching for an optimal encoding function, we would like to optimize the performance of an estimator in the asymptotic region (high h_r/σ_r), because in the prior region (low h_r/σ_r), the observations do not provide enough information relevant to the true value of the parameter; hence, the estimators are mostly based on prior information in that region. Therefore, in order to observe and minimize the penalty incurred to intended receiver's estimation performance due to the encoding operation, we need a metric that closely reflects the MSE of practical Bayesian estimators in the asymptotic region and the ECRB satisfies this requirement.

In order to illustrate these points, we provide an example in Fig. 2.2, where we consider a prior distribution $w(\theta) = 2\theta$ for $\theta \in [0, 1]$. In the case of no secrecy constraints, an optimal encoding function based on the ECRB metric for the given prior distribution is $f_{opt}(\theta) = \theta^{4/3}$ based on Proposition 1. We provide the MSE performance of the MAP and MMSE estimators for the optimal encoding function, together with the resulting ECRB, which can be calculated as $\frac{27}{32}(\sigma_r/h_r)^2$, versus $(h_r/\sigma_r)^2$. It is noted that $ECRB = 10^{(\log_{10}(\frac{27}{32}) - 2\log_{10}(h_r/\sigma_r))}$. Therefore, the ECRB versus ($(h_r/\sigma_r)^2$ in dB) appears linear when we use the log-scale for the ECRB as well, as can be seen in Fig. 2.2. We also provide the MSE performance of the MAP and MMSE estimators for a non-optimal encoding function $f(\theta) = 0.2\theta$ for comparison purposes, and the resulting ECRB, which can be calculated as $25(\sigma_r/h_r)^2$. There are three important observations from Fig. 2.2.

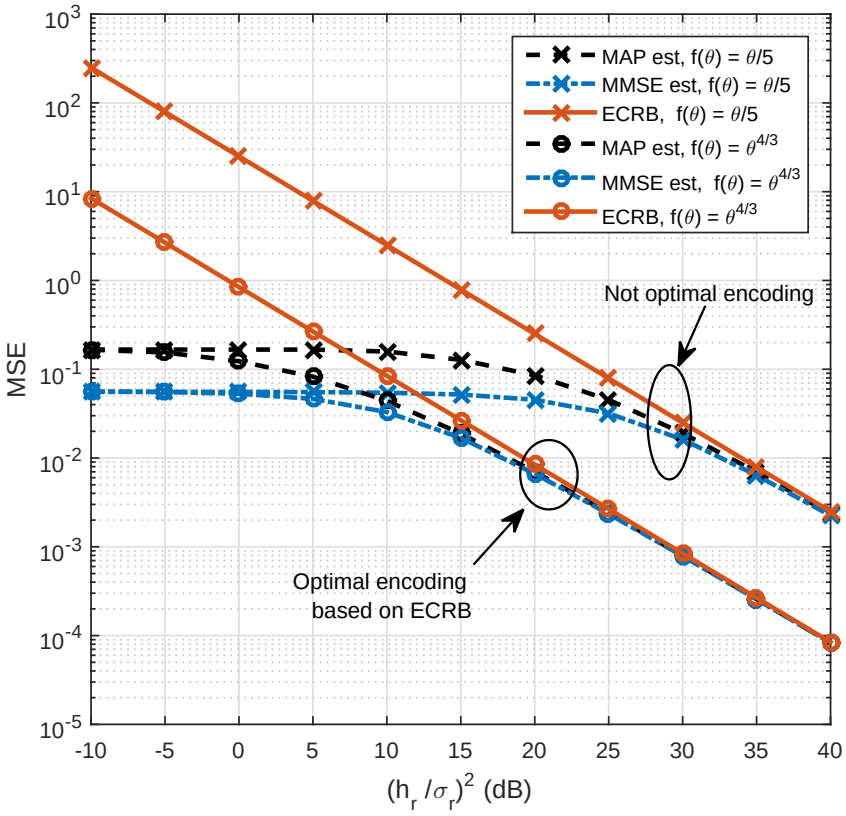

Figure 2.2: MSE versus $(h_r/\sigma_r)^2$ for MMSE, MAP estimators and ECRB when an optimal and non-optimal encoding functions are used for $w(\theta) = 2\theta$ for $\theta \in [0, 1]$. Note that $h_r = 1$.

- First, the minimization of the ECRB corresponds to shifting the linear line down in Fig. 2.2 while keeping the slope the same. We note the difference in the ECRBs corresponding to the optimal and the non-optimal encoding functions.

- Second, the MSE of the MAP and MMSE estimators converge to the ECRB as (h_r/σ_r) increases. Therefore, optimizing the ECRB effectively corresponds to optimizing the MSE of the MAP and MMSE estimators as h_r/σ_r increases.

- Third, the ECRB is not tight in the low h_r/σ_r region. However, it is noted that in the low h_r/σ_r region, the prior information mainly determines the estimator and judging the performance of different encoding functions in that region is less reliable due to noise. For example, $f(\theta) = 0.2\theta$ and $f(\theta) = \theta^{4/3}$ have similar MSE performance for the MMSE and MAP estimators when $(h_r/\sigma_r)^2 < 0\,\text{dB}$. However, as h_r/σ_r increases, a significant performance gap occurs between the two encoders. Therefore, the ECRB being loose at low h_r/σ_r's is not critical in terms of encoder design.

Finally, we should emphasize that the resulting ECRB of the optimal encoding function should not be taken as equal to the MSE of the MAP or MMSE estimator directly unless h_r/σ_r is large enough[5] due to the third item above. The performance of a specific estimator for a given encoding function in the low h_r/σ_r region can be calculated by using the expressions specific to that estimator.

2.4.2 Results for ECRB Based Design

In this part, numerical examples are provided to investigate the theoretical results in Section 2.2.1 and to compare the proposed approaches in Section 2.2.2. Throughout the simulations, h_r and σ_r^2 are set as $h_r = \sigma_r^2 = 1$.

[5]We note that the definition of the high $(h_r/\sigma_r)^2$ region depends on the employed encoding function. An approximation to indicate the start of the asymptotic region may be taken as $20 - 10\log_{10} E(f(\theta)^2)\,\text{dB}$.

First, we consider a scenario in which the channel parameters for the receiver and the eavesdropper are fixed, and investigate the relation between the ECRB and the secrecy limit α by using different encoding strategies. It is assumed that the parameter θ has uniform distribution over $[0, 1]$ and $h = h_e/\sigma_e = 1$. Also, the eavesdropper employs the linear MMSE estimator for the encoded parameter $\beta = f(\theta)$. The theoretical results derived in Section 2.2.1 can be applied for this example. In particular, based on Proposition 1, it is known that if there is no secrecy constraint, either $f(\theta) = \theta$ or $f(\theta) = 1-\theta$ is an optimal encoding function. Also, Proposition 3 states that the optimal encoding function can be searched among monotonically decreasing functions as the uniform distribution satisfies the symmetry condition. In addition, Corollary 2 reveals that if $\alpha \leq 4/39 = 0.1026$, then $f(\theta) = 1 - \theta$ is the optimal encoding function since such a secrecy level can be guaranteed by using $f(\theta) = 1 - \theta$. Furthermore, Proposition 4 claims that it is not possible to achieve a secrecy limit α higher than $1/3$ as $\gamma = 1 < 2/h = 2$ in this scenario.

For obtaining the encoding function based on the proposed approaches in Section 2.2.2, the linear and power encoding functions, and the polynomial and piecewise linear (PWL) approximations are considered. For the linear encoding, $f(\theta) = 1 - m\theta$ is used due to Proposition 5. Then, (2.39) provides a simple tool for the solution. For the power encoding function, $f(\theta) = p(1 - \theta)^q$ is employed, and the optimal p and q values are found for a given target α value (see Remark 4). For the polynomial approximation (with a degree of $K = 10$) and the piecewise linear approximation (with $M = 100$ intervals), the formulations in (2.41) and (2.42) are utilized, respectively.

In Fig. 2.3, the relation between the target level α and the optimal ECRB value can be observed. When $\alpha = 0.10$, it is noted that the optimal ECRB is 1, which can be achieved with $f(\theta) = 1 - \theta$. As α increases, the optimal ECRB increases exponentially. For example, when $\alpha = 0.25$, the optimal ECRB is found to be 25.06 and it becomes 1182.3 when $\alpha = 0.32$ for the piecewise linear approximation. Hence, the ECRB goes to infinity as α goes to the theoretical bound of $1/3$, as

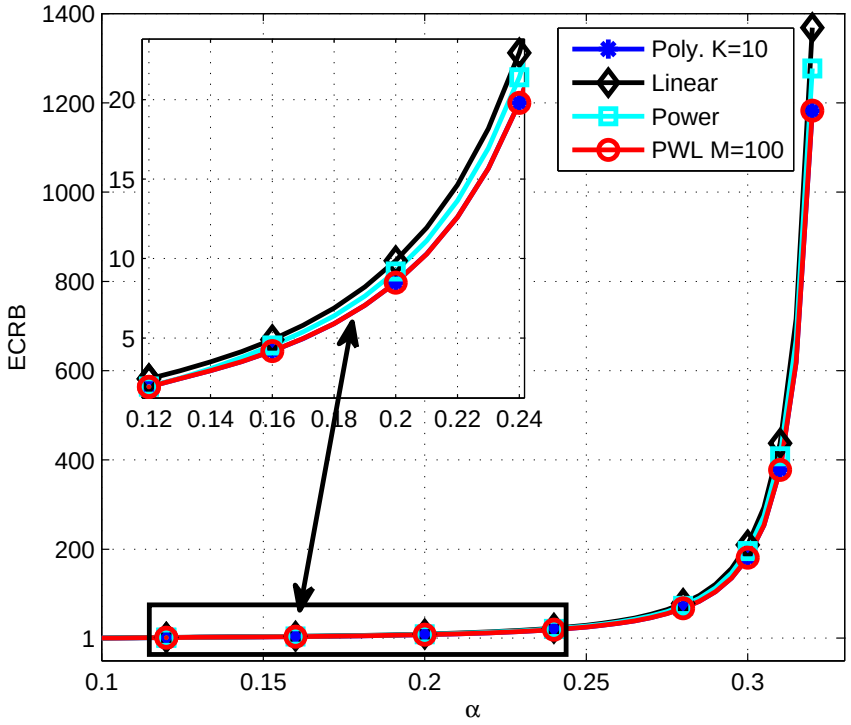

Figure 2.3: ECRB versus α for various solution approaches, where $h = 1$ and $0.1 \leq \alpha \leq 0.32$.

expected.[6] In Fig. 2.4, the encoding functions corresponding to the proposed solution approaches are presented for various values of α. It is observed that the polynomial approximation and the piecewise linear approximation yield almost the same function, which can also be deduced from the performance graph in Fig. 2.3. It is also seen that when $\alpha = 0.1$, all the methods lead to $f(\theta) = 1 - \theta$. When $\alpha = 0.2$, the difference between the solutions of the linear encoding and the approximation methods becomes noticeable. Since the approximation methods can use higher degrees of freedom than the linear encoding, they can achieve lower ECRBs. However, the linear encoding provides a simple solution for this scenario. For example, when $\alpha = 0.2$, the optimal linear encoding function can be obtained by finding the largest $m \in (0, 1]$ that satisfies $m^4 - m^3 + 9.4m - 18m + 4.8 \geq 0$, yielding $m = 0.3184$ due to (2.40); hence, $f(\theta) = 1 - 0.3184\,\theta$. It is also observed that the performance of the optimal power encoding approach in terms of the ECRB and the computational complexity is in between those of the optimal linear encoding and the other two approaches.

Next, the effects of the channel quality h of the eavesdropper on the optimal ECRB and encoding function are investigated for a given value of α. For this purpose, $\alpha = 0.15$ is used and the ECRB performance is evaluated versus $h = h_e/\sigma_e$ in Fig. 2.5. As discussed before, as h increases, the distortion due to encoding is transmitted to the eavesdropper more effectively and the intended MSE can be generated with a lower ECRB. Some interesting observations can be made in Fig. 2.5. First, three different regions are noted for the ECRB. In the first region, the ECRB slowly decreases as h increases for all the solution approaches. In the second region, for the power and the approximation approaches, the ECRB decreases more rapidly and finally when h is above some threshold value, $f(\theta) = 1 - \theta$ becomes sufficient to generate the MSE value of $\alpha = 0.15$ at the eavesdropper. Actually, this threshold can be calculated analytically based on Corollary 2. For the parameters in the considered scenario, $V_u = 1/12$, $E(\theta) = 1/2$, and $\alpha = 0.15$; hence, $h_{th} = \sqrt{48/11} = 2.09$. It is observed that the performance of the polynomial approximation is very similar to that of the piecewise linear approximation;

[6] In this example, the optimal ECRB value should not directly be taken as equal to the MSE at the estimator of the intended receiver since h_r/σ_r is not sufficiently high. Here, the ECRB is merely used as an objective function to represent generic estimation accuracy.

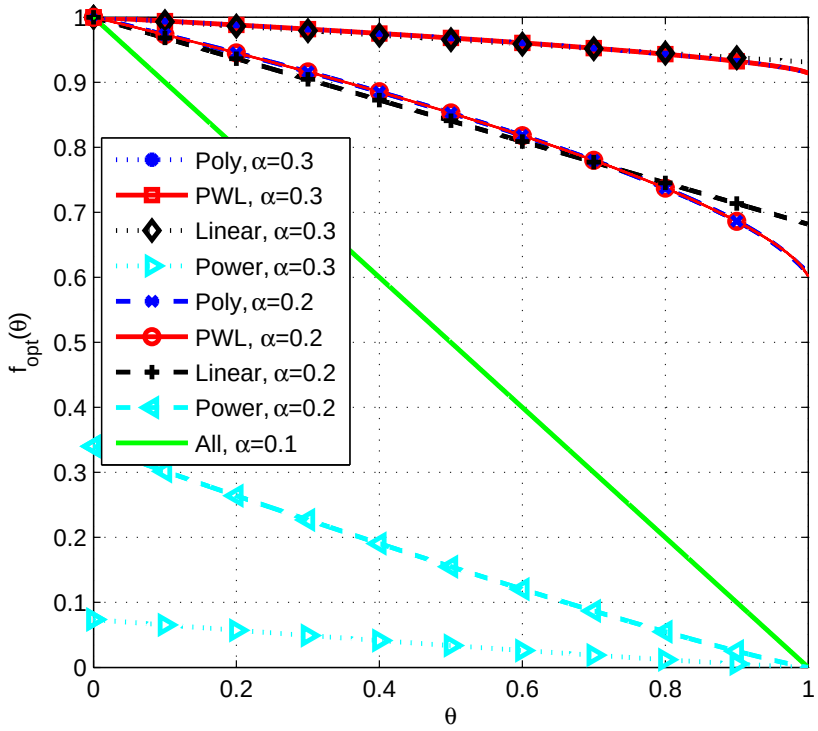

Figure 2.4: $f_{opt}(\theta)$ versus θ for various solution approaches, where $\alpha = 0.1$, 0.2, and 0.3.

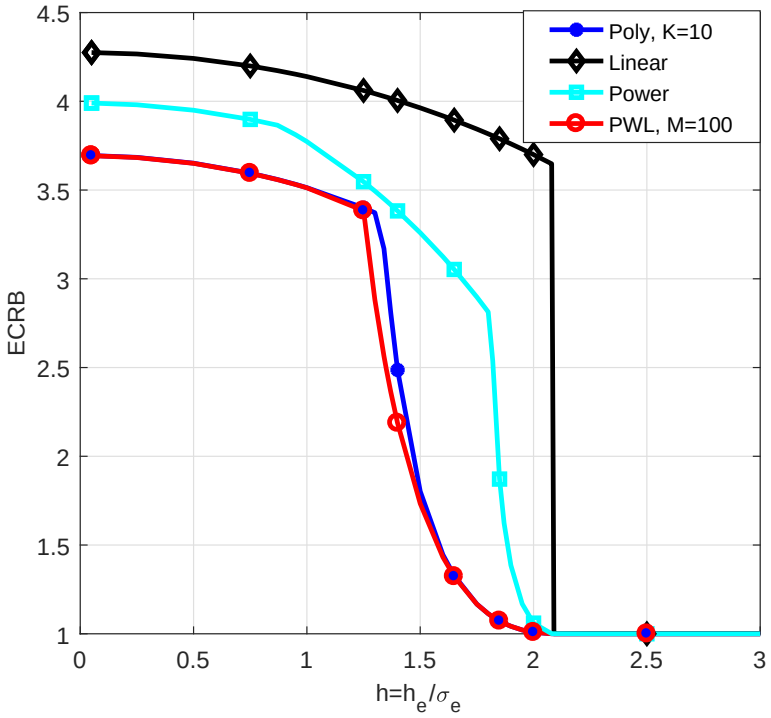

Figure 2.5: ECRB versus h for various solution approaches when $\alpha = 0.15$ with uniform prior distribution.

however, in the second region, it is slightly worse than that of the piecewise linear approximation. The optimal encoding function corresponding to the piecewise linear approximation approach is presented in Fig. 2.6, which reveals that the encoding function changes characteristics as h increases. This also explains why the polynomial approximation is slightly worse than the piecewise linear approximation for medium values of h. Namely, for the polynomial approximation, it is harder to correctly implement the sudden decrease around $\theta = 0.5$ while it has sufficient degrees of freedom to produce an encoding function required for smaller h values as it can also be observed in Fig. 2.4. It is also noted that the encoding function is still continuous; that is, it has a finite but large derivative around $\theta = 0.5$. In addition, it is seen that when $h = 2$, the encoding function is almost linear.

Next, a scenario with a nonuniform prior distribution is considered, and the prior PDF of parameter θ is modeled as $w(\theta) = 2\theta$ for $\theta \in [0, 1]$. Similar to the

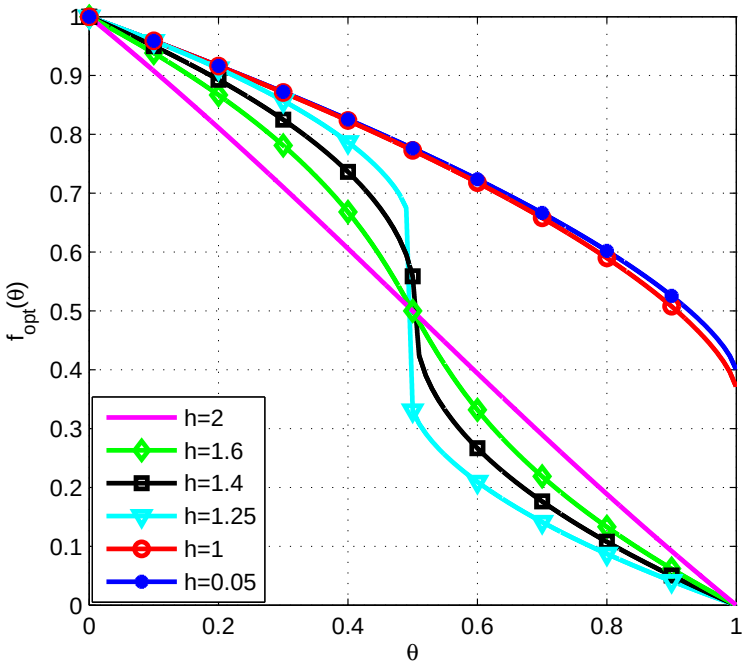

Figure 2.6: $f_{opt}(\theta)$ versus θ for the piecewise linear approximation when $\alpha = 0.15$ with uniform prior distribution.

uniform distribution case, the characteristics of the optimal encoding function are investigated for the fixed α and fixed h cases. First, it is assumed that $h = 1$ and the optimal encoding function is presented for various α values in Fig. 2.7 by using the piecewise linear approximation approach. The theoretical optimal solution $f(\theta) = 1 - \theta^{4/3}$ for the no constraint case is also shown in the figure, which is calculated based on Proposition 1 for the given prior distribution. It is observed that when the target level is small; i.e., $\alpha = 0.1$, the optimal encoding function calculated via the piecewise linear approximation is exactly the same as the theoretical solution. As α increases, in order to satisfy the target secrecy level, the optimal encoding function maps θ to lower values. It is noted that higher target α levels are achievable for this prior distribution as compared to the uniform distribution when $h = 1$. In particular, the secrecy limit is $1/2$ instead of $1/3$ in this example. Then, α is fixed as $\alpha = 0.34$, and the ECRB performance is investigated with respect to h in Fig. 2.8. It is noted that the performance trends of the different solution approaches are similar to those in the uniform case presented in Fig. 2.5; however, unlike the uniform

distribution case, a sharp decrease to the minimum ECRB does not exist in this scenario (see Fig. 2.8). This is mainly due to the fact the optimal functions for the various function families yielding that minimum ECRB in the absence of an eavesdropper actually could not satisfy the secrecy requirement even if h gets large. For example, if the linear encoding with $m = 1$ is used, it can be shown that as $h \to \infty$, the resulting MSE is $1/3$, and if the theoretical solution for the no constraint case (that is, $f(\theta) = 1 - \theta^{4/3}$) is used, the resulting MSE is 0.318 as $h \to \infty$. It is known that the linear encoding with $m = 1$ would yield an ECRB value of 1 and $f(\theta) = 1 - \theta^{4/3}$ would yield an ECRB value of $27/32 = 0.844$. However, unlike the previous example, since the target α value is too high to achieve with those encoding functions, these minimum ECRB values cannot be attained even if h gets arbitrarily large; hence, a slow decay with a floor is observed in the ECRB instead of a sudden decrease. The ECRB floor values are found to be 1.5625, 1.0482, and 0.8835 for the linear encoding, the power encoding, and the piecewise linear approximation, respectively. Also, it is noted that the performance differences between the different solution approaches are small when h is low, which become more significant for medium values of h. Finally, the optimal solutions via the piecewise linear approximation are provided for various h values in Fig. 2.9. It is noted that the characteristics of the optimal encoding function are different for small, medium, and large values of h. One interesting observation is that for medium values of h, it is seen that the sudden decrease in the optimal encoding function does not necessarily happen at 0.5 unlike the uniform prior distribution case.

Finally, we provide the simulation times for obtaining the solutions of the various methods and the resulting ECRB values in Table 2.1 for the scenario considered in Fig. 2 with $\alpha = 0.15$.[7] We observe that the linear and power encoding approaches have shorter solution times while they provide suboptimal solutions. For the polynomial and piecewise linear approximations, as K and M increase, the simulation times increase and lower ECRB values can be obtained. However, it is observed that after a certain value, the improvement in the ECRB is not significant. Therefore, it makes sense to choose the values of these parameters

[7]The simulations are performed with Intel Core i5-4590 CPU 3.30 GHz processor and Matlab R2017B.

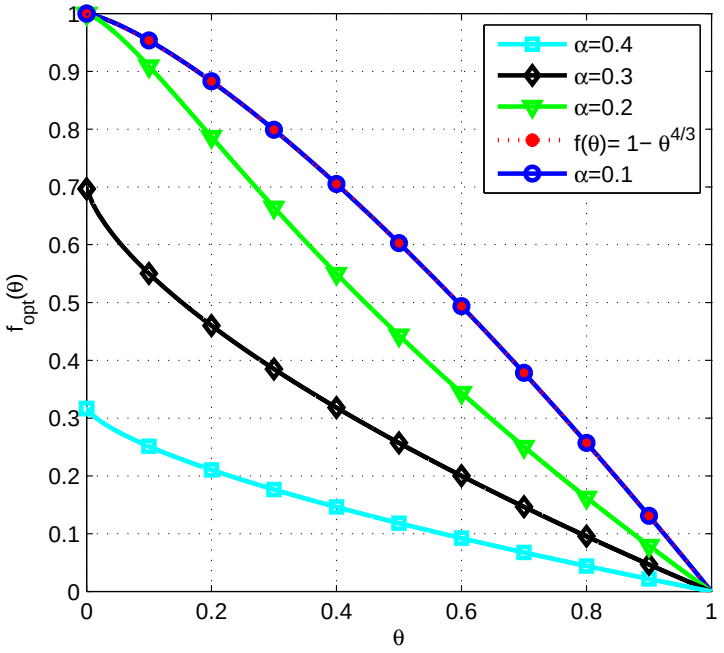

Figure 2.7: $f_{opt}(\theta)$ versus θ for piecewise linear approximation ($M = 100$), where $\alpha = 0.1$, 0.2, 0.3, and 0.4. $f(\theta) = 1-\theta^{4/3}$ is the optimal function under no secrecy constraints according to Proposition 1.

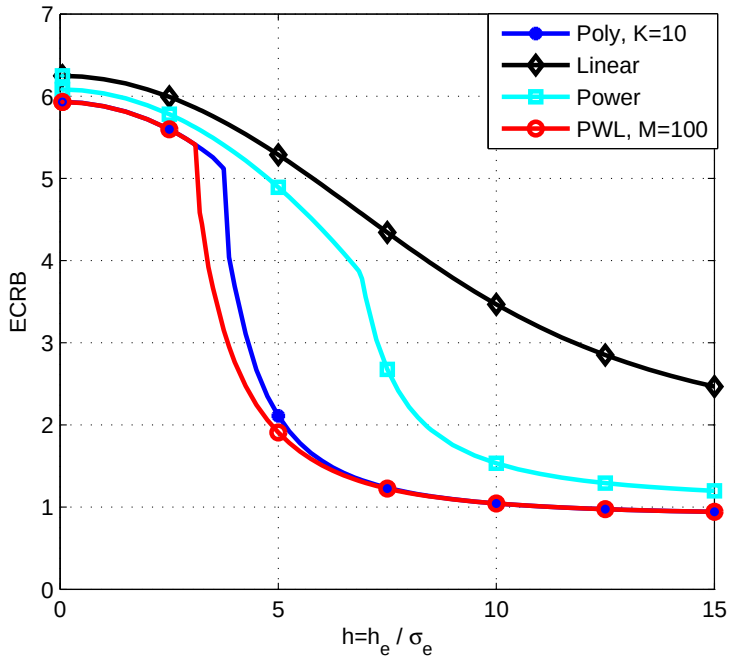

Figure 2.8: ECRB versus h for various solution approaches when $\alpha = 0.34$ for $w(\theta) = 2\theta$ for $\theta \in [0, 1]$.

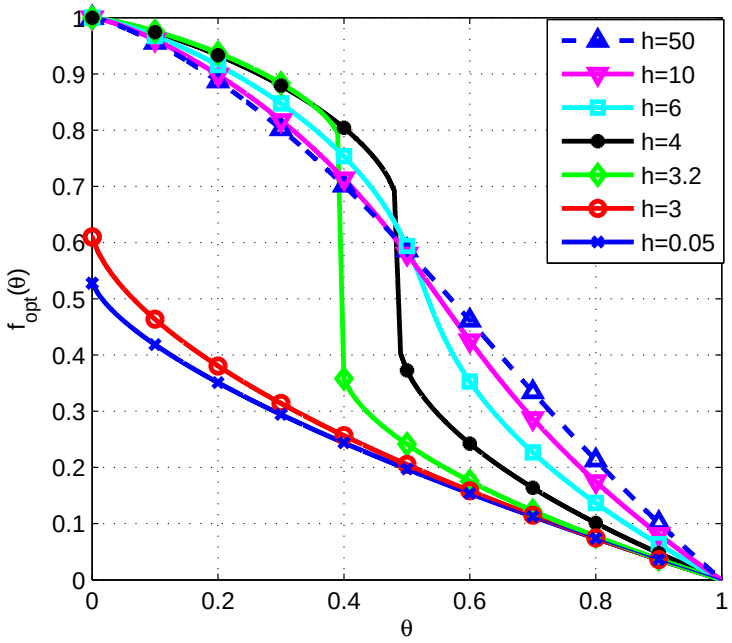

Figure 2.9: $f_{opt}(\theta)$ versus θ for piecewise linear approximation when $\alpha = 0.34$ with $w(\theta) = 2\theta$ for $\theta \in [0, 1]$.

considering the solution times as well. In this study, we have used $K = 10$ and $M = 100$.

Solution Method	**ECRB**	**Time (ms.)**
Linear Encoding	4.1395	0.35
Power Encoding	3.7730	35
Poly. App. ($K = 2$)	3.6170	33
Poly. App. ($K = 4$)	3.5263	142
Poly. App. ($K = 6$)	3.5163	763
Poly. App. ($K = 8$)	3.5139	4680
Poly. App. ($K = 10$)	3.5135	5540
Poly. App. ($K = 14$)	3.5129	18102
PWL App ($M = 5$)	3.5634	159
PWL App ($M = 10$)	3.5289	302
PWL App ($M = 25$)	3.5159	750
PWL App ($M = 50$)	3.5134	1483
PWL App ($M = 100$)	3.5125	6220
PWL App ($M = 200$)	3.5123	23687

Table 2.1: ECRB values and simulation times for various approaches, where $\alpha = 0.15$.

.

2.4.3 Results for Worst-Case Fisher Information Based Design

In this part, a numerical example is provided based on the theoretical results and the proposed algorithm in Section 2.3.1. The channel parameters are selected as $h_r = \sigma_r = 1$ for the intended receiver and $h = 0.5$ and $h = 1.5$ for the eavesdropper. The parameter θ is assumed to be uniformly distributed in the interval of $[0, 2]$; i.e., $\gamma = 2$. The eavesdropper employs the LMMSE estimator by using the observations based on the encoded parameter $\beta = f(\theta)$. Also, Δ is set to 0.001 in the proposed algorithm for calculating the optimal encoding functions. In Fig. 2.10, the worst-case Fisher information values achieved by the proposed algorithm are presented with respect to the target secrecy level for $h = 0.5$ and $h = 1.5$. For comparison purposes, the worst-case Fisher information values corresponding to the ECRB based encoding algorithm in Section 2.2 are

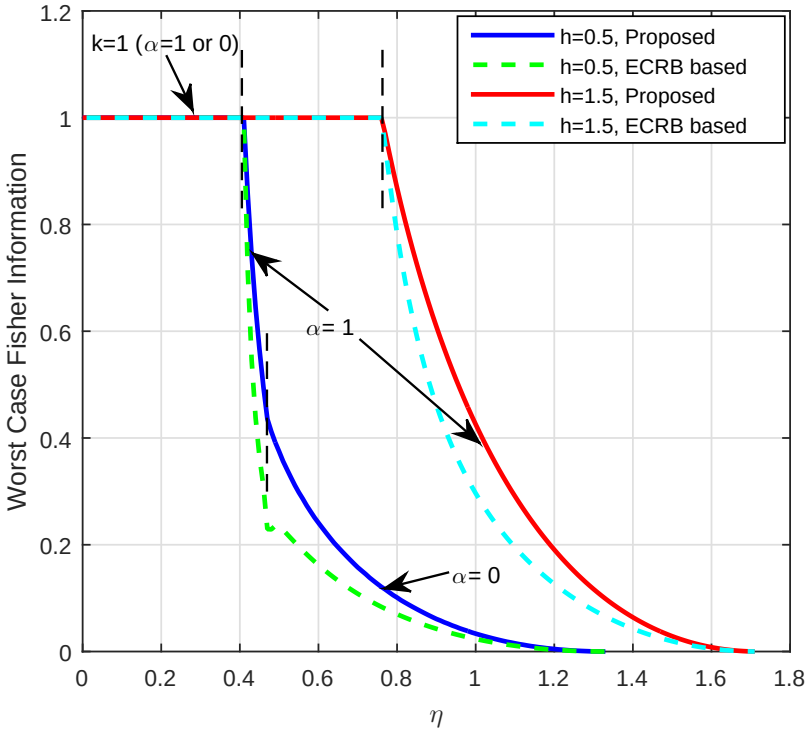

Figure 2.10: Worst-case Fisher information versus η.

also provided in the same figure. (The proposed scheme provides higher worst-case Fisher information than the ECRB based scheme since the latter aims to optimize the average CRB.) In Fig. 2.11, the optimal encoding functions based on the worst-case Fisher information metric are provided for various η values for $h = 0.5$. As justified in Proposition 7, the optimal encoding function is either linear with a certain slope between 0 and 1, or piecewise linear with a single discontinuity at $\theta = \gamma/2$ depending on the target secrecy level η.

In Fig. 2.10, it is observed that as the target secrecy level increases, the worst-case Fisher information achieved by the proposed algorithm decreases, as expected. In addition, it is possible to obtain higher worst-case Fisher information values when $h = 1.5$ for the same MSE target compared to the case of $h = 0.5$ since the distortion due to the encoding is transmitted to the eavesdropper more effectively under better channel conditions. Note that when $h = 0.5$, the three different regions are observable in the performance figure.

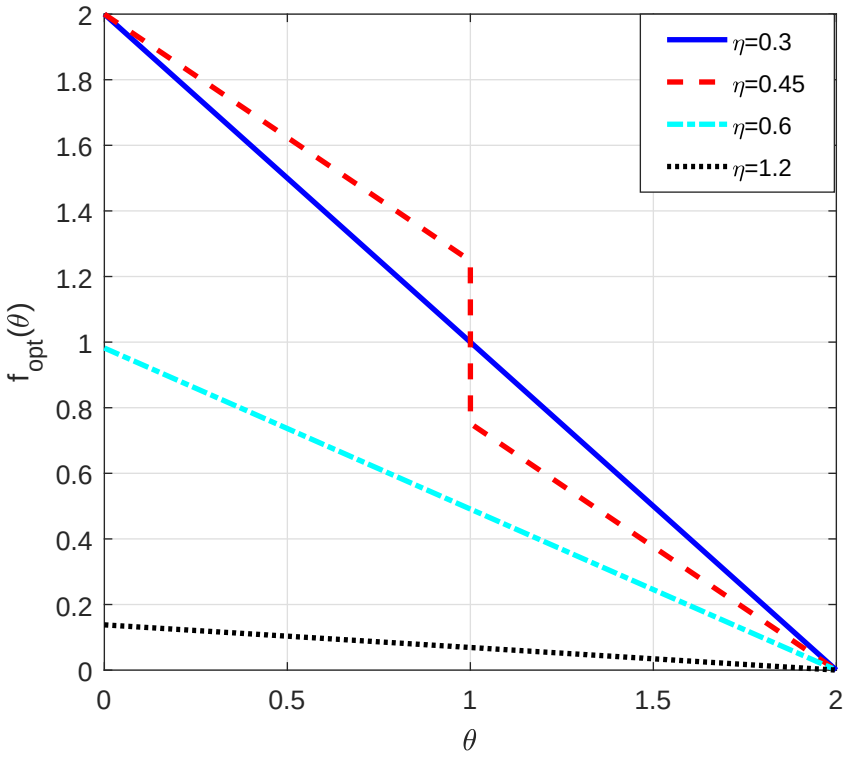

Figure 2.11: $f_{opt}(\theta)$ versus θ for $h = 0.5$.

When $\eta \leq \eta_1 = 16/39 = 0.4101$, employing $k = 1$, that is, $f_{opt}(\theta) = \gamma - \theta$, is sufficient to attain the target secrecy levels. In general, η_1 can be found as $\eta_1 = 0.25\gamma^2 \left(h^2\gamma^2/(h^2\gamma^2 + 12) + 1/3\right)$. When $\eta_1 < \eta \leq \eta_2$ with $\eta_2 = 0.4708$, it is observed that the optimal α value becomes $\gamma/2$. It is noted that η_2 can be found by determining the point at which (2.51) becomes an equality in general. Therefore, in this region, the optimal encoding function has a single discontinuity at $\theta = \gamma/2$. Finally, when $\eta_2 < \eta \leq 4/3$, the optimal α is 0; hence, the optimal encoding function is linear with no discontinuities. It is interesting to note that the worst-case Fisher information decreases faster in the second region, and it decays to zero in the third region more slowly as compared to the second region. On the other hand, when $h = 1.5$, only two of such regions are observed in Fig. 2.10.

2.5 Concluding Remarks

The optimal parameter encoding problem has been studied in the presence of an eavesdropper. An optimization problem is formulated to minimize the ECRB at the intended receiver under the constraint of a target MSE value at the eavesdropper. A closed-form expression has been derived for the optimal encoding function when there is no secrecy constraint. When a certain secrecy level is to be guaranteed at the eavesdropper, first a sufficient condition has been provided for the case in which the optimal encoding function under no secrecy constraints is still optimal. Next, a closed-form expression for the MSE of the eavesdropper has been derived under the assumption that the eavesdropper employs the linear MMSE estimation. Based on this result, the shift invariance property has been shown for generic prior PDFs, and it has been proved that it is sufficient to restrict the search to decreasing encoding functions if the prior distribution of the parameter has a certain symmetry property. In addition, an upper limit has been obtained for the MSE of the eavesdropper for the uniform prior distribution. This result implies that the optimal encoding function either maximizes or minimizes the variance of the encoded parameter depending on the channel quality parameter and the length of the range interval for the encoding function. In order to calculate the optimal encoding function numerically, various solution approaches

have been considered; namely, linear encoding, polynomial approximation, and piecewise linear approximation. It has also been shown that the optimal solution for the linear encoding function can be obtained algebraically.

Furthermore, optimal parameter encoding problem has been studied based on worst-case Fisher information to develop a robust encoding approach. An optimization problem is formulated to maximize the minimum Fisher information at the intended receiver while keeping the MSE at the eavedropper above a certain target. The solution set for the optimal encoding functions under no secrecy constraints is obtained analytically. Also, a closed-form expression for optimal encoding function which maximizes the MSE at the eavesdropper for a given level of minimum Fisher information at the intended receiver. This solution has been utilized to develop a low-complexity solution algorithm to obtain the solution for original problem with secrecy constraint.

As future work, an interesting extension would be to formulate the problem in a game theoretic framework, where the eavesdropper has some partial information about transmitter's strategy and the transmitter considers this possibility in the design of the encoding functions in both scenarios. Also, the extension of the analytical results to the cases in which the eavesdropper employs the MMSE estimator can be investigated, and the performance comparisons to the current model can be performed.

2.6 Appendices

2.6.1 Derivation of (2.24) and (2.25)

The linear MMSE estimator $\hat{\beta}(Z)$ to estimate β based on Z can be expressed as [64]

$$\hat{\beta}(Z) = E(\beta) + \frac{Cov(\beta, Z)}{Var(Z)}(Z - E(Z)) \tag{2.59}$$

From $Z = h_e\beta + N_e$, the following relations are obtained:

$$\begin{aligned}\hat{\beta}(Z) &= E(\beta) + \frac{Cov(\beta, h_e\beta + N_e)}{Var(h_e\beta + N_e)}(Z - h_eE(\beta)) \qquad (2.60)\\ &= E(\beta) + \frac{h_eVar(\beta)}{h_e^2Var(\beta) + \sigma_e^2}(Z - h_eE(\beta))\\ &= \frac{h_eVar(\beta)}{h_e^2Var(\beta) + \sigma_e^2}Z + \left(1 - h_e\frac{h_eVar(\beta)}{h_e^2Var(\beta) + \sigma_e^2}\right)E(\beta)\end{aligned}$$

where the second inequality is due to the independence of β and N_e.

2.6.2 Derivation of (2.26)

The eavesdropper is modeled to employ the linear MMSE estimator specified by $\hat{\beta}(z) = k_0 + k_1 z$, where k_1 and k_0 are given by (2.24) and (2.25), respectively. Defining $\beta = f(\theta)$, $V = Var(\beta)$, $C = Cov(\beta, \theta)$, and $h = h_e/\sigma_e$, the MSE at the eavesdropper can be written as

$$\begin{aligned}E\left(\left|\hat{\beta}(Z) - \theta\right|^2\right) &= E\left((k_1Z + k_0 - \theta)^2\right) \qquad (2.61)\\ &= E\left(k_1^2Z^2 + 2k_1k_0Z + k_0^2 + \theta^2 - 2(k_1Z + k_0)\theta\right) \qquad (2.62)\\ &= k_1^2E\left(h_e^2\beta^2 + 2h_e\beta N + N^2\right) + 2k_1k_0h_eE(\beta)\\ &+ k_0^2 + E(\theta^2) - 2k_1h_eE(\theta\beta) - 2k_0E(\theta) \qquad (2.63)\end{aligned}$$

where (2.63) follows from that facts that $Z = h_e\beta + N$ and $E(Z) = h_eE(\beta)$. In addition, it is known that $E(N^2) = \sigma_e^2$, and θ and N are independent random variables with $E(N) = 0$; hence, $E(\beta N) = 0$. Then, the expression in (2.63) is

further processed as follows:

$$
\begin{aligned}
& E\left(\left|\hat{\beta}(Z)-\theta\right|^2\right) \\
& = k_1^2 h_e^2 E(\beta^2) + k_1^2\sigma_e^2 + 2k_1 k_0 h_e E(\beta) \\
& + k_0^2 + E(\theta^2) - 2k_1 h_e E(\theta\beta) - 2k_0 E(\theta) && (2.64) \\
& = k_1^2 h_e^2 E(\beta^2) + k_1^2\sigma_e^2 + 2k_1(1-k_1 h_e) h_e E(\beta)^2 \\
& + E(\beta)^2(1 + k_1^2 h_e^2 - 2k_1 h) + E(\theta^2) \\
& - 2k_1 h_e E(\theta\beta) - 2(1-k_1 h_e) E(\beta) E(\theta) && (2.65) \\
& = k_1^2 h_e^2 (E(\beta^2) - E(\beta)^2) + k_1^2\sigma_e^2 + E(\theta^2) \\
& + E(\beta)^2 - 2k_1 h_e (E(\beta\theta) - E(\beta)E(\theta)) \\
& - 2E(\beta)E(\theta) = k_1^2 h_e^2 V + k_1^2\sigma_e^2 - 2k_1 h_e C \\
& + E(\theta^2) - E(\theta)^2 + E(\theta)^2 + E(\beta)^2 - 2E(\beta)E(\theta) && (2.66) \\
& = k_1^2(h_e^2 V + \sigma_e^2) - 2k_1 h_e C + Var(\theta) + (E(\beta) - E(\theta))^2 && (2.67) \\
& = \frac{h_e^2 V^2 - 2h_e^2 VC}{h_e^2 V + \sigma_e^2} + Var(\theta) + (E(\beta) - E(\theta))^2 && (2.68) \\
& = \frac{(h_e/\sigma_e)^2 V(V-2C)}{(h_e/\sigma_e)^2 V + 1} + Var(\theta) + (E(\beta) - E(\theta))^2 && (2.69) \\
& = \frac{h^2 V(V-2C)}{h^2 V + 1} + Var(\theta) + (E(\beta) - E(\theta))^2 && (2.70)
\end{aligned}
$$

where (2.64) follows directly from (2.63), (2.65) is obtained by inserting (2.25) into (2.64), (2.66) follows by rearranging the terms and adding and subtracting $E(\theta)^2$ in (2.65), (2.68) is obtained by inserting (2.24) into (2.67), and finally (2.70) is due to the use of $h = h_e/\sigma_e$ in (2.69). ■

Chapter 3

Estimation Theoretic Optimal Encoding Design for Secure Transmission of Multiple Parameters

In this chapter, optimal deterministic encoding of a vector parameter is investigated in the presence of an eavesdropper, and two practical solution strategies are developed based on nonlinear individual encoding and affine joint encoding of parameters [44]. The main contributions of this chapter can be summarized as follows:

- The optimal encoding of multiple parameters is proposed by utilizing the ECRB metric at the intended receiver and a MSE target at the eavesdropper. Two practical encoding strategies, nonlinear individual encoding and affine joint encoding, are introduced as possible encoding solutions.
- For nonlinear individual encoding, it is shown that the optimization problem can be decoupled into independent problems if the channel noise for the eavesdropper is white and parameters are independent. It is also proved

that if the prior distribution of a given parameter is symmetric on the domain, then the corresponding encoding function can be limited to decreasing functions.

- For affine joint encoding, the optimal encoding function is provided when there is no secrecy constraints and the channel noise for intended receiver is white.
- It is shown that the search for the optimal affine encoding strategy can be converted to a precoding matrix search; that is, the constant term can be eliminated from the optimization problem.

The rest of this chapter is organized as follows: The optimal encoding problem for multiple parameters is formulated in Section 3.1. The nonlinear individual encoding strategy and affine joint encoding strategies are studied in Sections 3.2 and 3.3, respectively. Numerical results are presented in Section 3.4 and concluding remarks are given in Section 3.5.

3.1 Problem Formulation

Consider a scenario in which N-dimensional random vector parameter $\boldsymbol{\theta} = [\theta_1\ \theta_2 \cdots \theta_N]^T \in \Lambda$ is to be transmitted to an intended receiver over N channels, and $w(\boldsymbol{\theta})$ denotes the joint probability density function (PDF) of $\boldsymbol{\theta}$. A block fading channel model is assumed such that the instantaneous fading coefficient at each channel is independent and denoted by constant $h_{r,i}$ for $i = 1, 2, \ldots, N$. As this model considers a slowly fading channel, it is assumed that the channel coefficients are constant during the transmission of the parameters. In addition to the transmitter and the intended receiver, there exists an eavesdropper that tries to estimate the parameter $\boldsymbol{\theta}$. The objective is to perform accurate estimation of the parameter at the intended receiver while keeping the estimation error at the eavesdropper above a certain level. Therefore, vector parameter $\boldsymbol{\theta}$ is encoded by using a vector-valued encoding function $\boldsymbol{f} : \Lambda \to \Gamma$ before the transmission of

the parameter. [1] Let $\boldsymbol{\beta} \in \Gamma$ be the encoded version of the parameter, which is defined as

$$\boldsymbol{\beta} \triangleq \boldsymbol{f}(\boldsymbol{\theta}) = \begin{bmatrix} f_1(\theta_1, \theta_2, \ldots, \theta_N) \\ f_2(\theta_1, \theta_2, \ldots, \theta_N) \\ \vdots \\ f_N(\theta_1, \theta_2, \ldots, \theta_N) \end{bmatrix}. \tag{3.1}$$

Then, the received signal at the intended receiver is expressed as

$$\boldsymbol{Y} = \boldsymbol{H}_r \boldsymbol{\beta} + \boldsymbol{N}_r \tag{3.2}$$

where $\boldsymbol{H}_r = \text{diag}\{h_{r,1}, h_{r,2}, \ldots, h_{r,N}\}$ is an $N \times N$ diagonal matrix of channel coefficients and $\boldsymbol{N}_r$ is the N-dimensional channel noise which is modeled as a zero-mean Gaussian random vector with covariance matrix $\boldsymbol{\Sigma}_r$ and is independent of $\boldsymbol{\theta}$. On the other hand, the eavesdropper observes

$$\boldsymbol{Z} = \boldsymbol{H}_e \boldsymbol{\beta} + \boldsymbol{N}_e \tag{3.3}$$

where $\boldsymbol{N}_e$ is zero-mean Gaussian noise with covariance matrix $\boldsymbol{\Sigma}_e$, which is also independent of $\boldsymbol{\theta}$, and $\boldsymbol{H}_e = \text{diag}\{h_{e,1}, h_{e,2}, \ldots, h_{e,N}\}$ is an $N \times N$ diagonal matrix representing the channel between the transmitter and the eavesdropper under a block fading channel model. The intended receiver tries to estimate parameter $\boldsymbol{\theta}$ based on observation $\boldsymbol{Y}$ whereas the eavesdropper employs observation $\boldsymbol{Z}$ for estimating $\boldsymbol{\theta}$, as illustrated in Fig. 3.1. Note that the eavesdropper is not aware of encoding; hence, it effectively tries to estimate $\boldsymbol{\beta}$.

In order to measure estimation accuracy at the intended receiver, the expectation of Cramer-Rao bound (ECRB) is employed similarly to Section 2.2. It is also assumed that the eavesdropper employs the LMMSE estimator $\hat{\boldsymbol{\beta}}(\boldsymbol{Z})$ whose coefficients are selected to estimate $\boldsymbol{\beta} = \boldsymbol{f}(\boldsymbol{\theta})$ based on $\boldsymbol{Z}$. The secrecy goal is achieved when the MSE at the eavesdropper for each θ_i is above a certain

[1] The encoder is designed for each transmission block and should be updated when the channel realization changes.

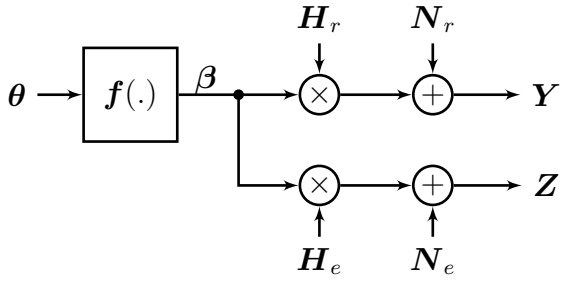

Figure 3.1: System model.

threshold. The ECRB for vector parameters can be expressed as [48]

$$E_{\boldsymbol{\theta}}\big(\mathbf{I}(\boldsymbol{\theta})^{-1}\big) = \int_{\Lambda} w(\boldsymbol{\theta})\ \mathbf{I}(\boldsymbol{\theta})^{-1} d\boldsymbol{\theta} = ECRB \tag{3.4}$$

where $\mathbf{I}(\boldsymbol{\theta})$ represents the Fisher information matrix (FIM), which is given by

$$I(\boldsymbol{\theta}) = E\left(\left(\frac{\partial p_{\boldsymbol{Y}|\boldsymbol{\theta}}(\mathbf{y}|\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right)\left(\frac{\partial p_{\boldsymbol{Y}|\boldsymbol{\theta}}(\mathbf{y}|\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right)^{T}\right) \tag{3.5}$$

with $p_{\boldsymbol{Y}|\boldsymbol{\theta}}(\mathbf{y}|\boldsymbol{\theta})$ representing the conditional PDF of $\boldsymbol{Y}$ for a given value of $\boldsymbol{\theta}$ [47]. Also, the error covariance matrix at the eavesdropper, who is unaware of the encoding, based on the estimate of the eavesdropper $\hat{\boldsymbol{\beta}}(\boldsymbol{Z})$ and the true value of the parameter $\boldsymbol{\theta}$ is defined as

$$\boldsymbol{\Sigma}_{err} = E\Big(\big(\hat{\boldsymbol{\beta}}(\boldsymbol{Z}) - \boldsymbol{\theta}\big)\big(\hat{\boldsymbol{\beta}}(\boldsymbol{Z}) - \boldsymbol{\theta}\big)^{T}\Big). \tag{3.6}$$

The expression in (3.4) is a matrix with each diagonal element representing the estimation accuracy limit for an individual parameter. Therefore, to determine the optimal encoding function for the overall vector parameter, the cost function is based on the sum of the diagonal elements of the inverse FIM, and the optimal parameter encoding problem is proposed as follows:

$$\begin{aligned} \boldsymbol{f}_{opt} = &\arg\min_{\boldsymbol{f}} \int_{\Lambda} w(\boldsymbol{\theta})\ tr\{\mathbf{I}(\boldsymbol{\theta})^{-1}\} d\boldsymbol{\theta} \\ &s.t. \quad \boldsymbol{\Sigma}_{err}(i) \geq \eta_i\,, \quad i = 1, 2, \ldots, N. \end{aligned} \tag{3.7}$$

where $tr\{\cdot\}$ denotes the trace operator, $\boldsymbol{\Sigma}_{err}(i)$ is the ith diagonal element of

$\boldsymbol{\Sigma}_{err}$, and η_i is the MSE target for θ_i at the eavesdropper.

It is important to emphasize that (3.7) involves optimization in the space of vector-valued functions with multiple inputs, hence it is difficult to solve in general. In the following sections, two special cases of the generic form of the encoding function given in (3.1) are considered as practical solution approaches.

Remark 1: Note that the closed-form expression for $\boldsymbol{\Sigma}_{err}$ can be derived in the following way (similarly to the derivation for the scalar case in Section 2.2.1). The LMMSE estimator $\hat{\boldsymbol{\beta}}(\boldsymbol{Z})$ is expressed as $\hat{\boldsymbol{\beta}}(\boldsymbol{Z}) = \mathbf{A}\boldsymbol{Z} + \mathbf{b}$, where $\mathbf{A}$ and $\mathbf{b}$ are chosen to minimize $E\left(\left|\left|\hat{\boldsymbol{\beta}}(\boldsymbol{Z}) - \boldsymbol{\beta}\right|\right|^2\right)$, as the eavesdropper is unaware of the encoding, and are given by

$$\mathbf{A} = \boldsymbol{\Sigma}_{\boldsymbol{\beta},\boldsymbol{Z}} \left(\boldsymbol{H}_e \boldsymbol{\Sigma}_{\boldsymbol{\beta}} \boldsymbol{H}_e^T + \boldsymbol{\Sigma}_e\right)^{-1}, \tag{3.8}$$

and

$$\mathbf{b} = \left(\mathbf{I} - \mathbf{A}\boldsymbol{H}_{\mathbf{e}}\right) E(\boldsymbol{\beta}), \tag{3.9}$$

with

$$\boldsymbol{\Sigma}_{\boldsymbol{\beta},\boldsymbol{Z}} = E\Big(\big(\boldsymbol{\beta} - E(\boldsymbol{\beta})\big)\big(\boldsymbol{Z} - E(\boldsymbol{Z})^T\big)\Big). \tag{3.10}$$

Based on (3.8)–(3.10), $\boldsymbol{\Sigma}_{err}$ can be obtained as

$$\begin{aligned}\boldsymbol{\Sigma}_{err} = {} & \boldsymbol{\Sigma}_{\boldsymbol{\beta}}\mathbf{R}\boldsymbol{\Sigma}_{\boldsymbol{\beta}}^T - \boldsymbol{\Sigma}_{\boldsymbol{\beta}}\mathbf{R}\boldsymbol{\Sigma}_{\boldsymbol{\beta},\boldsymbol{\theta}} - \boldsymbol{\Sigma}_{\boldsymbol{\beta},\boldsymbol{\theta}}^T\mathbf{R}\boldsymbol{\Sigma}_{\boldsymbol{\beta}}^T + \boldsymbol{\Sigma}_{\boldsymbol{\theta}} \\ & + \Big(\big(E(\boldsymbol{\beta}) - E(\boldsymbol{\theta})\big)\big(E(\boldsymbol{\beta}) - E(\boldsymbol{\theta})\big)^T\Big),\end{aligned} \tag{3.11}$$

where

$$\begin{aligned}\boldsymbol{\Sigma}_{\boldsymbol{\beta}} &= E\big(\boldsymbol{\beta}\boldsymbol{\beta}^T\big) - E(\boldsymbol{\beta})E(\boldsymbol{\beta})^T, \\ \boldsymbol{\Sigma}_{\boldsymbol{\beta},\boldsymbol{\theta}} &= E\big(\boldsymbol{\beta}\boldsymbol{\theta}^T\big) - E(\boldsymbol{\beta})E(\boldsymbol{\theta})^T, \\ \boldsymbol{\Sigma}_{\boldsymbol{\theta}} &= E\big(\boldsymbol{\theta}\boldsymbol{\theta}^T\big) - E(\boldsymbol{\theta})E(\boldsymbol{\theta})^T, \\ \mathbf{R} &= \boldsymbol{H}_e^T \left(\boldsymbol{H}_e \boldsymbol{\Sigma}_{\boldsymbol{\beta}} \boldsymbol{H}_e^T + \boldsymbol{\Sigma}_e\right)^{-1} \boldsymbol{H}_e.\end{aligned} \tag{3.12}$$

3.2 Nonlinear Individual Encoding

In this section, the proposed problem in Section 3.1 is investigated for an encoding approach such that each parameter θ_i is encoded *individually* by a nonlinear scalar function such that

$$\boldsymbol{\beta} \triangleq \boldsymbol{f}(\boldsymbol{\theta}) = \begin{bmatrix} f_1(\theta_1) \\ f_2(\theta_2) \\ \vdots \\ f_N(\theta_N) \end{bmatrix}. \tag{3.13}$$

Furthermore, as motivated in Section 2.2, the parameter space and the intrinsic constraints on each encoding function $f_i(\theta_i)$ are specified as follows:

- $\theta_i \in [a_i, b_i]$ for $i = 1, 2, \ldots, N$.
- $\beta_i = f_i(\theta_i) \in [a_i, b_i]$ for $i = 1, 2, \ldots, N$.
- f_i is a continuous and one-to-one function.

Under these assumptions, the optimal encoding problem in (3.7) can be written as

$$\begin{aligned} \boldsymbol{f}_{opt} = \arg \min_{f_1(\theta_1),\ldots,f_N(\theta_N)} & \int_{\Lambda} w(\boldsymbol{\theta})\ tr\{\mathbf{I}(\boldsymbol{\theta})^{-1}\} d\boldsymbol{\theta} \\ s.t. \quad & \boldsymbol{\Sigma}_{err}(i) \geq \eta_i\,, \quad i = 1, 2, \ldots, N. \end{aligned} \tag{3.14}$$

In the remainder of this section, the solution of the problem in (3.14) is investigated. To that end, $tr\{\mathbf{I}(\boldsymbol{\theta})^{-1}\}$ for parameter $\boldsymbol{\theta}$ is derived for the system model specified by (3.2) and the error covariance matrix in (3.11) is employed. Note that for a fixed $\boldsymbol{f}$ and channel matrix $\boldsymbol{H}_r$, $\boldsymbol{Y}$ is a Gaussian random vector with

mean $\mu(\boldsymbol{\theta})$ expressed as

$$\mu(\boldsymbol{\theta}) = \boldsymbol{H}_r\boldsymbol{\beta} = \begin{bmatrix} h_{r,1}\, f_1(\theta_1) \\ h_{r,2}\, f_2(\theta_2) \\ \vdots \\ h_{r,N}\, f_N(\theta_N) \end{bmatrix} \tag{3.15}$$

and covariance matrix $\boldsymbol{\Sigma}_r$. Accordingly, each element of $\mathbf{I}(\boldsymbol{\theta})$ can explicitly be written as [68]

$$[\mathbf{I}(\boldsymbol{\theta})]_{i,j} = [\boldsymbol{\Sigma}_r^{-1}]_{i,j} \left(h_{r,i}\frac{df_i(\theta_i)}{d\theta_i}\right)\left(h_{r,j}\frac{df_j(\theta_j)}{d\theta_j}\right), \tag{3.16}$$

where $[\boldsymbol{\Sigma}_r^{-1}]_{i,j}$ denotes the (i,j)th element of $\boldsymbol{\Sigma}_r^{-1}$. Note that if $\alpha_i \triangleq h_{r,i}\frac{df_i(\theta_i)}{d\theta_i}$, then $[\mathbf{I}(\boldsymbol{\theta})]_{i,j} = \alpha_i\alpha_j[\boldsymbol{\Sigma}_r^{-1}]_{i,j}$; thus, the FIM can simply be expressed as $\mathbf{I}(\boldsymbol{\theta}) = \text{diag}\{\alpha_1, \alpha_2, \ldots, \alpha_N\}\boldsymbol{\Sigma}_r^{-1}\text{diag}\{\alpha_1, \alpha_2, \ldots, \alpha_N\}$. Therefore, the following expression is obtained:

$$tr\{\mathbf{I}(\boldsymbol{\theta})^{-1}\} = \sum_{i=1}^{N}\frac{\sigma_{r,i}^2}{\alpha_i^2} = \sum_{i=1}^{N}\frac{\sigma_{r,i}^2}{{h_{r,i}}^2 f_i'(\theta_i)^2} \tag{3.17}$$

where $f_i'(\theta_i)$ denotes the derivative of $f_i(\theta_i)$. Note that (3.17) implies that even though the effective noise is not necessarily white, $tr\{\mathbf{I}(\boldsymbol{\theta})^{-1}\}$ can still be written as the sum of individual scalar inverse Fisher information corresponding to different parameters. Then, the cost function in (3.14) becomes

$$\begin{aligned} &\int_{a_1}^{b_1}\int_{a_2}^{b_2}\cdots\int_{a_N}^{b_N} w(\boldsymbol{\theta})\sum_{i=1}^{N}\frac{\sigma_{r,i}^2}{{h_{r,i}}^2 f_i'(\theta_i)^2}d\theta_1 d\theta_2\ldots d\theta_N \\ &= \sum_{i=1}^{N}\int_{a_1}^{b_1}\int_{a_2}^{b_2}\cdots\int_{a_N}^{b_N} w(\boldsymbol{\theta})\frac{\sigma_{r,i}^2}{{h_{r,i}}^2 f_i'(\theta_i)^2}d\theta_1 d\theta_2\ldots d\theta_N \\ &= \sum_{i=1}^{N}\frac{\sigma_{r,i}^2}{{h_{r,i}}^2}\int_{a_i}^{b_i} w_i(\theta_i)\frac{1}{f_i'(\theta_i)^2}d\theta_i. \end{aligned} \tag{3.18}$$

It is observed that the overall cost function is actually the sum of individual ECRB values for any generic $w(\boldsymbol{\theta})$. Based on (3.11) and (3.18), one can calculate the cost function and the constraints in (3.14) for any given $w(\boldsymbol{\theta})$, $\boldsymbol{\beta}$ and channel

statistics. In the following, two specific scenarios are investigated in more detail.

3.2.1 Independent Parameters & White Gaussian Noise for Eavesdropper

We first consider the scenario in which the channel noise is zero-mean white Gaussian for the eavesdropper[2], that is, $\mathbf{\Sigma}_e = \text{diag}\{\sigma_{e,1}^2, \sigma_{e,2}^2, \ldots, \sigma_{e,N}^2\}$ and the parameters, θ_i's, are independent of each other with marginal distributions denoted by $w_i(\theta_i)$ for $i = 1, 2, \ldots, N$. (Note that $w(\boldsymbol{\theta}) = \prod_{i=1}^{N} w_i(\theta_i)$ in this scenario.) Under this setting, the following proposition reveals that the optimization problem be decoupled into independent scalar problems.

Proposition 1: *If the parameters are independent and the channel noise for the eavesdropper is white Gaussian, the optimization problem in* (3.14) *can be decoupled into independent problems as follows:*

$$\begin{aligned} f_{i,opt} &= \arg\min_{f_i} \int_{a_i}^{b_i} w_i(\theta_i) \frac{1}{f_i^{'}(\theta_i)^2} d\theta_i \\ s.t. \quad & \mathbf{\Sigma}_{err}(i) \geq \eta_i \,, \quad i = 1, 2, \ldots, N. \end{aligned} \tag{3.19}$$

where

$$\mathbf{\Sigma}_{err}(i) = \frac{h_i^2 V_i (V_i - 2C_i)}{h_i^2 V_i + 1} + Var(\theta_i) + \big(E(f_i(\theta_i)) - E(\theta_i)\big)^2 \tag{3.20}$$

$V_i = Var(f_i(\theta_i))$, $C_i = Cov(f_i(\theta_i), \theta_i)$ *and* $h_i = h_{e,i}/\sigma_{e,i}$.

Proof: First, we focus on the error covariance matrix $\mathbf{\Sigma}_{err}$. Note that $\mathbf{\Sigma}_{\boldsymbol{\beta}} = \text{diag}\{V_1, V_2, \ldots, V_N\}$ with $V_i = Var(f_i(\theta_i))$, $\mathbf{\Sigma}_{\boldsymbol{\beta},\boldsymbol{\theta}} = \text{diag}\{C_1, C_2, \ldots, C_N\}$ with $C_i = Cov(f_i(\theta_i), \theta_i)$ and $\mathbf{\Sigma}_{\boldsymbol{\theta}} = \text{diag}\{Var(\theta_1), Var(\theta_2), \ldots, Var(\theta_N)\}$ due to the independence of θ_i's. Also, $\mathbf{R} = \text{diag}\{\frac{{h_{e,1}}^2}{{h_{e,1}}^2 V_1 + \sigma_{e,1}^2}, \frac{{h_{e,2}}^2}{{h_{e,2}}^2 V_2 + \sigma_{e,2}^2}, \ldots, \frac{{h_{e,N}}^2}{{h_{e,N}}^2 V_N + \sigma_{e,N}^2}\}$ due to the independence of θ_i's and the white Gaussian noise assumption for the

[2]Note that there is no further assumption on the noise statistics for the intended receiver, as it does not effect the constraint and the cost function according to (3.11) and (3.17).

eavesdropper. Therefore $\mathbf{\Sigma}_{err} = \text{diag}\{\mathbf{\Sigma}_{err}(1), \mathbf{\Sigma}_{err}(2), \ldots, \mathbf{\Sigma}_{err}(N)\}$, where

$$\mathbf{\Sigma}_{err}(i) = \frac{h_i^2 V_i (V_i - 2C_i)}{h_i^2 V_i + 1} + Var(\theta_i) + \big(E(f_i(\theta_i)) - E(\theta_i)\big)^2 \tag{3.21}$$

and $h_i = h_{e,i}/\sigma_{e,i}$. Based on (3.18) and (3.21), the generic optimization problem in (3.14) reduces to

$$\begin{aligned} \boldsymbol{f}_{opt} &= \arg \min_{f_1, f_2, \ldots, f_N} \sum_{i=1}^{N} \frac{\sigma_{r,i}^2}{{h_{r,i}}^2} \int_{a_i}^{b_i} w_i(\theta_i) \frac{1}{f_i'(\theta_i)^2} d\theta_i \\ s.t. \quad & \mathbf{\Sigma}_{err}(i) \geq \eta_i \,, \quad i = 1, 2, \ldots, N. \end{aligned} \tag{3.22}$$

Note that the constraints are independent of each other and each element of the sum in the objective function has no effect on the others. Therefore, the optimization problem can be decoupled and each θ_i can be optimized individually, where the decoupled problems can be expressed as in (3.19). ■

Remark 2: The optimization problem in (3.19) has been investigated in Section 2.2.1 in detail and the results and the solution methods proposed in that part can directly be applied to the vector parameter problem, when the channel noise for the eavesdropper is white Gaussian and the parameters are independent of each other. Also, when the parameters are not independent, the constraints given in (3.14) include cross terms even if the eavesdropper has white Gaussian noise; therefore, the optimization problem needs to be solved based on (3.14) for correlated parameters.

3.2.2 Independent Parameters & Colored Gaussian Noise Vectors

In this part, we again assume that the parameters are independent of each other, i.e., $w(\boldsymbol{\theta}) = \prod_{i=1}^{N} w_i(\theta_i)$; however, we suppose that $\mathbf{\Sigma}_e$ is a symmetric, positive definite matrix which is not necessarily diagonal. Due to the independence of parameters, $\mathbf{\Sigma}_{\boldsymbol{\beta}}$, $\mathbf{\Sigma}_{\boldsymbol{\beta},\boldsymbol{\theta}}$ and $\mathbf{\Sigma}_{\boldsymbol{\theta}}$ take diagonal forms as in Section 3.2.1. Then, the

ith diagonal element $\mathbf{\Sigma}_{err}(i)$ of $\mathbf{\Sigma}_{err}$ can be written as

$$\begin{aligned}\mathbf{\Sigma}_{err}(i) = h_{e,i}^2 V_i (V_i - 2C_i)\gamma_i + Var(\theta_i) \\ + \big(E(f_i(\theta_i)) - E(\theta_i)\big)^2\end{aligned} \tag{3.23}$$

where V_i and C_i are as defined previously. Also, γ_i is the ith diagonal element of matrix $(\tilde{\mathbf{D}} + \mathbf{\Sigma}_e)^{-1}$, where $\tilde{\mathbf{D}} = \text{diag}\{h_{e,1}^2 V_1, h_{e,2}^2 V_2, \ldots, h_{e,N}^2 V_N\}$. Note that γ_i depends on $\boldsymbol{H}_e$ and the encoding function $\boldsymbol{f}$. Due to the cross terms in the constraints, the optimization problem cannot be decoupled anymore, hence it should be solved using (3.14) based on (3.17) and (3.23). However, it is possible to derive some theoretical results about the form of the solution in the considered scenario. Lemma 1 generalizes Proposition 3 in Chapter 2 for the multivariable case.

Lemma 1: *Suppose that the eavesdropper employs the linear MMSE estimator and $w_i(\theta_i)$ is symmetric around $(a_i + b_i)/2$. Then, for any given encoding function $\boldsymbol{f}(\boldsymbol{\theta})$ which consists of continuous and strictly increasing encoding functions $f_i(\theta_i)$, there exists a corresponding encoding function $\mathbf{s}(\boldsymbol{\theta})$ consisting of continuous and strictly decreasing encoding functions $s_i(\theta_i)$ that yields the same ECRB at the intended receiver with a higher MSE for the individual parameters at the eavesdropper.*

Proof: By using the arguments in Section 2.2.1, we consider two encoding functions $f_i(\theta_i)$ and $s_i(\theta_i) = f_i(a_i + b_i - \theta_i)$, where $\theta_i \in [a_i, b_i]$ and $f_i(\theta_i)$ is a continuous and monotonically increasing function. Since $s_i'(\theta_i) = -f_i'(a_i + b_i - \theta_i)$ by definition and due to the symmetry in $w_i(\theta_i)$, both encoding functions result in the same $tr\{\mathbf{I}(\boldsymbol{\theta})^{-1}\}$, which is given in (3.17). Furthermore, as shown in Section 2.2.1, $Cov(f_i(\theta_i), \theta_i) > Cov(s_i(\theta_i), \theta_i)$ and two encoders yield the same variance and expectation for the encoded version of the parameter. Also, $\mathbf{\Sigma}_e$ is a positive definite matrix and $\tilde{\mathbf{D}}$ has positive entries. Therefore, $(\tilde{\mathbf{D}} + \mathbf{\Sigma}_e)^{-1}$ is also a positive definite matrix[3] and $\gamma_i > 0$ always holds. Combining these results and via (3.23), it is obtained that a larger MSE for parameter θ_i, i.e., $\mathbf{\Sigma}_{err}(i)$, can be

[3]Since $\mathbf{\Sigma}_e$ is a positive definite symmetric matrix, it can be expressed as $\mathbf{\Sigma}_e = \sum_{k=1}^{N} \lambda_k \mathbf{v}_k \mathbf{v}_k^T$ and since $\tilde{\mathbf{D}}$ is diagonal, $(\tilde{\mathbf{D}} + \mathbf{\Sigma}_e)^{-1} = \sum_{k=1}^{N} \frac{1}{\lambda_k + h_{e,k}^2 V_k} \mathbf{v}_k \mathbf{v}_k^T$ can be obtained.

achieved by employing $s_i(\theta_i)$ instead of $f_i(\theta_i)$ while keeping the ECRB the same. ■

Lemma 1 has an important practical implication that the search space for the optimal encoding function for the ith parameter can be restricted to strictly decreasing functions when the sufficient condition given in the lemma is satisfied. Note that Lemma 1 can be applied if θ_i has a symmetric distribution on its domain. Some examples of continuous symmetric distributions on a bounded interval satisfying the condition include uniform distribution, beta distribution with both parameters of 1/2, and raised cosine distribution.

3.2.2.1 Two-Parameter Case ($N = 2$)

In this part, we investigate the case of $N = 2$; that is, $\boldsymbol{\theta} = [\theta_1, \theta_2]^T$. Therefore, the channel noise $\boldsymbol{N}_e$ for the eavesdropper can be modeled as zero-mean Gaussian with covariance matrix $\boldsymbol{\Sigma}_e = \begin{bmatrix} \sigma_{e,1}^2 & \rho \\ \rho & \sigma_{e,2}^2 \end{bmatrix}$. For this particular case, γ_i in (3.23) can explicitly be written as

$$\gamma_1 = \frac{h_{e,2}^2 V_2 + \sigma_{e,2}^2}{(h_{e,1}^2 V_1 + \sigma_{e,1}^2)(h_{e,2}^2 V_2 + \sigma_{e,2}^2) - \rho^2} \tag{3.24}$$

and γ_2 can be obtained by replacing the numerator in (3.24) with $h_{e,1}^2 V_1 + \sigma_{e,1}^2$. After some manipulation, $\boldsymbol{\Sigma}_{err}(1)$ can be derived as

$$\begin{aligned} \boldsymbol{\Sigma}_{err}(1) =& \lambda \; E(|\beta_1 - \theta_1|^2) \\ &+ (1-\lambda)\left((E(\beta_1) - E(\theta_1))^2 + Var(\theta_1)\right) \end{aligned} \tag{3.25}$$

where

$$\lambda = \frac{h_1^2 V_1}{h_1^2 V_1 + 1 - r_2(\rho)}$$

with

$$r_2(\rho) = \frac{\rho^2/\sigma_{e,1}^2}{h_{e,2}^2 V_2 + \sigma_{e,2}^2}$$

and $h_1 = h_{e,1}/\sigma_{e,1}$.

It is possible to gain practical intuition about the behavior of the optimal encoding function as a closed-form expression for $\mathbf{\Sigma}_{err}(1)$ (and $\mathbf{\Sigma}_{err}(2)$) is available. There are several important observations related to (3.25).

- For a fixed $r_2(\rho)$, if we let $h_1^2 \to \infty$, then $\mathbf{\Sigma}_{err}(1) \approx E(|\beta_1 - \theta_1|^2)$; hence, it is maximized when $E(|\beta_1 - \theta_1|^2)$ is maximized. This mode can be called as the *variance maximizing mode* as in Section 2.2.1. If we let $h_1^2 \to 0$, then $\mathbf{\Sigma}_{err}(1) \approx (E(\beta_1) - E(\theta_1))^2 + Var(\theta_1)$; therefore, it is maximized if $\beta_1 \to a_1$ or $\beta_1 \to b_1$. This mode can be called as the *variance minimizing mode*.

- For a fixed h_1 (and relevant parameters for θ_2), as ρ^2 increases, $r_2(\rho)$ and λ also increase. According to (25), if λ is small enough, the encoder is in the variance minimizing mode; however, as λ increases and becomes large enough, maximizing $E(|\beta_1 - \theta_1|^2)$ becomes the priority. As ρ increases, after a certain threshold, which can be denoted as ρ_0, the mode of operation can change and the encoder can get into the variance maximizing mode when $\rho > \rho_0$.

Note that in the analysis above h_1^2 can be viewed as the signal-to-noise ratio (SNR) for the channel of θ_1 to the eavesdropper. As the SNR of this channel increases, the distortion due to encoding is transmitted to the eavesdropper more effectively and the main factor to create a large MSE at the eavesdropper is the distortion to the parameter via encoding in the variance maximizing mode. Also, when $h_1 \to 0$, this means that the channel is very noisy; hence, the only information available to the eavesdropper through its observation is the mean of the encoded version of the parameter. Therefore, the encoder tries to ensure that the mean of the encoded version is away from the true mean. Note that in practice, even if the SNR values are not necessarily in absolute limits, we can

still observe the aforementioned behavior in the encoding functions (see Figs. 3.3 and 3.5). Hence, it can be concluded that the form of encoding function depends on the parameters of the channel and the correlation between eavesdropper's noise components. Finally, we note that a similar derivation and analysis can be performed for $\boldsymbol{\Sigma}_{err}(2)$ based on γ_2 and (3.23).

3.3 Affine Joint Encoding Strategy

In this section, the encoding operation is assumed to be an affine function. Namely, the vector parameter $\boldsymbol{\theta}$ is encoded by using an $N \times N$ precoding matrix $\boldsymbol{P}$ and an N-dimensional constant vector $\boldsymbol{r}$ prior to transmission such that $\boldsymbol{\beta} = \boldsymbol{P}\boldsymbol{\theta} + \boldsymbol{r}$. Under this assumption, the optimal parameter encoding problem can be expressed as follows:

$$[\boldsymbol{P}_{opt}, \boldsymbol{r}_{opt}] = \arg\min_{\boldsymbol{P},\boldsymbol{r}} \int_{\Lambda} w(\boldsymbol{\theta})\ tr\{\mathbf{I}(\boldsymbol{\theta})^{-1}\} d\boldsymbol{\theta}$$
$$s.t. \quad \boldsymbol{\Sigma}_{err}(i) \geq \eta_i\,, \quad i = 1, 2, \ldots, N. \tag{3.26}$$

As in the previous section, the parameter space is specified as $\theta_i \in [a_i, b_i]$, for $i = 1, 2, \ldots, N$ for this strategy. If we define $a \triangleq \min\{a_1, a_2, \ldots a_N\}$ and $b \triangleq \max\{b_1, b_2, \ldots b_N\}$, then $\theta_i \in [a, b]$, for $i = 1, 2, \ldots, N$. In this section, it is assumed that the generalized domain of the parameters, i.e., $[a, b]$, needs to be preserved after the encoding operation; hence, it is assumed that $\beta_i \in [a, b]$, for $i = 1, 2, \ldots, N$. This condition can be guaranteed if the sum of the absolute values of the elements in each row of $\boldsymbol{P}$ is less than or equal to 1. This can formally be expressed as $\|\boldsymbol{P}^T \boldsymbol{e}_j\|_1 \leq 1$ for $j = 1, 2, \ldots, N$, where $\boldsymbol{e}_j$'s are standard basis vectors.[4] Finally, the precoding matrix $\boldsymbol{P}$ is taken to be full rank (invertible).

In the remainder of this section, the solution of the problem in (3.26) is investigated. First, $tr\{\mathbf{I}(\boldsymbol{\theta})^{-1}\}$ for parameter $\boldsymbol{\theta}$ is derived for the given system model and encoding strategy. Note that $\boldsymbol{Y}$ is a Gaussian random vector with mean

[4] $\|\mathbf{x}\|_1 \triangleq \sum_{i=1}^{N} |x_i|$ is called the l_1 norm of vector $\mathbf{x}$.

$\mu(\boldsymbol{\theta}) = \boldsymbol{H}_r\boldsymbol{\beta} = \boldsymbol{H}_r\boldsymbol{P}\boldsymbol{\theta} + \boldsymbol{H}_r\boldsymbol{r}$ and covariance matrix $\boldsymbol{\Sigma}_r$ for fixed $\boldsymbol{P}$, $\boldsymbol{r}$ and channel matrix $\boldsymbol{H}_r$. Therefore, each element of $\mathbf{I}(\boldsymbol{\theta})$ can explicitly be written as

$$\begin{aligned}[\mathbf{I}(\boldsymbol{\theta})]_{i,j} &= \left(\frac{d\boldsymbol{\mu}(\boldsymbol{\theta})}{d\theta_i}\right)^T \boldsymbol{\Sigma}_r^{-1} \left(\frac{d\boldsymbol{\mu}(\boldsymbol{\theta})}{d\theta_i}\right) \\ &= \boldsymbol{p}_i^T \boldsymbol{H}_r \boldsymbol{\Sigma}_r^{-1} \boldsymbol{H}_r \boldsymbol{p}_j \end{aligned} \tag{3.27}$$

where $\boldsymbol{p}_i$ denotes the ith column of precoding matrix $\boldsymbol{P}$. Accordingly, the FIM can be expressed as

$$\begin{aligned}\mathbf{I}(\boldsymbol{\theta}) &= \boldsymbol{P}^T \boldsymbol{H}_r \boldsymbol{\Sigma}_r^{-1} \boldsymbol{H}_r \boldsymbol{P} \\ &= \boldsymbol{P}^T \boldsymbol{D} \boldsymbol{P} \end{aligned} \tag{3.28}$$

where $\boldsymbol{D} \triangleq \boldsymbol{H}_r \boldsymbol{\Sigma}_r^{-1} \boldsymbol{H}_r$. Note that $\boldsymbol{D}$ and $I(\boldsymbol{\theta})$ are positive definite, invertible and symmetric matrices. Also, $\mathbf{I}(\boldsymbol{\theta})$ is not a function of $\boldsymbol{\theta}$. Therefore, the objective function in (3.26) simplifies to

$$\int_{\Lambda} w(\boldsymbol{\theta})\ tr\{\mathbf{I}(\boldsymbol{\theta})^{-1}\} d\boldsymbol{\theta} = tr\left\{\left(\boldsymbol{P}^T \boldsymbol{D} \boldsymbol{P}\right)^{-1}\right\}. \tag{3.29}$$

Note that the objective function depends only on $\boldsymbol{P}$ and the constant factor $\boldsymbol{r}$ in the encoding operation does not effect its value. Furthermore, if the zero-mean Gaussian random noise $\boldsymbol{N}_r$ in the received signal has independent components, then $\boldsymbol{D}$ becomes a diagonal matrix with its ith diagonal element being given by $h_{r,i}^2/\sigma_{r,i}^2$, where $\sigma_{r,i}^2$ is the variance of the ith noise component in $\boldsymbol{N}_r$.

The following proposition provides an optimal solution to the affine joint encoding problem without any secrecy constraints for a diagonal $\boldsymbol{D}$.

Proposition 2: *Assume $\boldsymbol{D}$ is a diagonal matrix. In the absence of secrecy constraints on the eavesdropper, any signed permutation matrix[5] is an optimal solution. Furthermore, any other precoding matrix with a different form is not optimal.*

[5] A signed permutation matrix is defined as a matrix whose every row and column has exactly one non-zero entry, which can be either 1 or -1.

Proof: In the absence of secrecy constraints, the optimization problem can be formulated as

$$\begin{aligned} \boldsymbol{P}_{opt} = \arg\min_{\boldsymbol{P}} & \, tr\left\{\left(\boldsymbol{P}^T\boldsymbol{D}\boldsymbol{P}\right)^{-1}\right\} \\ s.t. \quad & \|\boldsymbol{P}^T\boldsymbol{e}_j\|_1 \leq 1\,, \quad j = 1, 2, \ldots, N. \end{aligned} \tag{3.30}$$

Then, a lower bound for any given feasible $\boldsymbol{P}$ can be obtained as follows:

$$\begin{aligned} tr\left\{\left(\boldsymbol{P}^T\boldsymbol{D}\boldsymbol{P}\right)^{-1}\right\} &= tr\left\{\left(\boldsymbol{P}^{-1}\boldsymbol{D}^{-1}\boldsymbol{P}^{-T}\right)\right\} \\ &= \left\|\boldsymbol{P}^{-1}\boldsymbol{D}^{-1/2}\right\|_F^2 \\ &= \sum_{j=1}^{N} \frac{1}{\lambda_j}\|\boldsymbol{m}_j\|_2^2 \end{aligned} \tag{3.31}$$

where $\boldsymbol{M} \triangleq \boldsymbol{P}^{-1}$, $\boldsymbol{m}_j$ is the jth column of $\boldsymbol{M}$ and $\boldsymbol{D} = diag\{\lambda_1, \lambda_2, \ldots, \lambda_N\}$. Note that $\boldsymbol{P}\boldsymbol{M} = \mathbf{I}$, thus $\boldsymbol{p}_j^{(r)}\boldsymbol{m}_j = 1$ for $j = 1, 2, \ldots N$, and $\boldsymbol{p}_j^{(r)} = \boldsymbol{e}_j^T\boldsymbol{P}$ is the jth row of $\boldsymbol{P}$. As the sum of the absolute values of the elements in each row cannot be greater than 1, $\|\boldsymbol{p}_j^{(r)}\|_2 \leq \|\boldsymbol{p}_j^{(r)}\|_1 \leq 1$. Also, via Cauchy-Schwarz inequality, it can be obtained that $1 = |\boldsymbol{p}_j^{(r)}\boldsymbol{m}_j|^2 \leq \|\boldsymbol{p}_j^{(r)}\|_2^2\|\boldsymbol{m}_j\|_2^2$; hence, as $\|\boldsymbol{p}_j^{(r)}\|_2 \leq 1$, $\|\boldsymbol{m}_j\|_2 \geq 1$ for $j = 1, 2, \ldots, N$. Therefore,

$$tr\left\{\left(\boldsymbol{P}^T\boldsymbol{D}\boldsymbol{P}\right)^{-1}\right\} = \sum_{j=1}^{N} \frac{1}{\lambda_j}\|\boldsymbol{m}_j\|_2^2 \geq \sum_{j=1}^{N} \frac{1}{\lambda_j} \tag{3.32}$$

for any given feasible $\boldsymbol{P}$. Note that this lower bound can exactly be attained when $\|\boldsymbol{m}_j\|_2 = 1$, which implies $\|\boldsymbol{p}_j^{(r)}\|_2 = 1$ for an optimal solution. Also, due to the relation $1 = \|\boldsymbol{p}_j^{(r)}\|_2 \leq \|\boldsymbol{p}_j^{(r)}\|_1 \leq 1$ for $j = 1, 2, \ldots, N$, $\|\boldsymbol{p}_j^{(r)}\|_2 = \|\boldsymbol{p}_j^{(r)}\|_1 = 1$. This is satisfied if and only if $\boldsymbol{p}_j^{(r)}$ contains an element with a value of $+1$ or -1 and the rest of its elements are zero. Due to the rank constraint, each $\boldsymbol{p}_j^{(r)}$ should have the non-zero element at a different location and this is satisfied if and only if the precoding matrix is a signed permutation matrix. ■

Proposition 2 reveals that if there is no secrecy constraint for a given diagonal $\boldsymbol{D}$, then a signed permutation matrix can be used as the optimal precoding

matrix.

Next, the optimal affine joint encoding problem is considered in the presence of secrecy constraints. The error covariance matrix $\boldsymbol{\Sigma}_{err}$ in the constraint of (3.26) can be calculated based on the procedure in Remark 1. Specifically, it can be obtained by using the equations given in (3.11) and (3.12) and inserting $\boldsymbol{\Sigma}_{\boldsymbol{\beta}} = \boldsymbol{P}\boldsymbol{\Sigma}_{\boldsymbol{\theta}}\boldsymbol{P}^T$ and $\boldsymbol{\Sigma}_{\boldsymbol{\beta},\boldsymbol{\theta}} = \boldsymbol{P}\boldsymbol{\Sigma}_{\boldsymbol{\theta}}$. Note that only the last term in (3.11) depends on $\boldsymbol{r}$. As only the diagonal terms are taken into consideration for the secrecy targets, they can explicitly be calculated. The following lemma is provided regarding the relationship between $\boldsymbol{\Sigma}_{err}$ and $\boldsymbol{r}$ for any given $\boldsymbol{P}$ and $w(\boldsymbol{\theta})$.

Lemma 2: *When the eavesdropper employs the linear MMSE estimator, then* $\boldsymbol{\Sigma}_{err}(i)$*, (i.e., the ith diagonal element of* $\boldsymbol{\Sigma}_{err}$*) for the encoding operation* $\boldsymbol{\beta} = \boldsymbol{P}\boldsymbol{\theta} + \boldsymbol{r}$ *is a convex function of* r_i*, i.e., the ith element of* $\boldsymbol{r}$ *for a fixed* $\boldsymbol{P}$*.*

Proof: Consider the expression for $\boldsymbol{\Sigma}_{err}$ in Remark 1 (see (3.11) and (3.12)). It is noted that only the last term in (3.11) depends on $\boldsymbol{r}$, which can be written as

$$\begin{aligned}&\Big(\big(E(\boldsymbol{\beta}) - E(\boldsymbol{\theta})\big)\big(E(\boldsymbol{\beta}) - E(\boldsymbol{\theta})\big)^T\Big) = \\ &(\boldsymbol{P} - \mathbf{I})\, E(\boldsymbol{\theta})E(\boldsymbol{\theta})^T\, (\boldsymbol{P} - \mathbf{I})^T + \boldsymbol{r}E(\boldsymbol{\theta})^T\, (\boldsymbol{P} - \mathbf{I})^T \\ &+ (\boldsymbol{P} - \mathbf{I})\, E(\boldsymbol{\theta})\boldsymbol{r}^T + \boldsymbol{r}\boldsymbol{r}^T\,. \end{aligned} \tag{3.33}$$

For a given $\boldsymbol{P}$, the contribution of (3.33) (i.e., the last term of $\boldsymbol{\Sigma}_{err}$) to $\boldsymbol{\Sigma}_{err}(i)$, denoted as $g(i)$, can be calculated as

$$g(i) = \Big(r_i + \boldsymbol{p}_i^{(r)} E(\boldsymbol{\theta}) - E(\theta_i)\Big)^2\,, \tag{3.34}$$

where $\boldsymbol{p}_i^{(r)}$ is the ith row of $\boldsymbol{P}$. As the other terms of $\boldsymbol{\Sigma}_{err}$ does not depend on $\boldsymbol{r}$ (see (3.11)) and $\frac{d^2 g(i)}{dr_i^2} = 2 > 0$, the convexity claim in the lemma holds. ■

As a result of Lemma 2, $\boldsymbol{\Sigma}_{err}(i)$ is maximized either at r_i^{min} or r_i^{max}, where r_i^{min} and r_i^{max} are, respectively, the lowest and highest possible values of r_i for a given $\boldsymbol{P}$, while ensuring that the ith element of $\boldsymbol{P}\boldsymbol{\theta}+\boldsymbol{r}$, i.e., β_i, is in $[a, b]$. For example,

if $\theta_1, \theta_2 \in [0,1]$ and $\boldsymbol{P} = \begin{bmatrix} 0.1 & 0.5 \\ 0 & -0.8 \end{bmatrix}$, then $0 \leq r_1 \leq 0.4$ and $0.8 \leq r_2 \leq 1$ to ensure $\beta_1, \beta_2 \in [0,1]$. Therefore, $r_1^{min} = 0$, $r_1^{max} = 1$, $r_2^{min} = 0.8$ and $r_2^{max} = 1$ for this particular example. Among r_i^{min} or r_i^{max}, the one that yields a higher $\boldsymbol{\Sigma}_{err}(i)$ can be selected. As the objective function in (3.26) does not depend on $\boldsymbol{r}$, it can freely be selected to maximize $\boldsymbol{\Sigma}_{err}(i)$ for a given $\boldsymbol{P}$; therefore, it is sufficient to search over precoding matrices for the optimal strategy.

Corollary 1: *Suppose that eavesdropper's noise has independent components, and $\beta_i = w_i\theta_j + r_i$ for some $i \neq j$. If either of $E(\theta_i)$ or $E(\theta_j)$ is equal to $\frac{a+b}{2}$, then, the sign of w_i does not effect $\boldsymbol{\Sigma}_{err}(i)$.*

Proof: We prove the statement for the case of $E(\theta_i) = (a+b)/2$, as it can be shown for $E(\theta_j) = (a+b)/2$ in a similar fashion. First, we note that $\boldsymbol{\Sigma}_{err} = \boldsymbol{\Sigma}_{err}^{(1)} + \boldsymbol{\Sigma}_{err}^{(2)}$ such that $\boldsymbol{\Sigma}_{err}^{(1)}$ represents the first four terms of the sum in (3.11) and $\boldsymbol{\Sigma}_{err}^{(2)}$ denotes the last term. Under the condition in the corollary, w_i's appear in the form of w_i^2's in the diagonals of $\boldsymbol{\Sigma}_{err}^{(1)}$. Therefore, the sign of w_i does not have any effect on $\boldsymbol{\Sigma}_{err}^{(1)}$. For $\boldsymbol{\Sigma}_{err}^{(2)}$, if $\beta_i = w_i\theta_j + r_i$, then we know that $\boldsymbol{\Sigma}_{err}^{(2)}(i) = (r_i + w_i E(\theta_j) - E(\theta_i))^2$. As $\boldsymbol{\Sigma}_{err}^{(2)}(i)$ is maximized either at r_i^{min} or r_i^{max} due to Lemma 2, we have

$$\boldsymbol{\Sigma}_{err}^{(2)}(i) = \max\left\{ \left(\frac{b-a}{2} + \alpha(E(\theta_j) - b)\right)^2, \left(\frac{a-b}{2} + \alpha(E(\theta_j) - a)\right)^2 \right\}$$

for $w_i = \alpha > 0$ and

$$\boldsymbol{\Sigma}_{err}^{(2)}(i) = \max\left\{ \left(\frac{b-a}{2} - \alpha(E(\theta_j) - a)\right)^2, \left(\frac{a-b}{2} - \alpha(E(\theta_j) - b)\right)^2 \right\}$$

for $w_i = -\alpha < 0$. Note that the $\boldsymbol{\Sigma}_{err}^{(2)}(i)$ expressions are exactly the same for both sign options for w_i as long as $|w_i|$ does not change. Therefore, $\boldsymbol{\Sigma}_{err}(i)$ does not depend on the sign of w_i. ■

Lemma 3: *Suppose the encoding matrix* $\boldsymbol{P}$ *has the form of* $\boldsymbol{P} = \boldsymbol{W}_1\boldsymbol{W}_2$, *where* $\boldsymbol{W}_1 = diag\{w_1, w_2, \dots w_N\}$ *is a diagonal matrix and* $\boldsymbol{W}_2$ *is a permutation matrix. Then,* $tr\{\left(\boldsymbol{P}^T\boldsymbol{D}\boldsymbol{P}\right)^{-1}\}$ *does not depend on the signs of the elements in* $\boldsymbol{P}$.

Proof: Note that if $\boldsymbol{P} = \boldsymbol{W}_1\boldsymbol{W}_2$, then

$$\begin{aligned} tr\left\{\left(\boldsymbol{P}^T\boldsymbol{D}\boldsymbol{P}\right)^{-1}\right\} &= tr\left\{\left(\boldsymbol{W}_2^T\boldsymbol{W}_1\boldsymbol{D}\boldsymbol{W}_1\boldsymbol{W}_2\right)^{-1}\right\} \\ &= tr\{\boldsymbol{W}_2^T\hat{\boldsymbol{W}}_1\boldsymbol{D}^{-1}\hat{\boldsymbol{W}}_1\boldsymbol{W}_2\} \\ &= tr\{\boldsymbol{W}_2\boldsymbol{W}_2^T\hat{\boldsymbol{W}}_1\boldsymbol{D}^{-1}\hat{\boldsymbol{W}}_1\} \\ &= tr\{\hat{\boldsymbol{W}}_1\boldsymbol{D}^{-1}\hat{\boldsymbol{W}}_1\} \\ &= \sum_{j=1}^{N}\frac{\hat{d}_j}{w_j^2} \end{aligned} \tag{3.35}$$

where $\hat{\boldsymbol{W}}_1 = \boldsymbol{W}_1^{-1} = diag\{1/w_1, 1/w_2, \dots, 1/w_N\}$ and $\hat{d}_j$ is the jth diagonal element of $\boldsymbol{D}^{-1}$. As $tr\{(\boldsymbol{P}^T\boldsymbol{D}\boldsymbol{P})^{-1}\}$ is the sum of squares, the signs of w_i's do not effect its value. ■

Corollary 1 and Lemma 3 imply that if the encoder applies the method of simple shuffle and scale, then the sign of the scaling factor does not matter in terms of the cost and objective of the optimization. Therefore, optimal scaling factors can be assumed to be positive without loss of generality, which reduces the search space.

Remark 3: By Proposition 2, we know that when $\boldsymbol{D}$ is a diagonal matrix, permutation matrices (with $+1$ or -1 as nonzero elements) are optimal precoding matrices. Also, the optimal precoder belongs to this family of matrices up to a certain secrecy target level $\eta^\dagger$ for each parameter. In other words, if the secrecy target for a given parameter is larger than $\eta^\dagger$, then the objective will be larger and the optimal precoder will not be a permutation matrix anymore. The exact value of $\eta^\dagger$ can be found by solving the following optimization problem:

$$\eta^\dagger = \max_{\boldsymbol{P}\in\mathcal{P}} \min_i \boldsymbol{\Sigma}_{err}(i) \tag{3.36}$$

where $\mathcal{P}$ denotes the set of permutation matrices with $+1$ or -1 as non-zero elements and $\mathbf{\Sigma}_{err}$ is as given in (3.11). Note that there are $2^N N!$ elements in $\mathcal{P}$; therefore, as N gets larger, it gets challenging to solve the optimization problem in (3.36). However, for small N's, it can be solved and provides a practical limit for the secrecy level that can be satisfied without increasing the ECRB values of the case without any secrecy concerns.

3.4 Numerical Results

In this section, numerical results are provided for both strategies proposed in Section 3.2 and Section 3.3.

3.4.1 Nonlinear Individual Encoding

In all the numerical examples for the individual encoding strategy, $\boldsymbol{\theta}$ is modeled as $\boldsymbol{\theta} = [\theta_1 \ \theta_2]^T$, where both θ_1 and θ_2 are uniformly distributed in $[0, 1]$ and are independent of each other. The channel parameters for the intended receiver are taken to be $h_{r,1} = h_{r,2} = 2$ and $\sigma_{r,1}^2 = \sigma_{r,2}^2 = 1$. As the conditions in Lemma 1 are satisfied, the optimal encoding functions are searched among decreasing functions. For the first example, the eavesdropper fading coefficients are taken as $h_{e,2} = 1.5$ and $h_{e,1} \in \{1, 1.2\}$. The channel noise for the eavesdropper is modeled as zero-mean multivariate Gaussian random variable with the covariance matrix $\mathbf{\Sigma}_e = \begin{bmatrix} \sigma_{e,1}^2 & \rho \\ \rho & \sigma_{e,2}^2 \end{bmatrix}$, where $\sigma_{e,1}^2 = \sigma_{e,2}^2 = 1$. The target secrecy levels are $\eta_1 = \eta_2 = 0.15$. In order to solve the optimization problem in (3.14), the approximation methods described in Section 2.2.2 can be used. In this chapter, the piecewise linear approximation method is employed. Namely, for each $f_i(\theta_i)$, $\Delta x_k^{(i)} \triangleq f_i(a_i + k\Delta\theta_i) - f_i(a_i + (k-1)\Delta\theta_i)$ is defined, and the optimization is performed over MN variables; that is, the increments/decrements for each parameter ($\Delta x^{(i)} = [\Delta x_1^{(i)}, \Delta x_2^{(i)}, \ldots, \Delta x_M^{(i)}]$ for $i = 1, 2, \ldots, N$) are obtained. For the numerical results, M is taken to be 50 and Global Optimization Toolbox of

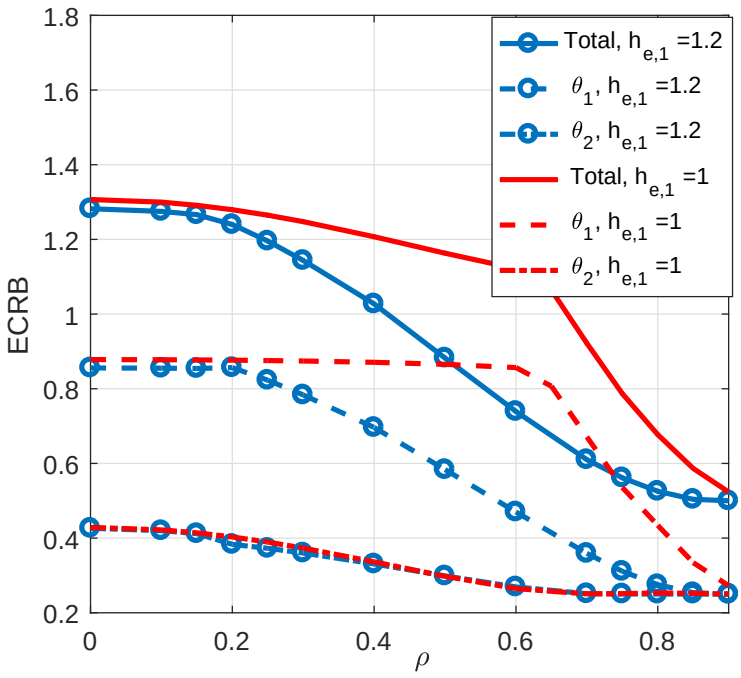

Figure 3.2: Total and individual ECRB values versus ρ for $h_{e,1} = 1$ and $h_{e,1} = 1.2$.

MATLAB is used.

In Fig. 3.2, the total and individual ECRB values for θ_1 and θ_2 are plotted for various ρ values. It is observed that as ρ increases, the total and individual ECRB values decrease, which implies that the correlation between the noise components of the eavesdropper for each parameter is useful for our design purposes. Also, the ECRB for θ_1 decreases very slightly until a certain value of ρ_0 (i.e., $\rho_0 \approx 0.2$ and 0.6 for $h_{e,1} = 1.2$ and 1, respectively), and then a sharper decrease in the ECRB is observed. This is due to the fact that the encoding mode for θ_1 changes as explained in Section 3.2.2.1. Another interesting observation is that for $h_{e,1} = 1.2$, the total and individual ECRB for θ_1 is lower than that in the case of $h_{e,1} = 1$ and the ECRB for θ_2 stays almost the same. The reason for having a lower total ECRB for a larger $h_{e,1}$ is the fact that the eavesdropper is unaware of encoding; hence, the distortion due to the encoding function is transmitted more effectively to the eavesdropper. Also, for larger values of ρ, the ECRB values for both parameters converge to each other.

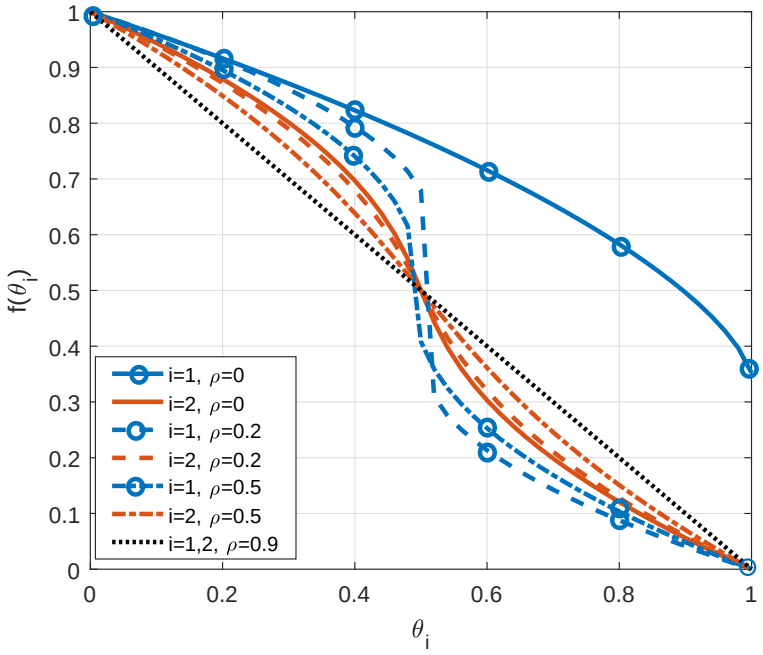

Figure 3.3: The optimal encoding functions for θ_1 and θ_2 for $\rho = \{0, 0.2, 0.5, 0.9\}$ when $h_{e,1} = 1.2$.

In Fig. 3.3, the optimal encoding functions for θ_1 and θ_2 are presented for $\rho \in \{0, 0.2, 0.5, 0.9\}$ when $h_{e,1} = 1.2$. This figure explains some of the behaviors observed in Fig. 3.2. For example, when $\rho = 0$, $f_1(\theta_1)$ is in the variance minimizing mode and $f_2(\theta_2)$ is in the variance maximizing mode. [6] As ρ increases, the changes in $f_2(\theta_2)$ are not significant and there is no mode change. On the other hand, the characteristics of $f_1(\theta_1)$ change when ρ increases, and it gets into the variance maximizing mode for $\rho \in \{0.2, 0.5, 0.9\}$. Also, both encoding functions are linear, $f_i(\theta_i) = 1 - \theta_i$, for $\rho = 0.9$, yielding the same ECRB.

For the second example, $h_{e,1} = 1.2$, $h_{e,2} = 1.5$, $\sigma^2_{e,1} = \sigma^2_{e,2} = 1$ and $\rho = 0.3$. The target secrecy level for θ_2 is fixed to be $\eta_2 = 0.15$, and the target secrecy level for θ_1 is increased starting from 0.1. In Fig. 3.4, the total and individual ECRB values for θ_1 and θ_2 are plotted for various η_1 values. Note that the change in the secrecy target for θ_1 does not have any significant effect on the ECRB performance

[6] Practically, in the variance minimizing mode, the encoder effectively decreases the transmitted signal power to hide the parameter; and in the variance maximizing mode, it has a two-level *quantizer-like* behavior to ensure secrecy.

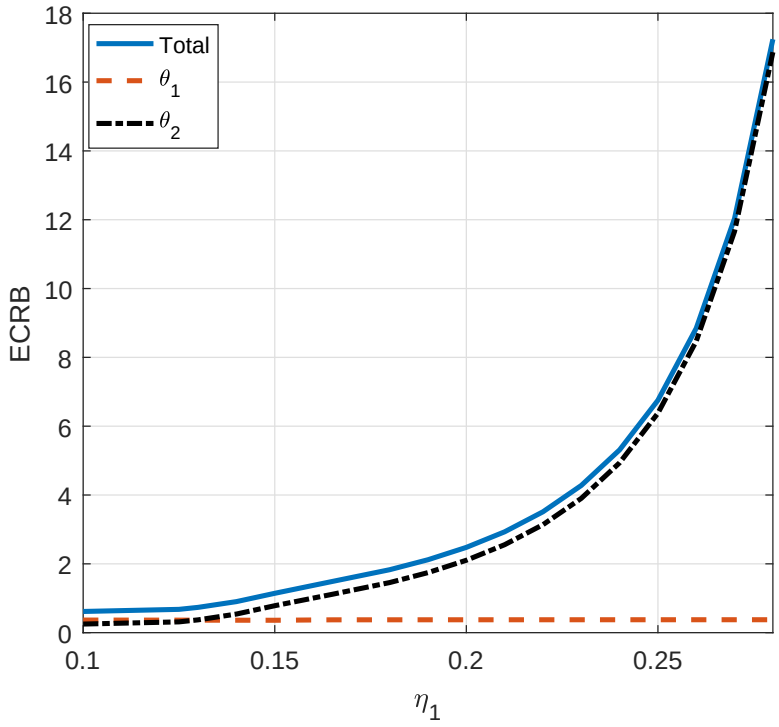

Figure 3.4: Total and individual ECRB values versus η_1.

of θ_2. However, the ECRB for θ_1 and the total ECRB increase exponentially as η_1 increases. The reason of this can be deduced from Fig. 3.5. In Fig. 3.5, the optimal encoding functions for θ_1 and θ_2 are given for $\eta_1 \in \{0.1, 0.15, 0.2, 0.25\}$. It is observed that when $\eta_1 = 0.1$, $f_1(\theta_1) = 1 - \theta_1$. When $\eta_1 = 0.15$, $f_1(\theta_1)$ operates in the variance maximizing mode, and for $\eta_1 = 0.2$ and 0.25, it is in the variance minimizing mode. Note that as η_1 increases, $f_1(\theta_1)$ approaches to 1. (Note that as $f_1(\theta_1) \to 1$, the ECRB goes to ∞). Also, note that the encoding function for θ_2 is insensitive to changes in η_1; that is, $f_2(\theta_2)$ does not change even though η_1 increases, and it is the same for all values of η_1 in this example.

In order to demonstrate the advantages of the proposed encoding scheme, the solution based on Section 2.2 is selected as a benchmark scheme, and a direct performance comparison between the optimal solution based on NIE and the solution based on Section 2.2 is provided in Fig. 3.6. Note that the individual encoding functions are obtained independently for each element of the vector parameter in the benchmark scheme as Section 2.2 provides a solution method for

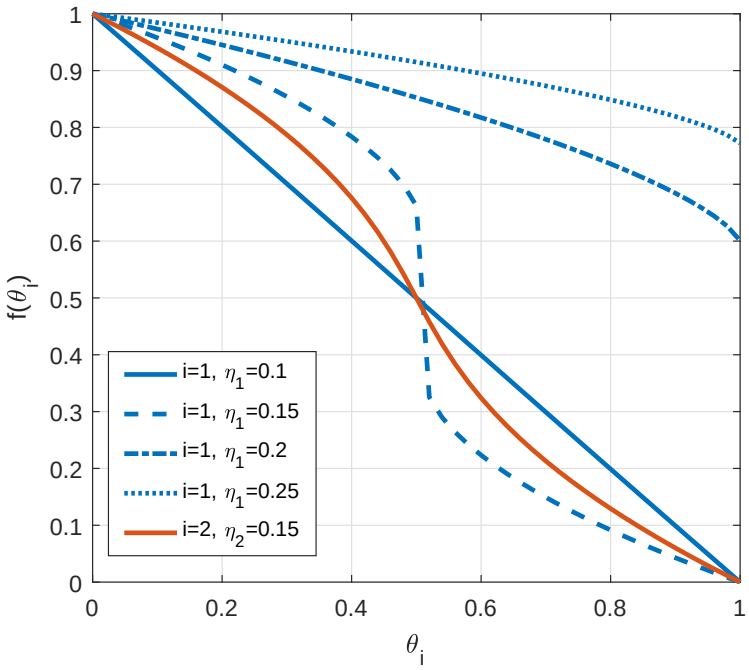

Figure 3.5: The optimal encoding functions for θ_1 and θ_2 for $\eta_1 \in \{0.1, 0.15, 0.2, 0.25\}$ and $\eta_2 = 0.15$.

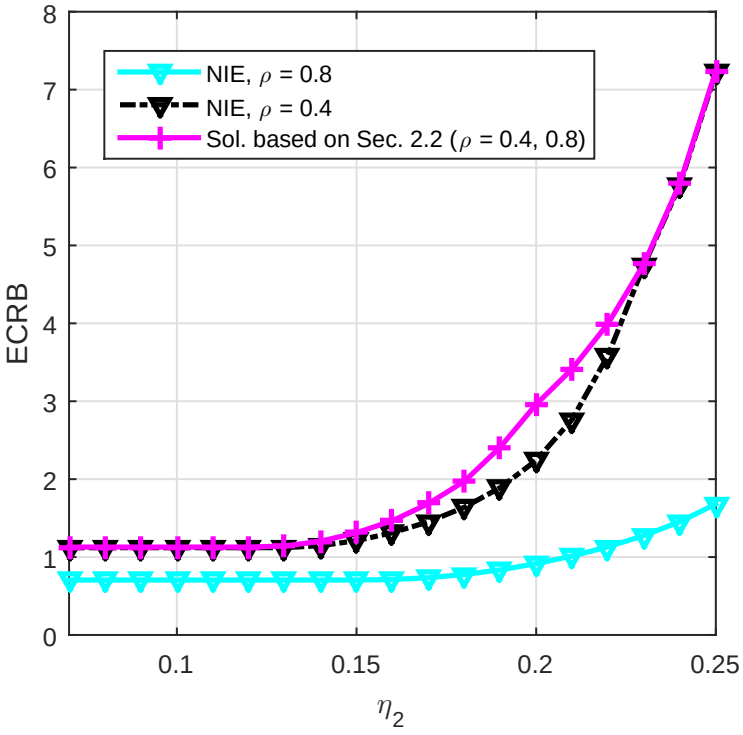

Figure 3.6: Total ECRB values versus η_2 for different approaches.

$h_{e,1}=1, h_{e,2}=1.5$	η_1	η_2	$h_{e,1}=1.2, h_{e,2}=1.5$	η_1	η_2
$\rho=0$	0.0769	0.0702	$\rho=0$	0.0744	0.0702
$\rho=0.3$	0.0764	0.0692	$\rho=0.3$	0.0738	0.0692
$\rho=0.5$	0.0754	0.0670	$\rho=0.5$	0.0723	0.0671
$\rho=0.7$	0.0730	0.0621	$\rho=0.7$	0.0692	0.0625
$\rho=0.9$	0.0660	0.0478	$\rho=0.9$	0.0605	0.0497

Table 3.1: Maximum secrecy target level values for θ_1 and θ_2, when $f_i(\theta_i)=\theta_i$ for $i=1,2$.

scalar problems. In this scenario, the ECRB is plotted versus η_2 for the solution based on Section 2.2 and NIE when $\rho=0.4$ and $\rho=0.8$ and the parameters are set to $h_{e,1}=1$, $h_{e,2}=1.5$, and $\eta_1=0.15$. Note that the solution based on Section 2.2 is the same for both ρ values, as it does not take ρ into account. It is observed that NIE has better performance than the solution based on Section 2.2, and the performance gap dramatically increases when the noise components have high correlation in this scenario. This is intuitive as optimizing the encoders in a joint manner makes sense in a correlated environment. However, if the correlation is decreased, the performance of NIE will converge to that of the solution based on Section 2.2 as proven in Proposition 1. Note that this can be observed in Fig. 3.2 as well. The performance of NIE and the solution based on Section 2.2 would be same for $\rho=0$, and as ρ increases, ECRB of NIE starts to decrease in Fig. 3.2, however the solution based on Section 2.2 would stay constant, yielding a non-negligible performance difference especially in scenarios with medium and high correlation in the noise components.

Finally, the maximum estimation error values at the the eavesdropper are given in Table I when the parameters are directly sent to the channel without any encoding, i.e., $f_i(\theta_i)=\theta_i$ for $i=1,2$, to further emphasize the importance of the encoding operation. If there exists no eavesdroppers, not applying any encoding is a logical option, as the encoding operation can cause a loss in receiver's estimation accuracy. However, under secrecy constraints, lack of encoding can compromise the security, and a limited error can be caused at the eavesdropper. It is observed from Table 3.1 that the achievable target error levels are around 0.07 or lower for the simulation parameters considered in this chapter; however, larger error values are possible if NIE is applied as illustrated in the examples.

3.4.2 Affine Joint Encoding

In this part, we investigate the affine joint encoding strategy and obtain the optimal precoding matrix $\boldsymbol{P}$ to satisfy certain secrecy constraints. In all the numerical examples, $\boldsymbol{\theta}$ is modeled as $\boldsymbol{\theta} = [\theta_1 \ \theta_2]^T$, and θ_1 and θ_2 are assumed to be independent of each other with $\theta_1, \theta_2 \in [0, 1]$. Also, the channel parameters for the intended receiver are taken to be $h_{r,1} = h_{r,2} = 2$. The precoding matrix is expressed as $\boldsymbol{P} = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix}$. Note that $|p_{11}| + |p_{12}| \leq 1$ and $|p_{21}| + |p_{22}| \leq 1$ should be satisfied to ensure $\beta_1, \beta_2 \in [0, 1]$. The strategies considered in the numerical results are given as follows:

- **Affine Joint Encoding (AJE):** This approach refers to the solution of the optimization problem in (3.26).
- **Nonlinear Individual Encoding (NIE):** This approach refers to the solution of the optimization problem in (3.14).
- **Affine Individual Encoding (AIE):** This is a simplified version of the AJE approach. In particular, precoding matrix $\boldsymbol{P}$ has the form of $\boldsymbol{P} = \boldsymbol{W}_1\boldsymbol{W}_2$, where $\boldsymbol{W}_1 = \text{diag}\{w_1, w_2, \ldots, w_N\}$ is a diagonal matrix and $\boldsymbol{W}_2$ is a permutation matrix. The AIE approach can further be grouped as follows:
  1. **AIE without permutation:** This refers to special case with $\boldsymbol{W}_2 = \mathbf{I}$. For $N = 2$, we assume $p_{12} = p_{21} = 0$.
  2. **AIE with permutation:** This refers to the scenario with $\boldsymbol{W}_2 \neq \mathbf{I}$. For $N = 2$, we assume $p_{11} = p_{22} = 0$.

We provide five different examples to investigate the affine joint encoding strategy numerically. In the examples, different values for eavesdropper's fading coefficients and prior distributions for θ_1 and θ_2 are used in order to show the advantages and disadvantages of certain encoding strategies over each other in terms of their performance and to corroborate the theoretical results provided in the

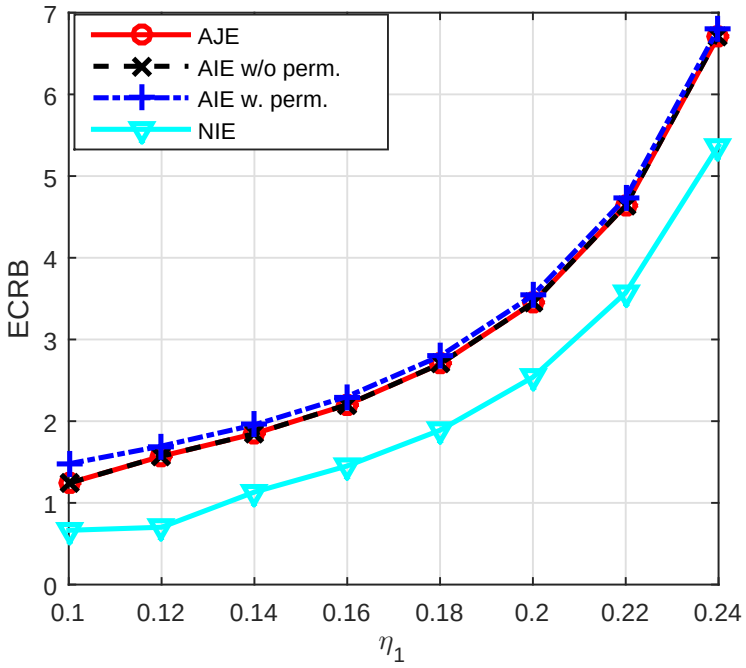

Figure 3.7: Total ECRB versus η_1 for different approaches.

chapter. For the first four examples, the channel noise for the eavesdropper and the intended receiver is taken to be zero-mean Gaussian random variables with independent components of unit variance, i.e., $\mathbf{\Sigma}_e = \mathbf{\Sigma}_r = \mathbf{I}$. In the first example, θ_1 and θ_2 are assumed to be uniformly distributed and the secrecy target for the second parameter, η_2, is set to be 0.15. Also, the eavesdropper fading coefficients are taken as $h_{e,1} = 1.2$ and $h_{e,2} = 1.5$. In Fig. 3.7, the total optimal ECRB values for θ_1 and θ_2 are plotted for various η_1 values. It is observed that NIE provides improved performance compared to the affine encoding options for this scenario. Also, the optimal AJE solution is the same as the optimal AIE without permutations and they perform slightly better than AIE with permutations.

For the second example, we investigate affine encoding strategies in more detail. The simulation parameters are the same as the first example except that the distribution of θ_2 is taken to be $w(\theta_2) = 2\theta_2$ and $w(\theta_2) = 7\theta_2^6$. The secrecy target for the second parameter, η_2, is set to 0.15. In Fig. 3.8, the total optimal ECRB values for θ_1 and θ_2 versus η_1 are plotted for various affine encoding strategies. For AIE with and without encoding strategies, we also investigate the case in which

the coefficients of the matrix are restricted to be positive and this is illustrated in the legend of Fig. 3.8 with (+) next to the name of the corresponding strategy, e.g., AIE w/o perm. (+). When $w(\theta_2) = 7\theta_2^6$, the solutions for the optimal AJE, AIE with permutation and AIE with permutation with positive coefficients are the same and yield the best performance, whereas AIE without permutation with positive coefficients gives the worst performance. AIE without permutation provides a moderate performance except for $\eta_1 < 0.11$, where it also provides the optimal performance. When $w(\theta_2) = 2\theta_2$, AIE with permutation and AIE with permutation with positive coefficients have the same performance, and they perform better than AIE without permutation when $\eta_1 > 0.111$; however, AIE without permutation is better when $\eta_1 < 0.111$. The optimal AJE solution achieves the minimum of these three strategies at all η_1 values. AIE without permutation with positive coefficients yields the worst performance in this case, as well. Note that Corollary 1 and Lemma 3 can be applied in this example for AIE with permutation strategy. As $E(\theta_1) = 1/2$, and eavesdropper's noise is white, Corollary 1 and Lemma 3 imply together that for the AIE with permutation strategy, the matrix elements can be restricted to be positive without loss of generality. Therefore, it is not a coincidence that AIE with permutation and AIE with permutation with positive coefficients yield the same performance in this example.

For the third example, θ_1 is assumed to be uniformly distributed and the distribution of θ_2 is taken to be $w(\theta_2) = 4\theta_2^3$. The secrecy targets for both parameters are set to 0.15. In Fig. 3.9, the total optimal ECRB values for θ_1 and θ_2 are plotted for various $h_{e,1}$ values when $h_{e,2} = 1.5$. It is observed that the performance of AIE with permutation and AIE with permutation with positive coefficients are the same as $E(\theta_1) = 1/2$ for this example, as well. Their performance stays constant as $h_{e,1}$ increases. The performance of AIE without permutation is initially worse than that of AIE with permutation; however, it improves as $h_{e,1}$ increases and performs better when $h_{e,1} > 2.57$. AIE without permutation with positive coefficients yields the worst performance, and its performance gets even worse as $h_{e,1}$ increases. The different responses of the strategies to the increase of $h_{e,1}$ are due to the fact that the structure of $\mathbf{\Sigma}_{err}$ varies as the encoding strategy changes.

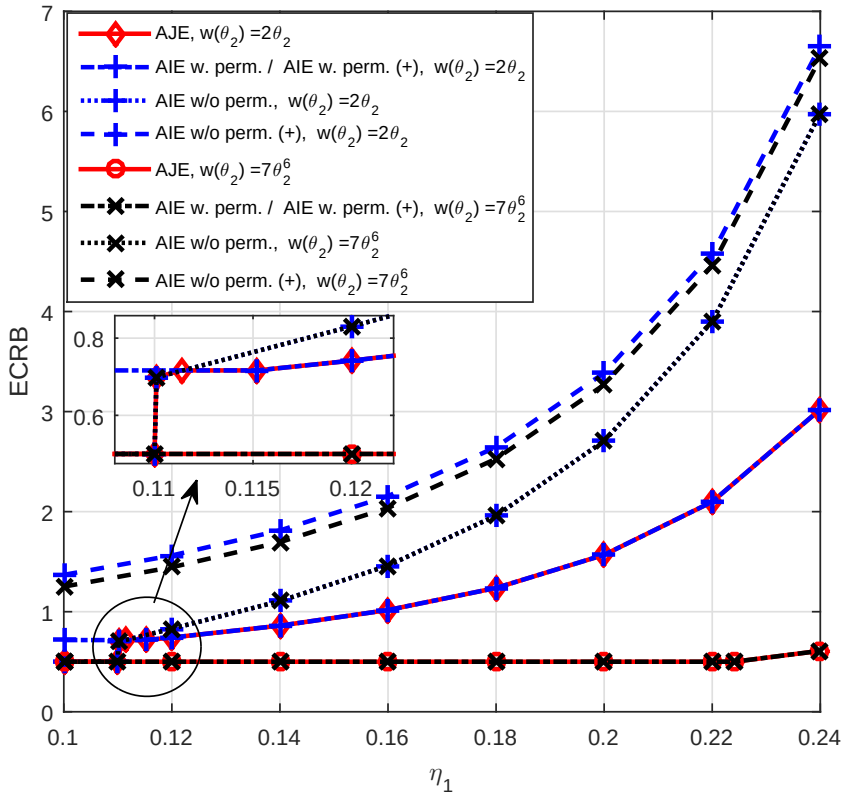

Figure 3.8: Total ECRB versus η_1 for different approaches.

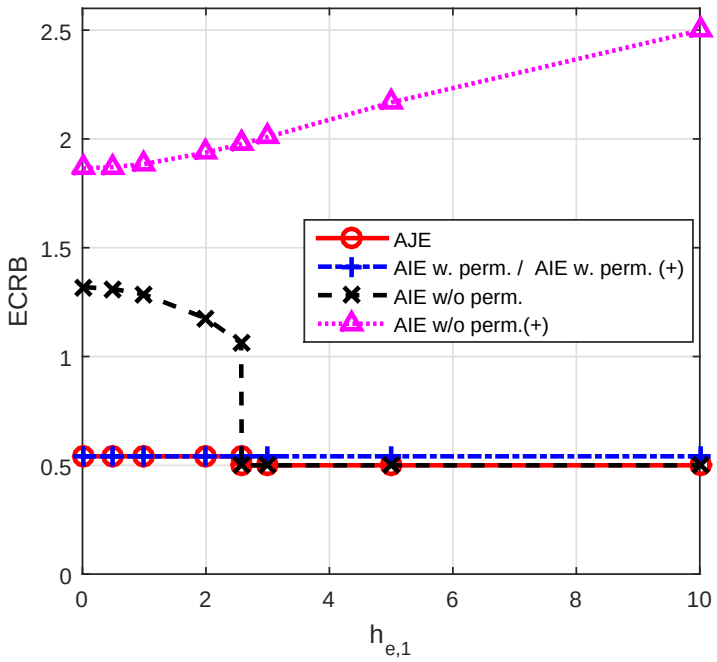

Figure 3.9: Total ECRB versus $h_{e,1}$ for different approaches.

The optimal AJE solution is the same as AIE with permutation when $h_{e,1} < 2.57$ and it is same as AIE without permutation when $h_{e,1} \geq 2.57$.

For the fourth example, the distribution of θ_1 is taken to be $w(\theta_1) = 2\theta_1$ and the distribution of θ_2 is given by $w(\theta_2) = 4\theta_2^3$. The secrecy target for the second parameter, η_2, is set to 0.2. In Fig. 3.10, the total optimal ECRB values for θ_1 and θ_2 are plotted for various η_1 values. It is observed that when $\eta_1 < 0.225$, the best performance is obtained by employing NIE; however, after $\eta_1 > 0.225$, the optimal AJE solution, which has the same performance as AIE with permutation, starts to yield the best performance. This shows that the simple flip and scale approach may be better than the individual nonlinear encoding function strategy in certain scenarios. AIE without permutation performs slightly worse than NIE. AIE with/without permutation with positive coefficients do not achieve a good performance in this scenario. As the conditions given in Corollary 1 are no longer satisfied, there is a significant performance gap between the optimal AIE solutions and the AIE solutions which are restricted to positive coefficients.

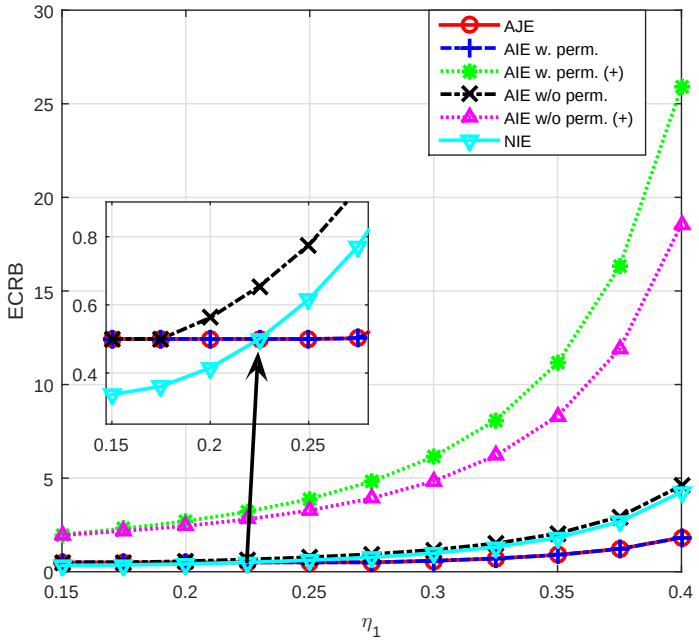

Figure 3.10: Total ECRB versus η_1 for different approaches.

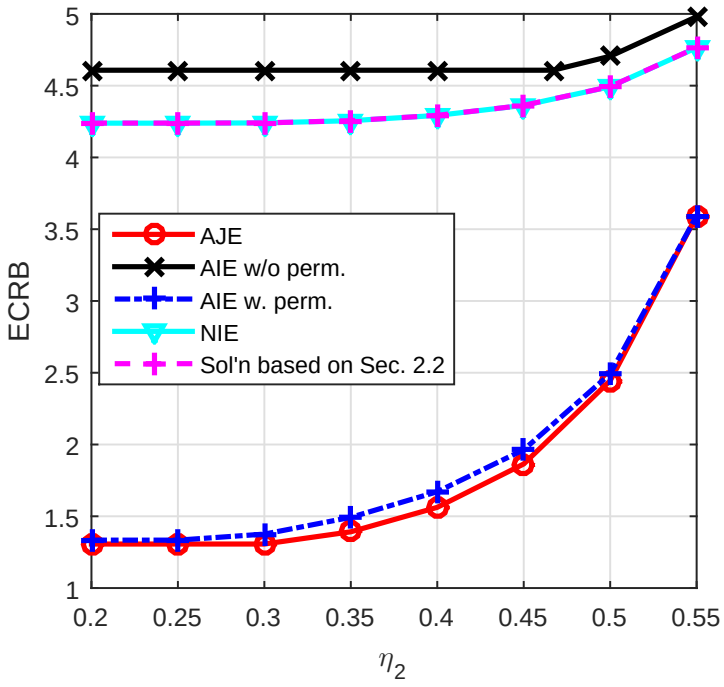

Figure 3.11: Total ECRB versus η_2 for different approaches.

In all the four examples, we have observed that the optimal AJE solution has the form of one of the AIE solutions. However, this does not have to be the case in all scenarios and the fifth example provides such an example. In this example, eavesdropper's fading coefficients are taken as $h_{e,1} = 0.8$ and $h_{e,2} = 1.25$. The channel noise for the eavesdropper is modeled as zero-mean multivariate Gaussian random variable with the covariance matrix $\mathbf{\Sigma}_e = \begin{bmatrix} \sigma_{e,1}^2 & \rho_e \\ \rho_e & \sigma_{e,2}^2 \end{bmatrix}$, where $\sigma_{e,1}^2 = \sigma_{e,2}^2 = 1$ and $\rho_e = -0.5$ and the channel noise for the eavesdropper is also modeled as zero-mean multivariate Gaussian random variable with the covariance matrix $\mathbf{\Sigma}_r = \begin{bmatrix} \sigma_{r,1}^2 & \rho_r \\ \rho_r & \sigma_{r,2}^2 \end{bmatrix}$, where $\sigma_{r,1}^2 = \sigma_{r,2}^2 = 1$ and $\rho_r = 0.7$. The distribution of θ_1 is taken to be $w(\theta_1) = 2\theta_1$ and the distribution of θ_2 is given by $w(\theta_2) = 5\theta_2^4$. The secrecy target for the first parameter, η_1, is set to be 0.4, and the total optimal ECRB values for θ_1 and θ_2 are plotted for various η_2 values. In Fig. 3.11, it is observed that the optimal AJE solution is better than both the optimal AIE with and without permutation solutions. For example, when $\eta_2 = 0.35$, the optimal precoding matrix for the AIE with permutation solution is $\begin{bmatrix} 0 & -0.4787 \\ -0.7807 & 0 \end{bmatrix}$, yielding an objective value of 1.5012. On the other hand, the optimal precoding matrix for the AJE strategy is $\begin{bmatrix} 0 & -0.48 \\ -0.6578 & -0.2439 \end{bmatrix}$, yielding an objective value of 1.4008.[7] Therefore, it is possible that joint encoding of parameters can outperform individual encoding depending on the channel and parameter statistics. It is also observed that NIE and the solution based on Section 2.2 have almost the same performance, and even though they are better than AIE without permutation, they perform worse than AIE with permutation and AJE. This implies that, in this particular scenario, the main source of performance improvement is to exploit the fact that there are multiple elements in the vector by shuffling the order of the elements or even jointly encoding them rather than individual encoding via a nonlinear function. Therefore, there might be cases in which it is not very critical to take the correlation in noise components into account in NIE, as the performance improvement can be negligible.

[7]The corresponding optimal $\boldsymbol{r}$ values for the AIE with permutation solution and the AJE solution can be found as $\boldsymbol{r}^T = [0.4787\ 0.7807]$ and $\boldsymbol{r}^T = [0.48\ 0.9017]$ respectively.

Parameters	η_1	η_2
$h_{e,1}=1.2, h_{e,2}=1.5, w(\theta_1)=1, w(\theta_2)=1$	0.0744	0.0702
$h_{e,1}=1.2, h_{e,2}=1.5, w(\theta_1)=1, w(\theta_2)=2\theta_2$	0.0744	0.0494
$h_{e,1}=1.2, h_{e,2}=1.5, w(\theta_1)=1, w(\theta_2)=7\theta_2^6$	0.0744	0.0118
$h_{e,1}=1, h_{e,2}=1.5, w(\theta_1)=1, w(\theta_2)=4\theta_2^3$	0.0769	0.0252
$h_{e,1}=3, h_{e,2}=1.5, w(\theta_1)=1, w(\theta_2)=4\theta_2^3$	0.0476	0.0252
$h_{e,1}=5, h_{e,2}=1.5, w(\theta_1)=1, w(\theta_2)=4\theta_2^3$	0.0270	0.0252
$h_{e,1}=10, h_{e,2}=1.5, w(\theta_1)=1, w(\theta_2)=4\theta_2^3$	0.0089	0.0252
$h_{e,1}=1.2, h_{e,2}=1.5, w(\theta_1)=2\theta_1, w(\theta_2)=4\theta_2^3$	0.0514	0.0252
The parameters of Fig. 3.11	0.0531	0.0191

Table 3.2: Maximum secrecy target level values for θ_1 and θ_2 when $\boldsymbol{P}=\mathbf{I}$ and $\boldsymbol{r}=\mathbf{0}$

.

The maximum secrecy target levels with no encoding are provided for this encoding scheme as well; that is, $\boldsymbol{P}=\mathbf{I}$ and $\boldsymbol{r}=\mathbf{0}$ in Table 3.2 for all the considered scenarios. It is observed that the achievable secrecy levels are much lower than those of the AJE scheme. It is also interesting to note that as $h_{e,1}$ increases, the secrecy levels decrease in Table 3.2. This is because of the fact that the channel of the eavesdropper gets better and the error performance improves when the original parameter is transmitted. Such an issue does not occur if the optimal AJE is applied, and this can even be turned into an advantage according to Fig. 3.9 due to the secret encoder. Also, for the parameters of Fig. 3.11, the maximum error levels for θ_1 and θ_2 are 0.0531 and 0.0191, respectively; however, the optimal AJE can reach $\eta_1=0.4$ and $\eta_2=0.55$ (and possibly more) according to the fifth example. This shows the clear advantage of the proposed schemes as compared to not utilizing any encoder in the presence of an eavesdropper.

3.4.3 Computational Complexity

One of the main factors determining the computational complexity of the proposed algorithms is the dimension of the space in which the search is performed. When we use the piecewise linear approximation (PWL) method to obtain the optimal solution for nonlinear individual encoding, the search is performed over MN variables as described in Section 3.4.1. As M increases, lower ECRB values

can be obtained. However, it increases the search dimension and the complexity. For affine joint encoding, the original optimization problem in (3.26) requires a search over $\boldsymbol{P}$ and $\boldsymbol{r}$, yielding a search space over $N^2 + N$ variables. However, it is shown in Lemma 2 that it is enough to calculate $\boldsymbol{P}$ for optimal encoding reducing the space to N^2 variables.

Another important factor related to the computational complexity of encoder optimization is the number of multiplications at the calculation of the cost and objective functions for a given candidate encoder. For NIE, both the objective and cost functions require a calculation of an N dimensional integral. Let X denote the terms in the Riemann sum for a given step size. Then, the objective function requires $\mathcal{O}(NX)$ multiplications. To calculate $\boldsymbol{\Sigma}_{err}$, each of $\boldsymbol{\Sigma}_{\boldsymbol{\beta}}$ and $\boldsymbol{\Sigma}_{\boldsymbol{\beta},\boldsymbol{\theta}}$ needs $\mathcal{O}(N^2X)$ and $E(\boldsymbol{\beta})$ needs $\mathcal{O}(NX)$ calculations. Then, the overall complexity to calculate (3.11) becomes $\mathcal{O}(N^2X) + \mathcal{O}(N^3)$. For AJE, the complexity of calculating the cost function and the objective function are both $\mathcal{O}(N^3)$. Therefore, AJE has lower computational complexity especially if N is not very large. However, if N is large, then the optimal matrix calculation can become more costly than the NIE algorithm. Note that AJE is a type of a precoding based encoding strategy; hence it has a comparable complexity to the beamforming strategies in the literature, which are employed in different problems.

As a special case of AJE, AIE is also considered in the numerical examples. If AIE without permutation is employed, the search space reduces to N from N^2, and the complexity of the cost and objective function calculations also decreases relatively. For AIE with permutation, the search space is $N + 1$, where the extra variable indicates the permutation order. Note that when N increases, the possible values for the permutation order increases very quickly. However, it is always possible to prune the size of this set to a practical maximum size, and to choose the permutation order from it.

3.4.4 General Observations

We have investigated the optimal encoding of multiple parameters for secure communication for the two proposed practical encoding approaches. For the NIE scheme, it is observed that as the correlation between eavesdropper's noise components increases, the total ECRB cost decreases for a given target secrecy level implying that such a correlation is useful for the parameter encoding task. It is also observed that the encoding function is in either the variance minimizing or maximizing mode depending on the channel quality and the correlation values of the parameter. In the second part, the affine joint and individual encoding schemes are compared with each other for various parameter distributions. It is observed that in many scenarios, the solution of the AJE scheme is in the form of the AIE solution, which can be with or without permutations. This implies that individually encoding each parameter can be good enough to solve the optimization problem in most cases. However, it is important to emphasize that this is not a theory as it is possible to find counter examples. Also, when AJE and NIE are compared to each other, it is observed that one can have better performance than the other depending on the scenario. This means that in certain scenarios, simple permutation or/and scaling of the parameters can be the effective security solution and in some cases, using a nonlinear function without utilizing any permutation brings more benefits.

We note that the main goal behind the encoding operation is to achieve the desired secrecy levels by also providing a certain estimation quality at the receiver. It is observed that both of the proposed approaches have the ability to achieve large estimation errors at the eavesdropper, which could not have been possible if there were no encoding utilized. Another important contribution of this chapter is that we provide useful theoretical simplifications (and sufficient conditions to apply them) to obtain optimal encoders in practice and they have been used in all the examples. Proposition 1 is utilized when $\rho = 0$ and decoupled optimization problems are solved. Similarly, Lemma 1 is also applied and even though the encoding functions are obtained jointly for $\rho > 0$, they are searched over decreasing functions as it can be observed in Figs. 3.3 and 3.5. For AJE,

Lemma 2 is utilized to solve the optimization problem and the search is restricted to the optimal precoding matrix as the constant term can be found theoretically for a given encoder. It is observed that even though signed permutation matrices are optimal when there is no secrecy constraint according to Proposition 2, they are not optimal under secrecy requirements in general and either joint encoding and/or scaling of the parameters is required. Finally, Corollary 1 and Lemma 3 are utilized to show that in certain scenarios, matrix coefficients can be restricted to be positive as can be observed in Figs. 3.8 and 3.9.

3.5 Concluding Remarks

In this chapter, optimal encoding of multiple parameters has been investigated in the presence of an eavesdropper. An optimization problem has been proposed with an objective to minimize the ECRB at the intended receiver while satisfying the MSE targets at the eavesdropper. Two practical encoding schemes, i.e., NIE and AJE, have been proposed. It has been observed that both schemes are able to create large estimation errors at the eavesdropper, which is not possible when no encoding is applied, and they can be employed as a security measure. The performance of both schemes has also been compared to each other and it has been observed that one can have better performance than the other depending on the scenario. Another observation is that the optimal encoding function for NIE is in either the variance minimizing or maximizing mode, and in many scenarios individually encoding the parameters with an affine function (with or without permutation of the parameters) has as good performance as that of AJE. Also, theoretical results derived in the chapter prove useful in simplifying the optimal encoding problem.

Finally, we note that it is entirely possible that two strategies proposed in this chapter can be combined to further optimize the encoding function. For example, the first encoding block can perform nonlinear individual encoding and the second encoding block can perform affine joint encoding to the output of the first block. In that case, both the nonlinear individual encoding functions and the precoding

matrix should be optimized jointly and despite the increase in computational complexity, the performance can further be improved.

Chapter 4

Estimation Theoretic Secure Communication via Encoder Randomization

In this chapter, optimal encoding of a scalar random parameter is investigated in the presence of an eavesdropper, where the encoder at the transmitter is allowed to use a randomized mapping between two one-to-one and continuous functions and the eavesdropper is fully aware of the encoding strategy at the transmitter [45]. The main contributions of this chapter can be summarized as follows:

- The problem of parameter encoding via encoder randomization is analyzed to ensure estimation theoretic secure communication under the assumption that the encoding scheme is available to the eavesdropper.
- For small numbers of observations, a closed form expression for the MSE of the LMMSE estimator is derived for both the receiver and the eavesdropper for the considered transmission and encoding scheme. The optimization problems to minimize the MSE at the intended receiver for a given secrecy target at the eavesdropper and to maximize the MSE at the eavesdropper for a given estimation accuracy limit at the receiver are formulated.

The relationship between the solutions of those problems is characterized. An optimal solution of the optimization problems is obtained theoretically when the channel of the eavesdropper is noisier than the channel of the intended receiver. It is also shown that a simple deterministic affine function can attain the optimal value. For the case of affine functions, the monotonicity behavior of the MSE is obtained with respect to the randomization probability when the encoding functions are fixed.

- For large numbers of observations, the optimization problems to minimize the ECRB at the intended receiver for a given secrecy target at the eavesdropper and to maximize the ECRB at the eavesdropper for a given estimation accuracy limit at the receiver are formulated. The optimizations problems are theoretically solved when only deterministic encoding is considered. It is also shown that under symmetric mapping, the ECRB is maximized when the randomization probability is 1/2. Also, the monotonicity behavior of the ECRB is obtained with respect to the randomization probability when the encoding functions are fixed for this case, as well.

The rest of the chapter is organized as follows: The system setup is introduced in Section 4.1. The optimization problems are formulated and analyzed for small and large numbers of observations in Section 4.2 and Section 4.3, respectively. The numerical results are presented in Section 4.4, and the concluding remarks are given in Section 4.5.

4.1 System Setup

Consider the transmission of a scalar parameter $\theta \in \Lambda$ to an intended receiver in the presence of an eavesdropper who wants to estimate parameter θ. Both the intended receiver and the eavesdropper obtain n-dimensional observations over their respective additive noise channels. The aim is to achieve accurate estimation of the parameter at the intended receiver while keeping the estimation error at the eavesdropper above a certain level; or, alternatively, to ensure that the estimation

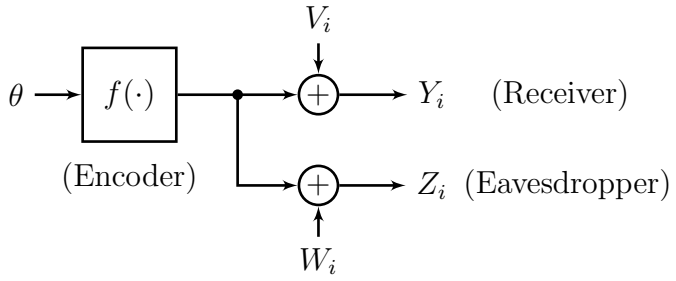

Figure 4.1: System model for the parameter encoding problem.

error at the eavesdropper is as large as possible while satisfying an estimation constraint at the intended receiver. To that aim, the parameter is encoded by an encoding function $f : \Lambda \to \Gamma$. Let $f(\theta)$ denote the encoded version of the parameter. Hence, the ith observation at the intended receiver can be written as

$$Y_i = f(\theta) + V_i \ , \quad i = 1, 2, \ldots, n. \tag{4.1}$$

where the noise V_i is modeled as a zero-mean Gaussian random variable with variance σ_V^2, and V_i and θ are assumed to be independent [18]. On the other hand, the ith observation at the eavesdropper is

$$Z_i = f(\theta) + W_i \ , \quad i = 1, 2, \ldots, n. \tag{4.2}$$

where W_i is zero-mean Gaussian noise with variance σ_W^2, which is independent of θ for $i = 1, 2, \ldots, n$. Also, the prior information on parameter θ is represented by a probability density function (PDF) denoted by $p_\theta(\theta)$ for $\theta \in \Lambda$. The signal model in (4.1) and (4.2) can also be employed for flat-fading channels assuming perfect channel estimation and appropriate equalization [69]. The intended receiver aims to estimate parameter θ based on observations $\mathbf{Y} \triangleq [Y_1, Y_2, \ldots, Y_n]^T$ whereas the eavesdropper uses observations $\mathbf{Z} \triangleq [Z_1, Z_2, \ldots, Z_n]^T$ for estimating θ. The system model is illustrated in Fig. 4.1.

The considered system model is also known as the Gaussian wiretap channel [12], [18], and has been studied extensively via information theoretical tools, as mentioned in Section 1.2. In that framework, it is assumed that the eavesdropper knows the codewords (mapping) in the encoder and has unlimited resources/time

for computation. Therefore, the encoder applies a stochastic mapping from messages to codewords to ensure that the message can be kept unknown to the eavesdropper by exploiting the degradedness of eavesdropper's channel while still being able to transmit the message to the intended receiver at a certain rate.[1] Motivated from such a setting, the following assumptions are made for the rest of this chapter:

- The encoding function at the transmitter is fully available to the eavesdropper and the receiver. Therefore, it is possible that both the eavesdropper and the receiver can utilize optimal estimators according to a certain metric.

- To enhance security, stochastic encoding is employed and the encoder is modeled to perform the following mapping:

$$f(\theta) = \begin{cases} f_1(\theta), & \text{with probability } \gamma \\ f_2(\theta), & \text{with probability } 1-\gamma \end{cases} \tag{4.3}$$

  where $f_k(\theta) : \Lambda \to \Gamma$ is a continuous and one-to-one function for $k = 1, 2$ and $\gamma \in [0, 1]$. [2]

- Each observation is corrupted by independent and identically distributed noise components. Therefore, based on this and the previous assumption, the conditional PDF of the n observations at the receiver given θ, denoted by $p(\mathbf{y}|\theta)$, can be expressed as

$$p(\mathbf{y}|\theta) = \prod_{i=1}^{n} p(y_i|\theta) \tag{4.4}$$

  where $\mathbf{y} \triangleq [y_1, y_2, \ldots, y_n]^T$, $p(y_i|\theta) = \gamma\, p_V(y_i - f_1(\theta)) + (1-\gamma)\, p_V(y_i - f_2(\theta))$ and $p_V(x) = \frac{1}{\sqrt{2\pi}\sigma_V} \exp\{-\frac{x^2}{2\sigma_V^2}\}$. Similarly, the conditional PDF of the n

[1] Unlike the classical Gaussian wiretap channel [12], [18], we consider a scenario in which the channel of the eavesdropper is not necessarily worse than that of the intended receiver.

[2] The stochastic encoder in (4.3) both facilitates practical implementations and allows for theoretical investigations. Note that it can also be represented as $f(\theta) = f_{2-X}(\theta)$, where X is a Bernoulli random variable with parameter γ and X is statistically independent of all other variables.

observations at the eavesdropper given θ, $p(\mathbf{z}|\theta)$, can be stated as

$$p(\mathbf{z}|\theta) = \prod_{i=1}^{n} p(z_i|\theta) \tag{4.5}$$

where $\mathbf{z} \triangleq [z_1, z_2, \ldots, z_n]^T$, $p(z_i|\theta) = \gamma\, p_W(z_i - f_1(\theta)) + (1-\gamma)\, p_W(z_i - f_2(\theta))$ and $p_W(x) = \frac{1}{\sqrt{2\pi}\sigma_W} \exp\{-\frac{x^2}{2\sigma_W^2}\}$.

In this setting, the encoder should be designed in such a way that the estimation errors at the eavesdropper or, alternatively, at the intended receiver satisfy the constraints. It is noted that the secrecy capacity in information theory is an asymptotic metric and assumes that $n \to \infty$. In practice, it is also important to investigate how much secrecy can be achieved in the finite regime with a small number of observations. For example, [70] provides new achievability results and converse bounds for the maximal secret communication rate of wiretap channels for a given finite blocklength n. Similarly, we focus on the optimal encoding design in the non-asymptotic region for both small and large numbers of observations in this work.

It is known that the optimal estimator for Bayesian parameter estimation in terms of the MSE metric is the MMSE estimator. However, in most scenarios, the MSE of the optimal MMSE estimator does not have a closed form expression. Therefore, even though the encoding operation can be performed with such an approach by using numerical methods, it does not allow theoretical investigations for achieving intuitive understanding of the parameter encoding problem. It is known that for a large number of observations, the MSE of the MMSE estimator converges to the ECRB [48], and for a small number of observations, the MSE of the LMMSE estimator is a close approximation to the optimal MMSE (see Figs. 4.2-4.4 for an illustration). (Note that the LMMSE estimator would actually be the optimal MMSE estimator if the parameter of interest and the observations were jointly Gaussian random variables.) Therefore, instead of the optimal MMSE, the ECRB and the LMMSE estimator will be considered in the rest of the chapter.

Remark 1: The main reason for employing the MSE metric in both the receiver and the eavesdropper is that we focus on a parameter estimation problem in the Bayesian setting in the presence of an eavesdropper and the MSE metric is widely used in practice with or without secrecy concerns in such problems. For example, estimation theoretic secrecy based on the MSE metric has been considered in various channel scenarios such as Gaussian interference channel [29], multiuser MIMO broadcast channel [37], sensor network systems with eavesdroppers [31] and MIMO Gaussian wiretap channel [36]. In addition to parameter estimation problems, the MSE metric is also utilized to design practical and implementable methods to degrade performance of eavesdroppers for enhancing security as an additional layer.

4.2 Small Number of Observations

In this section, it is assumed that a small number of observations are available to the intended receiver and the eavesdropper to estimate θ. As motivated in the previous section, both the eavesdropper and the intended receiver are modeled to employ LMMSE estimators for a given number of observations n.

4.2.1 Generic Encoding Functions

First, generic encoding functions are considered at the transmitter. To that end, as motivated in Section 2.2, the parameter space and the intrinsic constraints on the functions $f_1(\theta)$ and $f_2(\theta)$ are specified as follows:

- $\theta \in \Lambda = [a, b]$.
- $f_k(\theta) \in [a, b]$ for $k = 1, 2$.
- $f_1(\theta)$ and $f_2(\theta)$ are continuous and one-to-one functions.

The LMMSE estimator at the intended receiver can explicitly be written for given observations $\mathbf{y}$ as

$$\hat{\theta}_r = E(\theta) + \mathbf{\Sigma}_{\theta,\mathbf{Y}}\mathbf{\Sigma}_{\mathbf{Y}}^{-1}(\mathbf{y} - E(\mathbf{Y})), \tag{4.6}$$

and the corresponding MSE can be obtained as

$$\epsilon_r = MSE = Var(\theta) - \mathbf{\Sigma}_{\theta,\mathbf{Y}}\mathbf{\Sigma}_{\mathbf{Y}}^{-1}\mathbf{\Sigma}_{\theta,\mathbf{Y}}^{T}. \tag{4.7}$$

where $\mathbf{\Sigma}_{\theta,\mathbf{Y}} = [Cov(\theta, Y_1), Cov(\theta, Y_2) \dots Cov(\theta, Y_n)]$ and $\mathbf{\Sigma}_{\mathbf{Y}} = E\Big(\big(\mathbf{Y} - E(\mathbf{Y})\big)\big(\mathbf{Y} - E(\mathbf{Y})\big)^T\Big)$. Similarly, the MSE of the LMMSE estimator at the eavesdropper, ϵ_e, can be obtained for given observations $\mathbf{z}$ by using $\mathbf{Z}$ instead of $\mathbf{Y}$ in (4.7). Based on these MSE expressions, the optimization problems can be proposed as follows:

$$\min_{\gamma, f_1(\theta), f_2(\theta)} \epsilon_r \;\; \text{s.t.} \;\; \epsilon_e \geq \alpha_1 \tag{4.8}$$

and

$$\max_{\gamma, f_1(\theta), f_2(\theta)} \epsilon_e \;\; \text{s.t.} \;\; \epsilon_r \leq \alpha_2 \tag{4.9}$$

where α_1 and α_2 denote, respectively, the secrecy target for the first problem and the estimation accuracy (error) limit at the intended receiver for the second problem. The following proposition provides a closed form expression for the MSE of the LMMSE estimator at the intended receiver.

Proposition 1: *The MSE (ϵ_r) of the LMMSE estimator at the intended receiver for the encoding model specified in* (4.3) *with given $f_1(\theta)$, $f_2(\theta)$ and γ is*

$$\epsilon_r = Var(\theta) - \frac{n\,\left(\gamma\, c_1 + (1-\gamma)\, c_2\right)^2}{(n-1)x + \tau - nt} \tag{4.10}$$

where

$$
\begin{aligned}
x &\triangleq \gamma^2\, r_1 + (1-\gamma)^2 r_2 + 2\,\gamma\,(1-\gamma) E(f_1(\theta)\, f_2(\theta)) \\
\tau &\triangleq \gamma\, r_1 + (1-\gamma)\, r_2 + \sigma_V^2 \\
t &\triangleq (\gamma\, m_1 + (1-\gamma)\, m_2)^2
\end{aligned} \tag{4.11}
$$

with $m_i = E(f_i(\theta))$, $r_i = E(f_i(\theta)^2)$ *and* $c_i = Cov(f_i(\theta), \theta)$ *for* $i = 1, 2$.

Proof: Note that $\mathbf{\Sigma_Y} = E(\mathbf{Y}\,\mathbf{Y}^T) - E(\mathbf{Y})E(\mathbf{Y})^T$. Also, $E(Y_k|\theta) = \gamma\, f_1(\theta) + (1-\gamma)\, f_2(\theta)$. Then, $E(Y_k) = E\big(E(Y_k|\theta)\big) = \gamma\, m_1 + (1-\gamma)\, m_2$ for $k = 1, 2, \ldots, n$. Therefore, $E(\mathbf{Y}) = (\gamma\, m_1 + (1-\gamma)\, m_2)\mathbf{1}$, where $\mathbf{1}$ denotes the $n \times 1$ column vector of ones. Thus, $E(\mathbf{Y})E(\mathbf{Y})^T = (\gamma\, m_1 + (1-\gamma)\, m_2)^2 \mathbf{1}\mathbf{1}^T = t\mathbf{1}\mathbf{1}^T$.

In addition, $E(Y_k^2|\theta) = \gamma\,(f_1(\theta)^2 + \sigma_V^2) + (1-\gamma)\,(f_2(\theta)^2 + \sigma_V^2)$; hence, $E(Y_k^2) = \gamma\, r_1 + (1-\gamma)\, r_2 + \sigma_V^2 = \tau$ for $k = 1, 2, \ldots, n$. Similarly, $E(Y_j Y_k|\theta) = E(Y_j|\theta)E(Y_k|\theta) = (\gamma\, f_1(\theta) + (1-\gamma)\, f_2(\theta))^2$. Then, $E(Y_j Y_k) = \gamma^2\, r_1 + (1-\gamma)^2\, r_2 + 2\,\gamma\,(1-\gamma)\, E(f_1(\theta)\, f_2(\theta)) = x$ for $j, k = 1, 2, \ldots, n$ and $j \neq k$. Overall, the value of the diagonal elements of $\mathbf{\Sigma_Y}$ is $\tau - t$ and the rest of the elements are $x - t$.

Furthermore, $\mathbf{\Sigma}_{\theta,\mathbf{Y}} = Cov(\theta, Y_k)\mathbf{1}^T$ and $Cov(\theta, Y_k) = E(\theta Y_k) - E(\theta)E(Y_k)$. Note that $E(\theta\, Y_k) = E\big(E(\theta Y_k|\theta)\big) = E\big(\theta\, E(Y_k|\theta)\big) = \gamma\, E(\theta f_1(\theta)) + (1-\gamma)E(\theta f_2(\theta))$. Then, $Cov(\theta, Y_k) = \gamma\Big(E(\theta f_1(\theta)) - E(\theta)E(f_1(\theta))\Big) + (1-\gamma)\Big(E(\theta f_2(\theta)) - E(\theta)E(f_2(\theta))\Big) = \gamma\, c_1 + (1-\gamma)c_2$. Therefore, the MSE becomes $Var(\theta) - \mathbf{\Sigma}_{\theta,\mathbf{Y}}\mathbf{\Sigma}_{\mathbf{Y}}^{-1}\mathbf{\Sigma}_{\theta,\mathbf{Y}}^T = Var(\theta) - (\gamma\, c_1 + (1-\gamma)\, c_2)^2\, \mathbf{1}^T\mathbf{\Sigma}_{\mathbf{Y}}^{-1}\mathbf{1}$. Note that the sum of the elements in each row of $\mathbf{\Sigma_Y}$ is the same; therefore, $\mathbf{\Sigma_Y}\mathbf{1} = \lambda\mathbf{1}$, where $\lambda = (n-1)x + \tau - nt$. As λ is an eigenvalue of $\mathbf{\Sigma_Y}$ with a corresponding eigenvector $\mathbf{1}$, $\mathbf{\Sigma}_{\mathbf{Y}}^{-1}\mathbf{1} = (1/\lambda)\mathbf{1}$ holds. Then, $\mathbf{1}^T\mathbf{\Sigma}_{\mathbf{Y}}^{-1}\mathbf{1} = (1/\lambda)\,\mathbf{1}^T\mathbf{1} = n/\lambda$. Hence, the MSE becomes $Var(\theta) - (\gamma\, c_1 + (1-\gamma)\, c_2)^2\, n/\lambda$, and inserting the value of $\lambda = (n-1)x + \tau - nt$ concludes the proof. ■

Proposition 1 provides a tool to calculate the MSE for any given prior information $p_\theta(\theta)$, encoding scheme ($f_1(\theta)$, $f_2(\theta)$, γ) and number of observations n. Note that Proposition 1 can similarly be derived for the eavesdropper by using

σ_W^2 instead of σ_V^2 whenever necessary. It can be observed that the MSE in (4.10) increases when the noise variance increases; therefore, $\epsilon_r < \epsilon_e$ when $\sigma_V^2 < \sigma_W^2$.

It is noted that the optimization problems in (4.8) and (4.9) are related such that the expressions for ϵ_r and ϵ_e differ only in the noise variance terms. Therefore, it is possible to find a relationship between the solutions of (4.8) and (4.9), as stated in the following proposition.

Proposition 2: *Suppose that* $\mathcal{S} = \{(\gamma^*, f_1^*, f_2^*)\}$ *is the set of optimal solutions to* (4.8). *Let the optimal value of* (4.8) *be denoted as* ϵ_r^*. *If* α_2 *is set as* $\alpha_2 = \epsilon_r^*$ *in* (4.9), *then the optimal solutions of* (4.9) *satisfy the constraint in* (4.9) *with equality, and* $\epsilon_e^\dagger = \max_{(\gamma, f_1, f_2) \in \mathcal{S}} \epsilon_e$, *where* $\epsilon_e^\dagger$ *is the optimal value of* (4.9). *Similarly, let* $\bar{\mathcal{S}} = \{(\gamma^\dagger, f_1^\dagger, f_2^\dagger)\}$ *denote the set of optimal solutions to* (4.9). *If* $\alpha_1 = \epsilon_e^\dagger$ *in* (4.8), *then the optimal solutions to* (4.8) *satisfy the constraint in* (4.8) *with equality, and* $\epsilon_r^* = \min_{(\gamma, f_1, f_2) \in \bar{\mathcal{S}}} \epsilon_r$.

Proof: We provide a proof only for the first statement as the second one can be shown in a similar fashion. Let the MSEs of the intended receiver and the eavesdropper be denoted, respectively, as $\epsilon_r = T(\gamma, f_1, f_2, \sigma_V^2)$ and $\epsilon_e = T(\gamma, f_1, f_2, \sigma_W^2)$ for given γ, f_1, and f_2. Suppose that $(\gamma^\dagger, f_1^\dagger, f_2^\dagger)$ is an optimal solution to (4.9) with $T(\gamma^\dagger, f_1^\dagger, f_2^\dagger, \sigma_V^2) < \alpha_2 = \epsilon_r^*$. Then, $(\gamma^\dagger, f_1^\dagger, f_2^\dagger)$ cannot be in the feasible set of (4.8) as $\alpha_2 = \min \epsilon_r$ for $\epsilon_e \geq \alpha_1$ in (4.8), implying that $T(\gamma^\dagger, f_1^\dagger, f_2^\dagger, \sigma_W^2) < \alpha_1$. Note that any $(\gamma^*, f_1^*, f_2^*) \in \mathcal{S}$ satisfies $T(\gamma^*, f_1^*, f_2^*, \sigma_W^2) \geq \alpha_1 > T(\gamma^\dagger, f_1^\dagger, f_2^\dagger, \sigma_W^2)$, which shows that $(\gamma^\dagger, f_1^\dagger, f_2^\dagger)$ cannot be an optimal solution to (4.9). Therefore, the optimal solution to (4.9) should satisfy $T(\gamma^\dagger, f_1^\dagger, f_2^\dagger, \sigma_V^2) = \alpha_2 = T(\gamma^*, f_1^*, f_2^*, \sigma_V^2) = \epsilon_r^*$, and it needs to be in $\mathcal{S}$. Hence, the sufficient space to search for the optimal solution of (4.9) reduces to $\mathcal{S}$, and $\epsilon_e^\dagger = \max_{(\gamma, f_1, f_2) \in \mathcal{S}} \epsilon_e$. ■

The following corollaries immediately follow from Proposition 2.

Corollary 1: *If* (γ^*, f_1^*, f_2^*) *is a unique solution to* (4.8) *with the optimal value* ϵ_r^*, *then it is also a unique solution to* (4.9) *for* $\alpha_2 = \epsilon_r^*$.

Corollary 2: *If all the optimal solutions to* (4.8) *satisfy the constraint in* (4.8) *with equality, then the optimal value of* (4.9), $\epsilon_e^\dagger$, *is equal to* α_1 *for* $\alpha_2 = \epsilon_r^*$.

Corollary 3: *If* $(\gamma^\dagger, f_1^\dagger, f_2^\dagger)$ *is a unique solution to* (4.9) *with the optimal value* $\epsilon_e^\dagger$, *then it is also a unique solution to* (4.8) *for* $\alpha_1 = \epsilon_e^\dagger$.

Corollary 4: *If all the optimal solutions to* (4.9) *satisfy the constraint in* (4.9) *with equality, then the optimal value of* (4.8), ϵ_r^*, *is equal to* α_2 *for* $\alpha_1 = \epsilon_e^\dagger$.

As the optimization problems in (4.8) and (4.9) require a search over functions, characterizing the set of optimal solutions in every case may not be possible. However, Proposition 1 provides the required expressions to evaluate the objective and constraint functions for given σ_W^2 and σ_V^2. Based on those expressions, the following proposition provides a closed form expression for an optimal solution to (4.8) and (4.9) when the channel of eavesdropper is noisier than that of the intended receiver; that is, $\sigma_W^2 > \sigma_V^2$.

Proposition 3: *If* $\sigma_W^2 > \sigma_V^2$, *an optimal solution to* (4.8) *is a deterministic affine function, denoted by* $f^*(\theta) = k_1^*\theta + k_2^*$, *where*

$$k_1^* = \pm\sqrt{\frac{\sigma_V^2}{n}\left(\frac{1}{\alpha_1} - \frac{1}{Var(\theta)}\right)} \tag{4.12}$$

and k_2^* *can be anything as long as* $f^*(\theta) \in [a, b]$. *Then, the optimal value of* (4.8) *is*

$$\epsilon_r^* = \frac{\sigma_V^2\, Var(\theta)\, \alpha_1}{\sigma_W^2(Var(\theta) - \alpha_1) + \sigma_V^2\alpha_1}\,. \tag{4.13}$$

Similarly, an optimal solution to (4.9) *is a deterministic affine function,* $f^\dagger(\theta) = k_1^\dagger\theta + k_2^\dagger$, *where*

$$k_1^\dagger = \pm\sqrt{\frac{\sigma_W^2}{n}\left(\frac{1}{\alpha_2} - \frac{1}{Var(\theta)}\right)} \tag{4.14}$$

and $k_2^\dagger$ *can be anything as long as* $f^\dagger(\theta) \in [a, b]$. *Then, the optimal value of* (4.9)

is

$$\epsilon_e^{\dagger} = \frac{\sigma_W^2 \, Var(\theta) \, \alpha_2}{\sigma_V^2 (Var(\theta) - \alpha_2) + \sigma_W^2 \alpha_2} \cdot \tag{4.15}$$

Proof: First, we focus on the optimization problem in (4.9). The denominator of the second term in (4.10) can be rewritten as $n(x-t)+\tau-x$, where $x - t = Var(\gamma f_1(\theta) + (1-\gamma) f_2(\theta))$ and $\tau - x = \gamma\,(1-\gamma) E\left(|f_1(\theta) - f_2(\theta)|^2\right) + \sigma_V^2$. Also, the numerator of the second term in (4.10) can be expressed as $n\,Cov(\gamma f_1(\theta) + (1-\gamma) f_2(\theta), \theta)^2$. Therefore, ϵ_e and ϵ_r become

$$\epsilon_e = Var(\theta) - \frac{n\,Cov(\tilde{f}, \theta)^2}{n\,Var(\tilde{f}) + \gamma\,(1-\gamma) E\left(|f_1(\theta) - f_2(\theta)|^2\right) + \sigma_W^2}$$
$$\epsilon_r = Var(\theta) - \frac{n\,Cov(\tilde{f}, \theta)^2}{n\,Var(\tilde{f}) + \gamma\,(1-\gamma) E\left(|f_1(\theta) - f_2(\theta)|^2\right) + \sigma_V^2}$$

respectively, where $\tilde{f} \triangleq \gamma f_1(\theta) + (1-\gamma) f_2(\theta)$. It is noted that unless we have the trivial case of $\tilde{f} = 0$, the following equation holds:

$$\frac{\epsilon_r - V}{\epsilon_e - V} = \frac{\Delta + \sigma_W^2}{\Delta + \sigma_V^2}$$

where $V = Var(\theta)$ and $\Delta \triangleq n\,Var(\tilde{f}) + \gamma\,(1-\gamma) E\left(|f_1(\theta) - f_2(\theta)|^2\right)$. Then, for all feasible $\gamma, f_1(\theta), f_2(\theta)$,

$$\begin{aligned} \epsilon_e = V - (V - \epsilon_r)\frac{\Delta + \sigma_V^2}{\Delta + \sigma_W^2} &\leq V - (V - \alpha_2)\frac{\Delta + \sigma_V^2}{\Delta + \sigma_W^2} \\ &\leq V - (V - \alpha_2)\frac{\Delta^* + \sigma_V^2}{\Delta^* + \sigma_W^2} \end{aligned} \tag{4.16}$$

where $\Delta^* = \min\limits_{\gamma, f_1, f_2} \Delta$ s.t., $\epsilon_r \leq \alpha_2$. Note that the first inequality in (4.16) is due to the fact that $\epsilon_r \leq \alpha_2$ in the feasible region, and the second inequality is due to the fact that $(\Delta + \sigma_V^2)/(\Delta + \sigma_W^2)$ is an increasing function of Δ as $\sigma_W^2 > \sigma_V^2$ with $\Delta \geq 0$. As (4.16) provides a global upper bound for ϵ_e, if there exists a feasible (γ, f_1, f_2) such that ϵ_e attains the global bound, then it is concluded that ϵ_e is maximized with it. A sufficient condition for the existence of such a case is that the solution of $\min\limits_{\gamma, f_1, f_2, \epsilon_r \leq \alpha_2} \Delta$ satisfies the constraint with equality, i.e., $\epsilon_r = \alpha_2$.

Therefore, we aim to obtain the solution of the following problem:

$$\min_{\gamma, f_1(\theta), f_2(\theta)} nVar(\tilde{f}) + \gamma(1-\gamma)E\left(|f_1(\theta) - f_2(\theta)|^2\right) \text{ s.t.}$$
$$\frac{n\,Cov(\tilde{f},\theta)^2}{n\,Var(\tilde{f}) + \gamma\,(1-\gamma)E\left(|f_1(\theta) - f_2(\theta)|^2\right) + \sigma_V^2} \geq V - \alpha_2 \tag{4.17}$$

Note that for any possible $\tilde{f}$, which is obtained using a feasible (γ, f_1, f_2), there are infinitely many alternative ways of constructing it with other feasible (γ, f_1, f_2)'s. Among all constructions, choosing $\tilde{f} = f_1 = f_2$ yields a smaller objective value and a larger value for the left side of the constraint in (4.17), implying that it is the optimal selection. Therefore, the problem reduces to

$$\min_{\tilde{f}} Var(\tilde{f}) \text{ s.t. } V - \frac{n\,Cov(\tilde{f},\theta)^2}{n\,Var(\tilde{f}) + \sigma_V^2} \leq \alpha_2 \tag{4.18}$$

The constraint in (4.18) can be expressed as

$$\frac{n\left(Var(\theta)Var(\tilde{f}) - Cov(\tilde{f},\theta)^2\right) + \sigma_V^2 Var(\theta)}{n\,Var(\tilde{f}) + \sigma_V^2} \leq \alpha_2$$

Note that $Var(\theta)Var(\tilde{f}) - Cov(\tilde{f},\theta)^2 \geq 0$ for any $\tilde{f}$ due to Cauchy-Schwarz inequality. Therefore, $Var(\tilde{f}) \geq \sigma_V^2(Var(\theta) - \alpha_2)/(n\alpha_2)$ for any $\tilde{f}$. This global lower bound can be achieved via $\tilde{f}(\theta) = k_1^\dagger\theta + k_2^\dagger$ with $k_1^\dagger$ being given by (4.14) and $k_2^\dagger$ being selected as any value to guarantee $\tilde{f}(\theta) \in [a, b]$. It is noted that when (4.9) is a feasible problem, $|k_1^\dagger| \leq 1$. For such an encoding, $\Delta^* = \sigma_V^2(Var(\theta) - \alpha_2)/\alpha_2$ and $\epsilon_r = \alpha_2$, i.e., the constraint is satisfied with equality in (4.17). Therefore, an optimal solution of (4.17), which is a deterministic affine function, is also an optimal solution of (4.9), which yields the optimal value of $\epsilon_e^\dagger = \frac{\sigma_W^2\,Var(\theta)\,\alpha_2}{\sigma_V^2(Var(\theta)-\alpha_2)+\sigma_W^2\alpha_2}$.

Based on the preceding discussion and Corollary 4, it can be argued that an optimal solution to (4.8) is a deterministic affine function when $\sigma_W^2 > \sigma_V^2$. First, notice that any optimal solution to (4.9) should satisfy the constraint with equality, i.e., $\epsilon_r = \alpha_2$. This is due to the fact for any other solution which does not satisfy the constraint with equality, the inequality in (4.16) would strictly be

implying a gap between ϵ_e and the global bound, and it is already shown that this bound can actually be achieved. Therefore, the result of Corollary 4 can be applied to connect the solutions of (4.8) and (4.9) and to imply that the deterministic affine functions solve (4.8) as well under the conditions of Proposition 3. Via Corollary 4 and (4.15), the expression in (4.13) can be obtained after a rearrangement. ■

There are some interesting observations regarding the result in Proposition 3. First, randomization between two functions does not bring any benefits over deterministic encoding when the intended receiver has already a less noisy channel than the eavesdropper, and the encoding function can be selected as a simple affine function. Second, for a given α_1 (or, α_2) value, ϵ_r^* (and $\epsilon_e^\dagger$) does not depend on n; however, the slope of the deterministic affine optimal function decays with $1/\sqrt{n}$. This means that the transmit power per channel use should be decreased as n increases such that the total transmitted signal power to send θ with n channel uses stays constant. Also, the constant term in the deterministic affine optimal function does not have any effects; hence, it can be chosen freely as long as the function remains in the feasible set.

Even though Proposition 3 provides a closed-form expression for an optimal solution when $\sigma_W^2 > \sigma_V^2$, it does not bring any conclusions into the case of $\sigma_W^2 < \sigma_V^2$. In order to obtain the solutions of the optimization problems in (4.8) and (4.9) in this case, the solution methods provided in Section 2.2.2 can be adopted, and ϵ_e and ϵ_r can directly be calculated using (4.10). In this chapter, the piecewise linear approximation method described in Section 2.2.2 is utilized to obtain the optimal solutions when $\sigma_W^2 < \sigma_V^2$. In particular, for $f_i(\theta)$, the increment in the kth interval in $[a, b]$ is defined as $\Delta x_k^{(i)} \triangleq f_i(a + k\Delta\theta) - f_i(a + (k-1)\Delta\theta)$ for $k = 1, \dots M$, and the optimization is performed over $2M + 1$ variables, that is, $[\Delta x_1^{(1)}, \Delta x_2^{(1)}, \dots, \Delta x_M^{(1)}, \Delta x_1^{(2)}, \Delta x_2^{(2)}, \dots, \Delta x_M^{(2)}, \gamma]$, by using the Global Optimization Toolbox of MATLAB. In the numerical examples, M is taken to be 25, which seems to provide a good trade-off between accuracy and complexity.

Next, we investigate a special case in which the encoding function is restricted to be affine.

4.2.2 Affine Encoding Functions

In this section, it is assumed that encoding is performed via affine encoding functions such that $f_1(\theta) = k_1\theta + k_2$ and $f_2(\theta) = s_1\theta + s_2$.[3] For this case, the MSE of the intended receiver (and the eavesdropper by using σ_W^2) can be expressed in terms of k_1, k_2, s_1 and s_2 as a corollary to Proposition 1.

Corollary 5: *The MSE (ϵ_r) of the LMMSE estimator at the intended receiver for the encoding model specified in* (4.3) *when $f_1(\theta) = k_1\theta + k_2$ and $f_2(\theta) = s_1\theta + s_2$ is*

$$\epsilon_r = Var(\theta)\frac{\gamma\,(1-\gamma)\,\kappa + \sigma_V^2}{n\,Var(\theta)(\gamma\,k_1 + (1-\gamma)s_1)^2 + \gamma\,(1-\gamma)\,\kappa + \sigma_V^2} \tag{4.19}$$

where

$$\kappa \triangleq E\left(((k_1 - s_1)\theta + (k_2 - s_2))^2\right). \tag{4.20}$$

Proof: For the given f_1 and f_2, c_1 and c_2 defined in Proposition 1 become $k_1\,Var(\theta)$ and $s_1\,Var(\theta)$, respectively. Hence, the numerator of the second term in (4.10) becomes $n(\gamma\,k_1 + (1-\gamma)\,s_1)^2\,Var(\theta)^2$. Also, the denominator of (4.10) can be rewritten as $n(x-t) + \tau - x$, where x, τ and t are as defined in (4.11). Note that $(x-t) = \gamma^2 k_1^2 Var(\theta) + (1-\gamma)^2 s_1^2 Var(\theta) + 2\gamma(1-\gamma)k_1 s_1 Var(\theta) = (\gamma k_1 + (1-\gamma)s_1)^2\,Var(\theta)$, and $\tau - x = \gamma\,(1-\gamma)\kappa + \sigma_V^2$, where κ is as defined in (4.20). After arranging the terms, the final expression in (4.19) is obtained. ■

When the encoding functions are restricted to affine functions, the optimization problems in (4.8) and (4.9) involve a search over only 5 variables instead of functions. Let $\mathbf{x_a} \triangleq [\gamma, k_1, k_2, s_1, s_2]$ and $T_a(\mathbf{x_a}, \sigma_V^2) \triangleq \epsilon_r$, where ϵ_r is as defined

[3] k_1 and k_2 should be such that $k_1\theta + k_2 \in [a, b]$ for all $\theta \in [a, b]$. Similarly, $s_1\theta + s_2$ needs to be in $[a, b]$ for all $\theta \in [a, b]$. Note that this requires $|k_1| \leq 1$ and $|s_1| \leq 1$.

in (4.19). Then, the optimization problems can be written as

$$\min_{\mathbf{x_a}} T_a(\mathbf{x_a}, \sigma_V^2) \;\; \text{s.t.} \;\; T_a(\mathbf{x_a}, \sigma_W^2) \geq \alpha_1 \tag{4.21}$$

$$\max_{\mathbf{x_a}} T_a(\mathbf{x_a}, \sigma_W^2) \;\; \text{s.t.} \;\; T_a(\mathbf{x_a}, \sigma_V^2) \leq \alpha_2 \tag{4.22}$$

where $T_a(\mathbf{x_a}, \sigma_W^2) \triangleq \epsilon_e$. It is noted that the optimization problems in (4.21) and (4.22) are much easier to solve than those in the case of encoding with generic functions.

Finally, as the closed form expression for the MSE with affine encoding can be calculated based on given encoding coefficients, it is also possible to investigate its behavior as γ changes. Namely, the aim is to provide regions of $\gamma \in [0, 1]$ in which the MSE increases or decreases with respect to γ. Such a characterization is helpful for both theoretical analysis and gaining intuition on the benefits of randomization. In addition, it facilitates the specification of the exact optimal solution of γ for the given encoding functions, i.e., k_1, k_2, s_1, s_2, and secrecy target. The following proposition characterizes the behavior of the MSE with respect to γ, where γ is taken as a real number (the case of $\gamma \in [0, 1]$ immediately follows as a corollary).

Proposition 4: *Define* $\nu(\gamma) \triangleq \nu_2\gamma^2 + \nu_1\gamma + \nu_0$ *with*

$$\begin{aligned} \nu_2 &\triangleq -\kappa(k_1^2 - s_1^2) \\ \nu_1 &\triangleq -2\,\kappa\, s_1^2 - 2\sigma_V^2(k_1 - s_1)^2 \\ \nu_0 &\triangleq \kappa s_1^2 - 2\sigma_V^2(k_1 - s_1)s_1 \end{aligned} \tag{4.23}$$

where κ *is as defined in* (4.20)*. Then,*

- *if* $\nu_2 = 0$ *and* $\nu_1 > 0$*, then* ϵ_r *is an increasing (a decreasing) function of* γ *for* $\gamma > -\nu_0/\nu_1$ *(*$\gamma < -\nu_0/\nu_1$*);*
- *if* $\nu_2 = 0$ *and* $\nu_1 < 0$*, then* ϵ_r *is a decreasing (an increasing) function of* γ *for* $\gamma > -\nu_0/\nu_1$ *(*$\gamma < -\nu_0/\nu_1$*);*
- *if* $\nu_2 > 0$*, then* ϵ_r *is a decreasing function of* γ *when* γ *is in between the roots*

of $v(\gamma)$*=0, which are* $\frac{\kappa s_1 - 2\sigma_V^2(k_1-s_1)}{\kappa(k_1+s_1)}$ *and* $\frac{-s_1}{k_1-s_1}$*, and an increasing function elsewhere;*

- *if* $\nu_2 < 0$*, then* ϵ_r *is an increasing function of* γ *when* γ *is in between the roots of* $v(\gamma) = 0$*, and a decreasing function elsewhere;*
- *if* $\nu_1 = \nu_2 = 0$*, then* ϵ_r *is constant with respect to* γ*.*

Proof: From (4.19), the MSE can be expressed as $\epsilon_r = Var(\theta)h(\gamma)/(\xi g(\gamma)^2 + h(\gamma))$, where $h(\gamma) = \gamma(1-\gamma)\kappa + \sigma_V^2$, $g(\gamma) = (k_1 - s_1)\gamma + s_1$, and $\xi = nVar(\theta) > 0$. Consider the derivative of the MSE with respect to γ, i.e., $d\epsilon_r/d\gamma$. As the denominator of $d\epsilon_r/d\gamma$ is always positive, it is enough to characterize the sign of its numerator with respect to γ. Let $\hat{v}(\gamma)$ denote the numerator of $d\epsilon_r/d\gamma$.[4] Then,

$$\begin{aligned}\hat{\nu}(\gamma) &= h'(\gamma)\left(\xi g(\gamma)^2 + h(\gamma)\right) - h(\gamma)\left(2\xi g(\gamma)g'(\gamma) + h'(\gamma)\right) \\ &= \xi g(\gamma)\left(h'(\gamma)g(\gamma) - 2h(\gamma)g'(\gamma)\right) \triangleq \xi\, v(\gamma) \end{aligned} \tag{4.24}$$

where $h'(\gamma) = (1-2\gamma)\kappa$ and $g'(\gamma) = k_1 - s_1$. After inserting these into (4.24), $\nu(\gamma)$ becomes

$$\begin{aligned}\nu(\gamma) &= \left((k_1 - s_1)\gamma + s_1\right)\left(-\kappa(k_1+s_1)\gamma + \kappa s_1 - 2\sigma_V^2(k_1 - s_1)\right) \\ &= \nu_2\gamma^2 + \nu_1\gamma + \nu_0 \end{aligned} \tag{4.25}$$

where ν_2, ν_1, and ν_0 are as given in (4.23). As the roots of $v(\gamma)$ are $\frac{\kappa s_1 - 2\sigma_V^2(k_1-s_1)}{\kappa(k_1+s_1)}$ and $\frac{-s_1}{k_1-s_1}$, the conclusions in the proposition can be obtained by applying the sign test to $v(\gamma)$. ■

The result in Proposition 4 can be used to find the optimal γ directly when k_1, k_2, s_1 and s_2 are fixed. For example, consider a scenario with a single observation ($n = 1$), $\sigma_V = 0.01$, $\sigma_W = 0.5$, and a secrecy target of $\alpha_1 = 0.08$. If $f_1(\theta) = \theta$ and $f_2(\theta) = 1 - \theta$, where θ is uniformly distributed in $[0, 1]$, then $\nu_2 = 0$ and $\nu_1 < 0$ with $-\nu_0/\nu_1 = 1/2$ for both ϵ_r and ϵ_e. Therefore, when $\gamma > 1/2$, the MSE is a decreasing function of γ and when $\gamma < 1/2$ it is an increasing function of

[4]The $Var(\theta)$ term is omitted in the expression as it is always positive.

γ according to Proposition 4. Due to the symmetry in this specific problem, it is possible to restrict γ to $\gamma \in [0, 1/2]$. Therefore, when γ increases, the MSEs (both ϵ_r and ϵ_e) increase monotonically until $\gamma = 1/2$, as well. As the goal is to minimize ϵ_r, it is obvious that γ should be increased until it yields $\epsilon_e = \alpha_1 = 0.08$ but no more. Finally, $\gamma = 0.3$ can be obtained as the optimal probability, and the corresponding MSE at the intended receiver becomes $\epsilon_r = 0.07$.

4.3 Large Number of Observations

In this section, it is assumed that a large number of observations are available to the intended receiver and the eavesdropper to estimate θ.[5] As motivated in Section 4.1, the ECRB metric is employed for both the intended receiver and the eavesdropper in this scenario. The constraints on the parameter space and the encoding functions are the same as in the previous section.

The ECRB is defined as the expectation of the conditional CRB with respect to the unknown parameter [48], which is expressed as

$$E_\theta\big((I^{(n)}(\theta))^{-1}\big) = \int_a^b p_\theta(\theta) \frac{1}{I^{(n)}(\theta)}\, d\theta \triangleq ECRB \tag{4.26}$$

where $p_\theta(\theta)$ is the prior PDF of θ, $I^{(n)}(\theta)^{-1}$ corresponds to the conditional CRB for estimating θ and $I^{(n)}(\theta)$ denotes the Fisher information based on n observations. Therefore, for the intended receiver, $I_r^{(n)}(\theta)$ can be expressed as

$$I_r^{(n)}(\theta) = \int \left(\frac{\partial \log p(\mathbf{y}|\theta)}{\partial \theta} \right)^2 p(\mathbf{y}|\theta)\, d\mathbf{y} \tag{4.27}$$

with $p(\mathbf{y}|\theta)$ representing the conditional PDF of the n observations for a given value of θ [47]. Also, due to (4.4), $I_r^{(n)}(\theta) = nI_r(\theta)$, where $I_r(\theta)$ is the Fisher

[5]It should be emphasized that the ECRB approaches the MSE of the MMSE estimator in the asymptotic region, which refers to either a large number of observations or high SNR/SINR scenarios [48]. When stochastic encoding is employed, there exists a certain interference term in the received signal limiting the effective SINR. Therefore, the ECRB metric is not reliable for a small number of observations even for a small noise variance.

information based on $p(y|\theta) = \gamma\, p_V(y - f_1(\theta)) + (1-\gamma)\, p_V(y - f_2(\theta))$. Therefore,

$$I_r(\theta) = \int_{-\infty}^{\infty} \frac{u(\theta)^2}{p(y|\theta)}\, dy \tag{4.28}$$

where

$$\begin{aligned} u(\theta) = \gamma \frac{1}{\sqrt{2\pi}\sigma_V} e^{\frac{-(y-f_1(\theta))^2}{2\sigma_V^2}} \frac{(y - f_1(\theta))}{\sigma_V^2} f_1^{'}(\theta) \\ + (1-\gamma) \frac{1}{\sqrt{2\pi}\sigma_V} e^{\frac{-(y-f_2(\theta))^2}{2\sigma_V^2}} \frac{(y - f_2(\theta))}{\sigma^2} f_2^{'}(\theta) \end{aligned} \tag{4.29}$$

and

$$p(y|\theta) = \frac{\gamma}{\sqrt{2\pi}\sigma_V} e^{\frac{-(y-f_1(\theta))^2}{2\sigma_V^2}} + \frac{1-\gamma}{\sqrt{2\pi}\sigma_V} e^{\frac{-(y-f_2(\theta))^2}{2\sigma_V^2}} \tag{4.30}$$

In addition, when (4.28) is employed in (4.26), the ECRB at the intended receiver, E_r, is obtained as

$$E_r = \frac{1}{n} \int_a^b p_\theta(\theta) \frac{1}{I_r(\theta)}\, d\theta\,. \tag{4.31}$$

Similarly, the ECRB at the eavesdropper can be obtained by defining Fisher information $I_e(\theta)$ based on $p(z|\theta) = \gamma\, p_W(z - f_1(\theta)) + (1-\gamma)\, p_W(z - f_2(\theta))$, which can be calculated as in (4.28)-(4.30). Then, the ECRB at the eavesdropper, E_e, is

$$E_e = \frac{1}{n} \int_a^b p_\theta(\theta) \frac{1}{I_e(\theta)}\, d\theta\,. \tag{4.32}$$

Therefore, similarly to (4.8) and (4.9), the optimization problems can be proposed as follows:

$$\min_{\gamma, f_1(\theta), f_2(\theta)} E_r \quad \text{s.t.} \quad E_e \geq \eta_1 \tag{4.33}$$

$$\max_{\gamma, f_1(\theta), f_2(\theta)} E_e \quad \text{s.t.} \quad E_r \leq \eta_2 \tag{4.34}$$

where η_1 and η_2 denote the secrecy target for the first problem and the estimation

accuracy limit at the intended receiver for the second problem. Even though the simplification to (4.28) may not be possible for the generic case, calculating the ECRB is still easier and more practical for a large number of observations than calculating the MSEs of estimators such as the MAP or MMSE estimators.

Remark 2: Similarly to the results in Proposition 2 and Corollary 1–4, the exact relationship between the solutions of (4.33) and (4.34) can be obtained based on a similar approach, which is not repeated here for brevity.

It is noted that if the encoding function is deterministic, then simplification is possible for both E_r and E_e. The following proposition provides the solutions to the optimization problems in (4.33) and (4.34) in the absence of randomization.

Proposition 5: *Suppose that a deterministic encoding function $f(\theta)$ is employed at the transmitter. For a given feasible secrecy target η_1, the optimal value of the optimization problem in* (4.33) *is $\eta_1\,\sigma_V^2/\sigma_W^2$. Furthermore, any $f(\theta)$ with $(\sigma_W^2/n)\int_a^b p_\theta(\theta)/f'(\theta)^2 d\theta = \eta_1$ is an optimal deterministic encoding function for* (4.33). *Similarly, for a given estimation accuracy limit η_2, the optimal value of the optimization problem in* (4.34) *is $\eta_2\,\sigma_W^2/\sigma_V^2$. Furthermore, any $f(\theta)$ with $(\sigma_V^2/n)\int_a^b p_\theta(\theta)/f'(\theta)^2 d\theta = \eta_2$ is an optimal deterministic encoding function for* (4.34).

Proof: When a deterministic encoding function $f(\theta)$ is employed at the transmitter, $I_r(\theta)$ in (4.28) simplifies to $I_r(\theta) = f'(\theta)^2/\sigma_V^2$. Similarly, $I_e(\theta) = f'(\theta)^2/\sigma_W^2$. Then, the optimization problem in (4.33) becomes

$$\begin{aligned} \min_{f(\theta)} \ & \frac{\sigma_V^2}{n}\int_a^b p_\theta(\theta)\frac{1}{f'(\theta)^2}\,d\theta \\ \text{s.t. } & \frac{\sigma_W^2}{n}\int_a^b p_\theta(\theta)\frac{1}{f'(\theta)^2}\,d\theta \geq \eta_1\,. \end{aligned} \tag{4.35}$$

As the integral term is identical in both the objective and the constraint functions, the argument in Proposition 5 follows by choosing an encoding function that satisfies the constraint with equality. The result for (4.34) can be justified similarly. ■

Proposition 5 shows that if there is no randomization in the encoding function, then the ratio of E_r/E_e depends only on the noise variances in the channels of the eavesdropper and the intended receiver. Therefore, any deterministic encoding function can be used at the transmitter as long as it satisfies the constraints. Also, it is noted that the only difference between using a generic deterministic encoding function and an affine deterministic encoding function is that the former may support a larger set of feasible η_1 (or, η_2) values.

Finally, it is possible to obtain some theoretical and intuitive results for the generic stochastic encoding scheme in (4.3) by using the convexity of the Fisher information with respect to the conditional distribution [71]. Specifically, let the Fisher information based on $p_1(y|\theta)$ and $p_2(y|\theta)$ be denoted by $I_1(\theta)$ and $I_2(\theta)$, respectively. If $p_3(y|\theta) = \gamma p_1(y|\theta) + (1-\gamma)p_2(y|\theta)$, then the Fisher information $I_3(\theta)$ based on $p_3(y|\theta)$ satisfies $I_3(\theta) < \gamma I_1(\theta) + (1-\gamma)I_2(\theta)$ given that $\gamma \in (0,1)$ and $p_1(y|\theta) \neq p_2(y|\theta)$. This implies that $I_3(\theta)$ is also a convex function of γ for any given $\theta \in [a,b]$, and it always remains below the linear line connecting $I_1(\theta)$ and $I_2(\theta)$.

This convexity property is helpful for providing a few intuitive and analytical results. For example, a lower bound for the ECRB can be obtained when $f_1(\theta)$ and $f_2(\theta)$ correspond to affine encoding. To that end, consider the affine encoding scheme described in Section 4.2.2. Then, $I_1(\theta) = k_1^2/\sigma^2$ and $I_2(\theta) = k_2^2/\sigma^2$. Then, $I_3(\theta) < \left(\gamma k_1^2 + (1-\gamma)\, k_2^2\right)/\sigma^2 \ \forall \theta \in [a,b]$. Therefore, for the ECRB of the intended receiver, it is obtained that $E_r > \frac{\sigma_V^2}{n(\gamma k_1^2 + (1-\gamma)\, k_2^2)}$ and for the ECRB of the eavesdropper, it is obtained that $E_e > \frac{\sigma_W^2}{n(\gamma k_1^2 + (1-\gamma) k_2^2)}$. The following proposition provides a result for symmetric encoding:

Proposition 6: *Consider the symmetric mapping with $f_1(\theta) = g(\theta)$ and $f_2(\theta) = g_0 - g(\theta)$ such that $g(\theta) \in [a,b]$ and $g_0 - g(\theta) \in [a,b]$ for all $\theta \in [a,b]$. Then, the ECRB is maximized at $\gamma = 1/2$.*

Proof: Let $\gamma = \gamma_0 \in [0,1]$. For the given model, $I(\theta) =$

$g'(\theta)^2 \int_{-\infty}^{\infty} \hat{u}(\theta)^2/p(y|\theta)\,dy$, where

$$\begin{aligned}\hat{u}(\theta) &= \gamma_0 \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(y-g(\theta))^2}{2\sigma^2}} \frac{(y-g(\theta))}{\sigma^2} \\ &\quad - (1-\gamma_0)\frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(y+g(\theta)-g_0)^2}{2\sigma^2}} \frac{(y+g(\theta)-g_0)}{\sigma^2} \\ &\triangleq m(y,\theta,\gamma_0) \end{aligned} \tag{4.36}$$

and

$$\begin{aligned} p(y|\theta) &= \gamma_0 \frac{1}{\sqrt{2\pi}\sigma_V} e^{\frac{-(y-g(\theta))^2}{2\sigma_V^2}} \\ &\quad + (1-\gamma_0)\frac{1}{\sqrt{2\pi}\sigma_V} e^{\frac{-(y+g(\theta)-g_0)^2}{2\sigma_V^2}} \triangleq d(y,\theta,\gamma_0). \end{aligned} \tag{4.37}$$

If the change of variables with $g_0 - y = \hat{y}$ is applied in the integration for $I(\theta)$, it is obtained that $I(\theta) = g'(\theta)^2 \int_{-\infty}^{\infty} \frac{m(\hat{y},\theta,1-\gamma_0)^2}{d(\hat{y},\theta,1-\gamma_0)}\,d\hat{y}$. Therefore, $I(\theta)$ attains the same value for $\gamma = \gamma_0$ and $\gamma = 1-\gamma_0$; hence, it is a symmetric function of γ around $\gamma = 1/2$ for any $\theta \in [a,b]$. Due to this fact and the convexity of $I(\theta)$ with respect to γ, its minimum occurs at $\gamma = 1/2$ for all $\theta \in [a,b]$, implying that the ECRB is maximized at $\gamma = 1/2$. ■

Finally, the behavior of the ECRB with respect to γ can be investigated for the general encoding scheme in (4.3) based on the convexity property, as stated in the following proposition. (Similar results can also be derived for $I_e(\theta)$.)

Proposition 7: *Let* $\frac{dI_r(\theta)}{d\gamma}\big|_{\gamma=0^+} \triangleq d_0$ *and* $\frac{dI_r(\theta)}{d\gamma}\big|_{\gamma=1^-} \triangleq d_1$. *Then,*

- *if* $d_1 < 0$ *for all* $\theta \in [a,b]$, $I_r(\theta)$ *is monotone decreasing with* γ, *implying that the ECRB is monotone increasing with* $\gamma \in (0,1)$*;*
- *if* $d_0 > 0$ *for all* $\theta \in [a,b]$, $I_r(\theta)$ *is monotone increasing with* γ, *implying that the ECRB is monotone decreasing with* $\gamma \in (0,1)$*;*
- *if* $d_0 < 0$ *and* $d_1 > 0$ *for a given* $\theta \in [a,b]$, $I_r(\theta)$ *has a minimum* $\gamma^* \in (0,1)$. *Furthermore, if* γ^* *minimizes* $I_r(\theta)$ *for all* $\theta \in [a,b]$, *then* E_r *is maximized at* $\gamma = \gamma^*$

Proof: Due to the strict convexity of $I_r(\theta)$ with respect to γ, $\frac{d^2 I_r(\theta)}{d\gamma^2} > 0$ holds for $\gamma \in (0,1)$. If $d_1 < 0$ for all $\theta \in [a,b]$, then $\frac{dI_r(\theta)}{d\gamma} < 0$ for all $\gamma \in (0,1)$ as the value of the derivative only increases as γ increases. Hence, $I_r(\theta)$ is a monotone decreasing function of γ for all $\theta \in [a,b]$, which implies that E_r is monotone increasing. Similarly, if $d_0 > 0$ for all $\theta \in [a,b]$, $\frac{dI_r(\theta)}{d\gamma} > 0$ for all $\gamma \in (0,1)$; hence, $I_r(\theta)$ is a monotone increasing function of γ for all $\theta \in [a,b]$, which implies that E_r is monotone decreasing. Finally, if $d_0 < 0$ and $d_1 > 0$, then via a similar argument, there exists a $\gamma = \gamma^*$ such that $\frac{dI_r(\theta)}{d\gamma}\big|_{\gamma=\gamma^*} = 0$, and it is the minimum for $I_r(\theta)$, and the rest of the arguments in the proposition follow from (4.31). ■

The following point should be noted related to γ^* in Proposition 7. Even though there may not exist such a γ^* which is the minimum for all $\theta \in [a,b]$ in general, E_r can still have a maximizer in $\gamma \in (0,1)$. Hence, it is only a sufficient condition, and the symmetric mapping given in Proposition 6 is an example in which this condition is satisfied.

Remark 3: The monotonicity results are important to gain intuition about the benefits of randomization and provide a practical tool and guide to obtain the optimal value of γ for given functions $f_1(\theta)$ and $f_2(\theta)$. For example, if the designer fixes the encoding functions to decrease system complexity, then the problem reduces to finding the optimal γ to satisfy the secrecy targets. (In some other scenarios, it may help reduce the search space.) However, in order to obtain the solutions of the optimization problems in (4.33) and (4.34) in general, similarly to the previous section, the piecewise linear approximation method described in Section 2.2.2 can be utilized, and E_e and E_r are calculated based on (4.26)–(4.32).

Remark 4: Even though the ECRB metric is also utilized in Section 2.2, the current problem setup is significantly different as it considers encoder randomization, multiple observations ($n > 1$), and the availability of encoding information at the eavesdropper. ECRB is only an optimization metric for the performance of the estimator at the receiver in Section 2.2, i.e., optimizing it *implies* improved overall performance. However, in this chapter, ECRB is used only when n is sufficiently large; hence, it is rather directly a tight approximation of the optimal

MSE value in the asymptotic region. Also, in Section 2.2, different metrics are utilized in the receiver (ECRB) and the eavesdropper (MSE of LMMSE estimator) whereas in this section, ECRB is utilized both in the intended receiver and the eavesdropper. Due to these reasons, most of the theoretical discussions in Section 2.2 cannot be applied to the current chapter.

4.4 Numerical Results

In this section, numerical examples are provided to investigate the theoretical results and the solution of the optimization problems proposed in Section 4.2 and 4.3.

4.4.1 Justification for LMMSE Estimator and ECRB Metric

In this section, we provide numerical examples to illustrate the motivation behind using different approaches for the cases of small and large numbers of observations. In all examples, the corresponding ECRB and the MSEs for the MMSE and LMMSE estimators are plotted versus the number of observations n. The SNR is defined as $10\log_{10}(1/\sigma^2)$, where σ^2 is the variance of the zero-mean Gaussian noise. In the first example, we consider a simple scenario in which the parameter is not encoded, i.e., $f(\theta) = \theta$. In the second example, the parameter is encoded by a simple piecewise linear deterministic encoding function such that $f(\theta) = 2\theta/3$ for $\theta \in [0, 0.5]$ and $f(\theta) = (4\theta - 1)/3$ for $\theta \in [0.5, 1]$. In both examples, it is assumed that θ has uniform distribution in $\theta \in [0, 1]$ and the SNR is set to 5 dB. The results are shown in Fig. 4.2 (top and bottom figures), and the corresponding encoding functions are provided in the upper right corner of each figure. It is observed that the MSEs of the LMMSE and MMSE estimators are close to each other when n is small whereas the ECRB converges to the MSE of the MMSE estimator for large n values in both figures. In the absence of

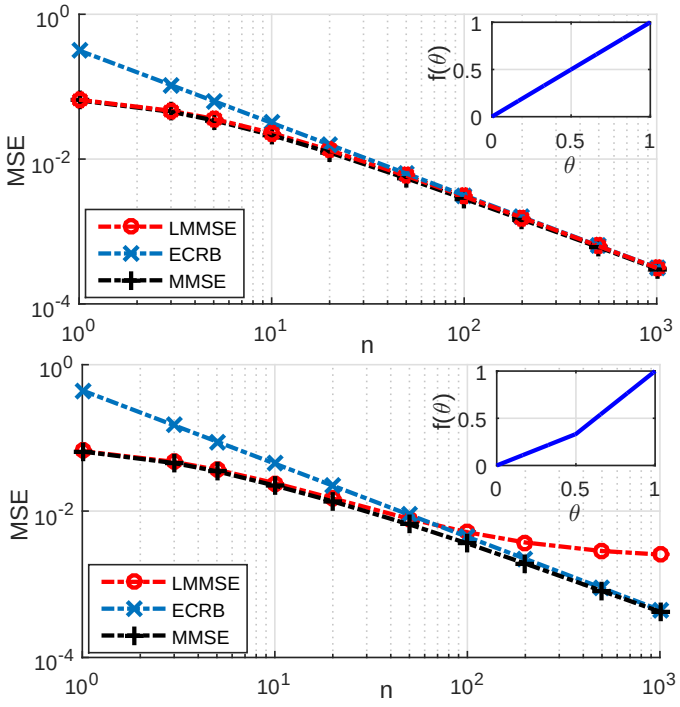

Figure 4.2: ECRB, LMMSE and MMSE versus n for two simple encoding scenarios.

encoding, the MSE performance of the MMSE and LMMSE estimators is almost the same for large numbers of observations, as well. However, the performance of the LMMSE estimator deviates from that of the MMSE estimator and the ECRB for large numbers of observations in the second example (with nonlinear encoding function), which motivates the use of ECRB in this regime in the general case. It is also noted that the ECRB is not a lower bound, and it rather identifies the optimal estimator behavior in asymptotic scenarios.

Next, we provide two numerical examples in Figs. 4.3 and 4.4 under stochastic encoding as modeled in (4.3). In both of the examples, it is assumed that $\gamma = 0.8$, $f_1(\theta) = \theta$, and $f_2(\theta) = 1 - \theta$ and $\theta \in [0, 1]$. Also, θ has uniform distribution in Fig. 4.3, and beta distribution with parameters $(2, 3)$, i.e., $p_\theta(\theta) = 12\,\theta\,(1-\theta)^2$, in Fig. 4.4. It is observed that for both SNR values in the figures, the MSE of the LMMSE estimator and the ECRB are close to the MSE of the MMSE estimator

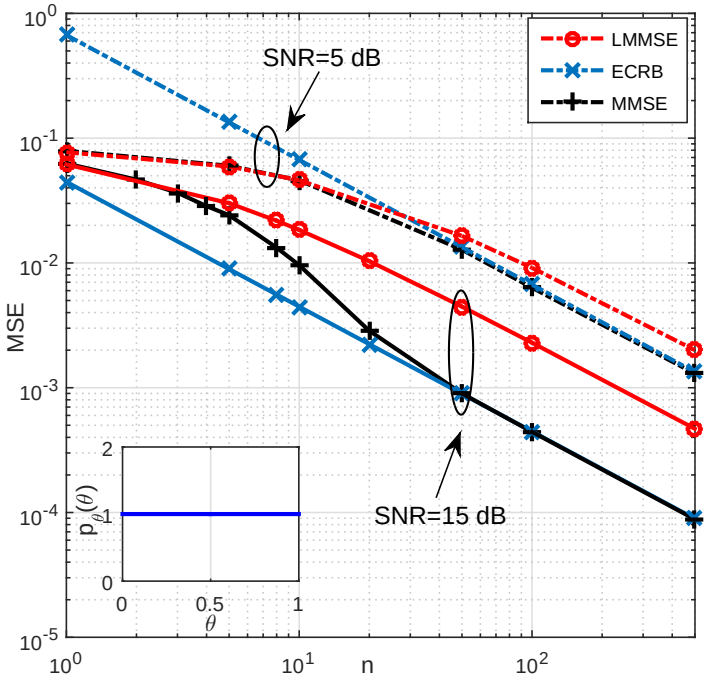

Figure 4.3: ECRB, LMMSE and MMSE versus n, where θ has uniform distribution in [0,1].

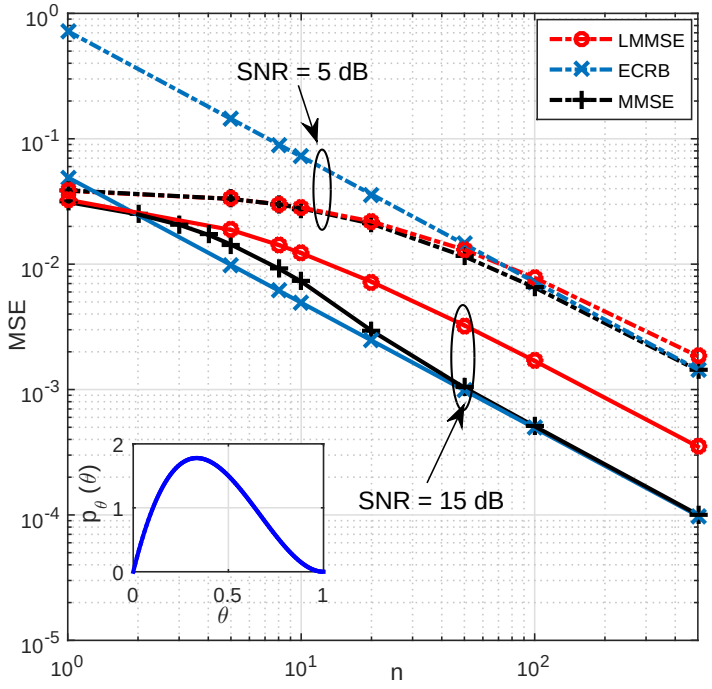

Figure 4.4: ECRB, LMMSE and MMSE versus n, where θ has beta distribution with parameters $(2, 3)$ in [0,1].

when n is small and large, respectively.[6] Another important observation is that as the noise variance decreases, the ECRB also reduces rapidly. For small values of n, the ECRB cannot capture the interference effect on the error due to the randomization employed in the encoder, and it can yield optimistic values for the MSE, which motivates the use of the LMMSE estimator in such scenarios. On the other hand, there is a performance gap between the LMMSE and MMSE estimators for large values of n. This is due to the fact that practical estimators start correctly deciding which mode of encoding (f_1 or f_2) is employed with larger observations. However, the LMMSE is unable to achieve such a decision, motivating the use of the ECRB in such scenarios as it is very tight in that region. Therefore, the LMMSE estimator and the ECRB can be utilized for small and large numbers of observations, respectively, at both the receiver and the eavesdropper.

Note that the MMSE solutions in these examples are obtained based on the following approach: For a given θ, $n-$dimensional realizations $\mathbf{y}$ are obtained empirically at each run of Monte-Carlo simulations. The conditional MSE is obtained using the MMSE estimator $\hat{\theta}(\mathbf{y}) = E(\theta|\mathbf{Y} = \mathbf{y})$, which is analytically calculated for a given $\mathbf{y}$ at each run. Finally, the MSE is obtained by taking the expectation of the conditional MSE over $p_\theta(\theta)$ analytically. The total number of Monte-Carlo runs is set to 10^5.

4.4.2 Small Number of Observations

In this section, numerical results are provided for the case of small number of observations. In all of the examples in this section, it is assumed that the number of observations is 5, i.e., $n = 5$, and θ is uniformly distributed in $[0, 2]$. The SNRs of the intended receiver and the eavesdropper are defined as $10\log_{10}(1/\sigma_V^2)$

[6] At high SNRs, the MSE of the MMSE estimator may be in between the ECRB and the MSE of the LMMSE estimator for medium values of n; hence, a more conservative approach can be taken and the ECRB can be used for the eavesdropper and the LMMSE metric can be used for the intended receiver in such a case.

and $10\log_{10}(1/\sigma_W^2)$, where σ_V^2 and σ_W^2 are the variances of the zero-mean Gaussian noise at each observation of the intended receiver and the eavesdropper, respectively. The following strategies are evaluated in the examples:

Stochastic generic: This strategy corresponds to the solution of (4.8) (and alternatively (4.9)), which provides optimal generic encoding functions $f_1(\theta)$ and $f_2(\theta)$, and the probability γ.

Stochastic affine: This strategy corresponds to the solution of (4.21) (and alternatively (4.22)), which provides the optimal affine encoding functions $f_1(\theta) = k_1\theta + k_2$ and $f_2(\theta) = s_1\theta + s_2$, and the probability γ.

Deterministic generic: This strategy corresponds to the solution of (4.8) (and alternatively (4.9)) when a deterministic generic encoding function $f(\theta)$ is employed at the transmitter.

Deterministic affine: This strategy corresponds to the solution of (4.21) (and alternatively (4.22)) when a deterministic encoding function $f(\theta) = k_1\theta + k_2$ is employed at the transmitter.

First, we consider the minimization of the MSE at the intended receiver for a given secrecy level at the eavesdropper, i.e., the optimization problems in (4.8) and (4.21). In the first example, two different scenarios are considered, and the MSE of the intended receiver is plotted versus the SNR of the intended receiver. In Scenario 1, the SNR of the eavesdropper is 20 dB, and the secrecy target $\alpha_1 = 0.26$ and in Scenario 2, the SNR of the eavesdropper is 15 dB, and the secrecy target $\alpha_1 = 0.04$. In Fig. 4.5, it is observed that when the SNR of the intended receiver is higher than the SNR of the eavesdropper, all strategies yield the same performance in both scenarios. This result is actually proved formally in Proposition 3, and the optimal value for the MSE of the intended receiver can be achieved by using a simple deterministic affine function. For example, when the SNR of the intended receiver is 30 dB, $f(\theta) = 0.013\,\theta$ is an optimal encoder for Scenario 1, yielding $\epsilon_r^* = 0.0872$, and $f(\theta) = 0.0663\,\theta$ is an optimal encoder for Scenario 2, yielding $\epsilon_r^* = 0.0014$ according to (4.12) and (4.13). It is also observed

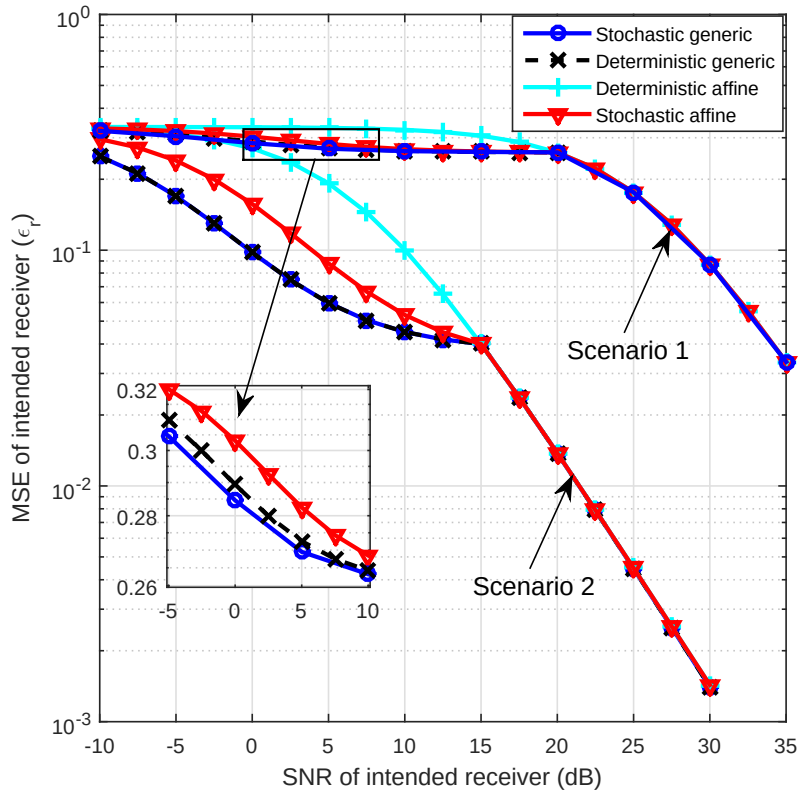

Figure 4.5: MSE of intended receiver (ϵ_r) versus SNR of intended receiver for two different scenarios.

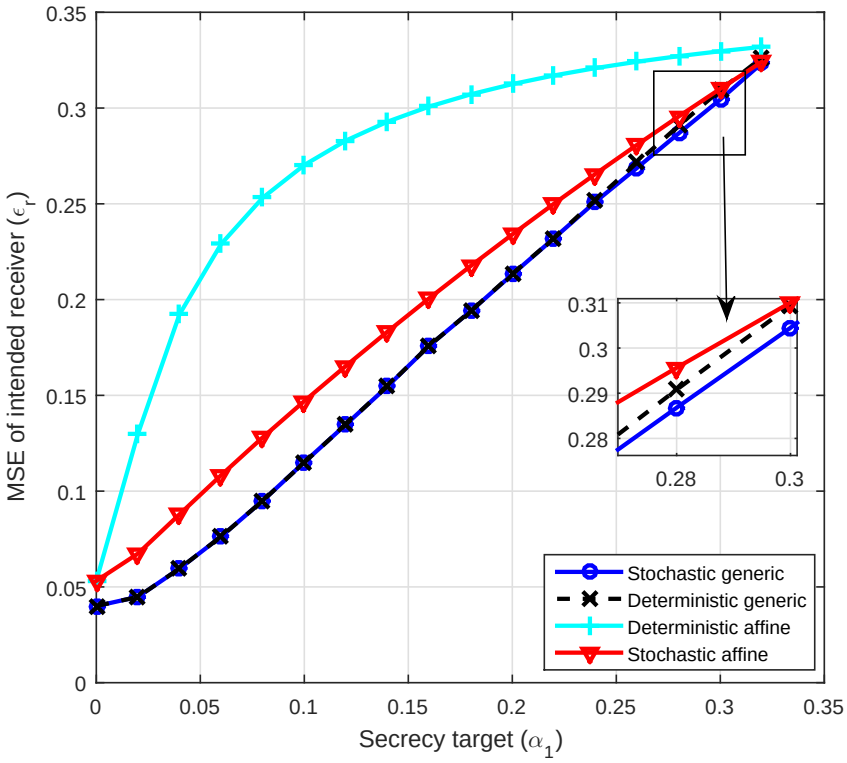

Figure 4.6: MSE of intended receiver (ϵ_r) versus secrecy target (α_1) when SNRs of eavesdropper and intended receiver are 15 and 5 dB, respectively.

in Fig. 4.5 that when the SNR of the intended receiver is lower than that of the eavesdropper, there is a performance gap between different strategies. In that region, the deterministic affine functions perform worse than the other strategies, and applying randomization to affine functions brings significant performance gains. Also, the generic functions yield lower MSE values than affine functions. In Scenario 1, stochastic generic functions bring a small performance gain over deterministic generic functions. However, stochastic and deterministic generic functions yield the same performance in Scenario 2, implying that randomization is not necessary if a generic function is employed in that scenario. Also, the MSE of the intended receiver is equal to α_1 for all strategies when the SNRs of the intended receiver and the eavesdropper are the same.

In Fig. 4.6, the MSE of the intended receiver is plotted versus the secrecy target at the eavesdropper when the SNRs of the eavesdropper and the intended receiver

are 15 and 5 dB, respectively. Obviously, as the secrecy target becomes larger, the MSE of the intended receiver increases, as well. When the secrecy target is very small (≈ 0) or very ambitious ($\approx Var(\theta)$), all the strategies have similar performance. For medium values of α_1, it is observed that the deterministic affine function strategy performs significantly worse than the other strategies. However, the stochastic affine strategy has significantly closer performance to that of generic functions. When α_1 is less than 0.24, randomization does not bring any improvements over the deterministic generic strategy. However, as α_1 gets larger (that is, a relatively large MSE is required at the eavesdropper), stochastic generic functions have slightly better performance than deterministic ones. This implies that it is not possible to claim that deterministic generic functions are an optimal class of functions in all settings even though their performance is not far from that of stochastic generic functions.

In Fig. 4.7, the optimal encoding functions for different strategies are plotted when the SNRs of the eavesdropper and the intended receiver are 10 and 0 dB, respectively, and the secrecy target α_1 is 0.28. Some important observations can be made from the figure related to the optimal functions. First, it is noticed that the deterministic affine function maps $\theta \in [0, 2]$ to a smaller interval ($[0, 0.213]$) to solve the optimization problem and has a low degrees of freedom in the mapping operation. On the other hand, the stochastic affine strategy sends an affine function $f_1(\theta) = 0.4625\,\theta + 1.075$ with probability 0.604 and nothing (i.e., $f_2(\theta) \approx 0$) with probability 0.396. Furthermore, the characteristics of the generic functions are quite different from those of the affine functions. The optimal deterministic generic function is $f(\theta) \approx 2$ if $\theta < 0.1232$, and $f(\theta) \approx 0$ otherwise.[7] This implies that the optimal deterministic function actually converges to a non-uniform quantizer such that θ values are mapped to 0 and 2. Furthermore, the stochastic generic function strategy randomizes between two *quantizer-like* generic functions to outperform the optimal deterministic encoding function strategy. The intuition behind such a scheme is that a quantizer-like encoder already assigns ≈ 2 and ≈ 0

[7]Note that the encoding functions are required to be one-to-one functions in this chapter; therefore, even though they are not allowed to stay constant over an interval, it is easy to make sure that they are arbitrarily close to being constant and still do not violate the one-to-one assumption. Also note that if $f(\theta)$ is an optimal deterministic solution, then $\bar{f}(\theta) = \bar{f}_0 \pm f(\theta)$ is also an optimal solution as long as $\bar{f}(\theta) \in [a, b]$, where $\bar{f}_0$ is a constant.

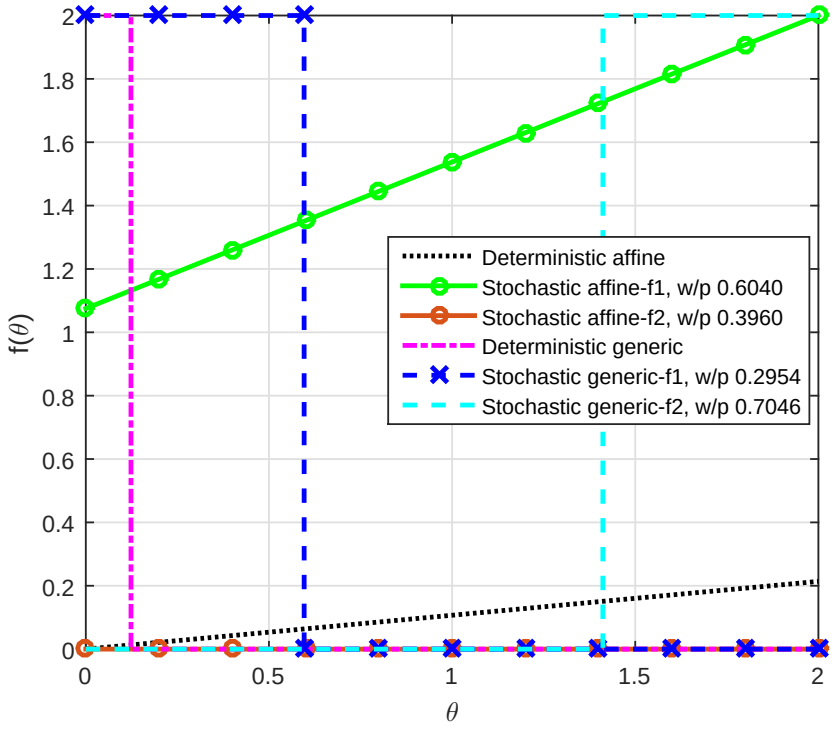

Figure 4.7: Optimal encoding functions for different strategies when SNRs of eavesdropper and intended receiver are 10 and 0 dB, respectively, and secrecy target α_1 is 0.28.

for a set of θ values and provides one layer of ambiguity. Then, randomization over these two quantizer-like functions provides an extra layer of ambiguity about the parameter to achieve required secrecy targets for the eavesdropper.

In Fig. 4.8, the optimal encoding functions for different strategies are plotted when the SNRs of the eavesdropper and the intended receiver are 15 and 5 dB, respectively, and the secrecy target α_1 is 0.04. In this case, the secrecy constraint is not as ambitious as the previous one. Similarly to the previous case, the deterministic affine function maps $\theta \in [0, 2]$ to a smaller interval ($[0, 0.746]$). The stochastic affine approach sends the original value of the parameter with probability 0.8775 but it maps θ to ≈ 2 with probability 0.1225. According to Fig. 4.5, the deterministic and stochastic affine approaches yield the MSE values of 0.1923 and 0.088, respectively, illustrating the benefits of randomization. In addition, the optimal deterministic generic function (and also the optimal

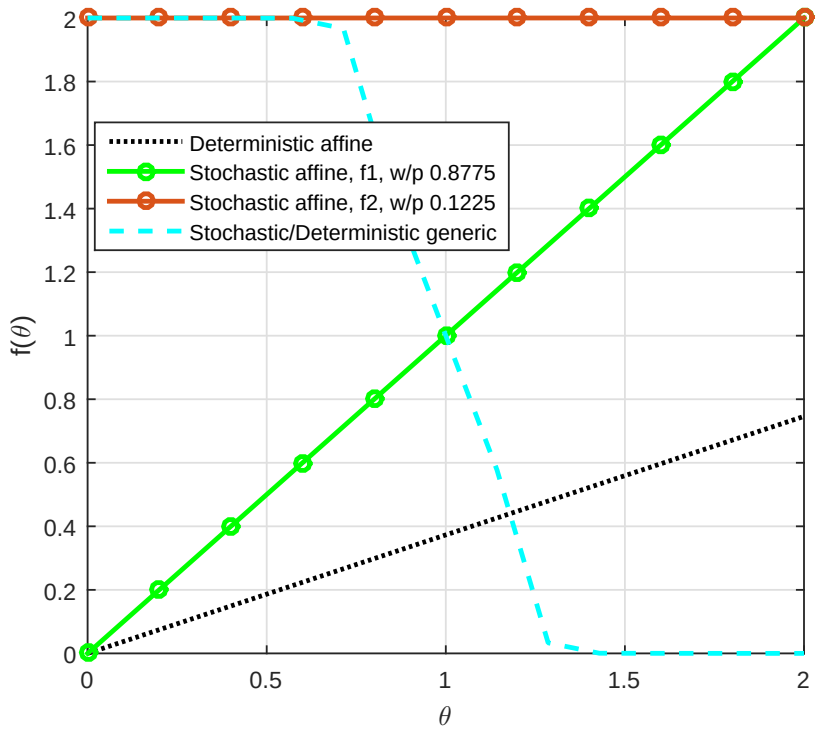

Figure 4.8: Optimal encoding functions for different strategies when SNRs of eavesdropper and intended receiver are 15 and 5 dB, respectively, and secrecy target α_1 is 0.04.

stochastic generic function) has different characteristics than the one in Fig. 4.7. In particular, it has three different regions; namely, $f(\theta) \approx 2$ for $\theta < 0.57$, $f(\theta) \approx 0$ for $\theta > 1.43$, and $f(\theta)$ decreases monotonically for $0.57 \leq \theta \leq 1.43$, yielding an MSE value 0.0597. This implies that when the secrecy target is not very high, the deterministic generic encoding function does not actually behave like a non-uniform quantizer.

Proposition 4 can be utilized to derive the probability values for the stochastic affine strategy theoretically for given affine functions f_1 and f_2. For example, if the parameters of Fig. 4.8 are used in Proposition 4, it is obtained that $\nu_2 < 0$ for the MSEs of both the eavesdropper and the intended receiver, and according to the root test given in the proposition, the MSE decreases as γ increases when $\gamma \in [0, 1]$. For $\gamma = 0$, ϵ_e is found as $1/3 > \alpha_1 = 0.04$; hence, γ has to be increased until $\epsilon_e = \alpha_1 = 0.04$ to minimize ϵ_r. After some algebra, γ can be obtained as 0.8775.

We also provide an example for the problem of maximizing the MSE at the eavesdropper for a given estimation accuracy limit at the intended receiver (i.e., the optimization problems in (4.9) and (4.22)). In Fig 4.9, the MSE of the eavesdropper is plotted versus the SNR of the eavesdropper when the SNR of the intended receiver is 5 dB and the estimation accuracy limit α_2 is 0.24. It is observed that when the SNR of the eavesdropper is lower than the SNR of the intended receiver, all the solutions have the same performance; that is, using an optimal deterministic affine function is sufficient as claimed in Proposition 3. However, when the SNR of the eavesdropper increases, the MSE of the eavesdropper keeps decreasing for the deterministic affine strategy. Performing randomization over affine functions stops such a decline in the MSE and creates an MSE floor at the eavesdropper. Using generic functions yields even a higher MSE floor, where the stochastic approach performs slightly better than the deterministic one.

Finally, in Fig. 4.10, the MSE of the intended receiver (ϵ_r) is plotted versus the secrecy target α_1, and the MSE of the eavesdropper (ϵ_e) is plotted versus the estimation accuracy limit α_2 when the SNRs of the eavesdropper and the intended receiver are 5 and 15 dB, respectively. In this scenario, it is already established

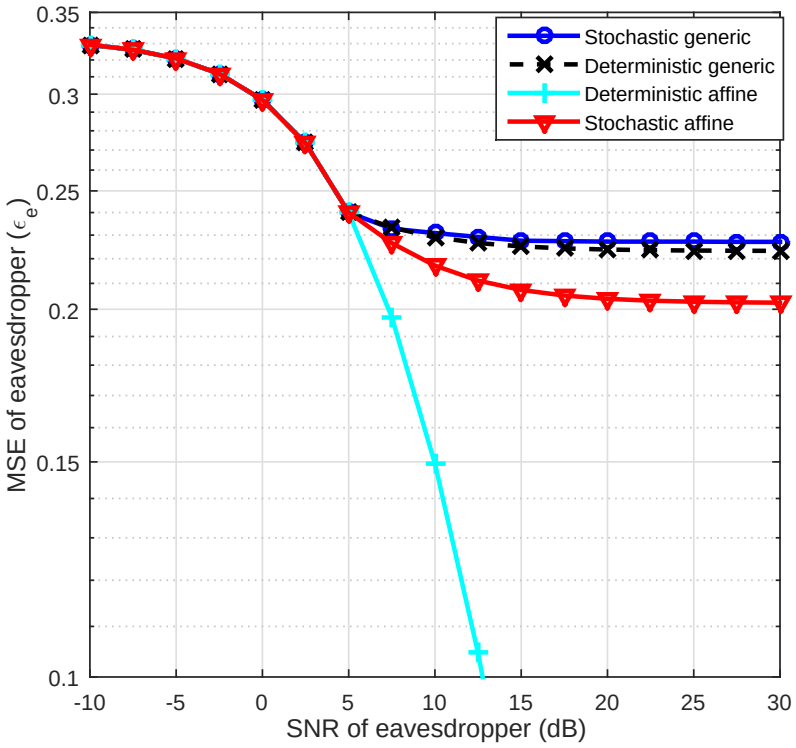

Figure 4.9: MSE of eavesdropper (ϵ_e) versus SNR of eavesdropper when SNR of intended receiver is 5 dB, and estimation accuracy limit α_2 is 0.24.

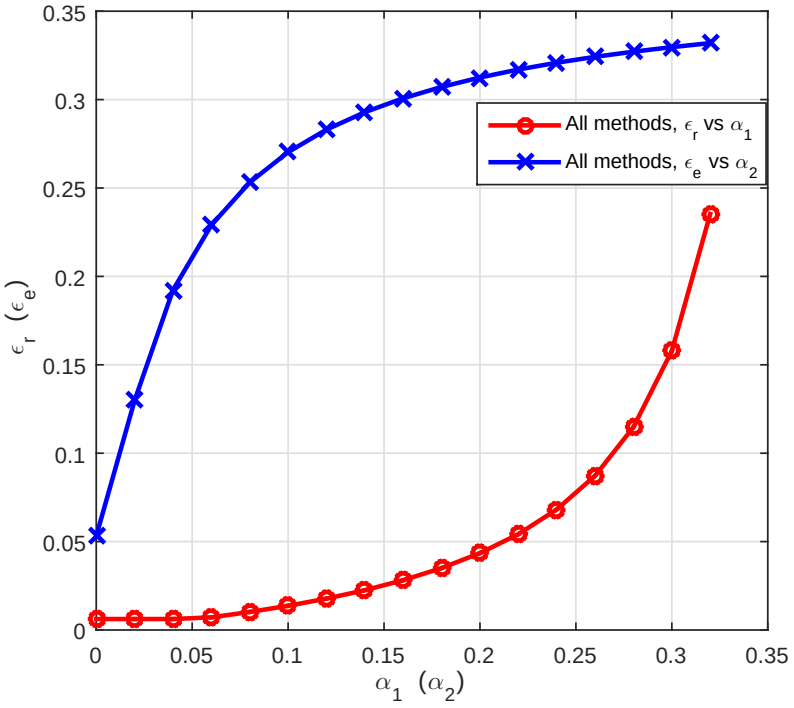

Figure 4.10: ϵ_r versus α_1 and ϵ_e versus α_2 when SNRs of eavesdropper and intended receiver are 5 and 15 dB, respectively.

that all the methods have the same performance. Note that ϵ_r can be kept at relatively low levels for $\alpha_1 < 0.2$; then, it increases rapidly as the secrecy demand becomes more ambitious. Also, ϵ_e increases at a high rate when α_2 is lower than 0.15, but further relaxing the estimation accuracy limit at the intended receiver does not bring significant benefits in terms of the MSE level at the eavesdropper.

It is noted that Proposition 2 and Corollary 1-4 establish the direct relationship between the optimization problems in (4.8) and (4.9). Also, based on Proposition 3, it has already been established that the conditions of Corollary 2 and 4 are satisfied when the SNR of the intended receiver is higher than the SNR of the eavesdropper; hence, their results can be applied. This can also be verified in Fig. 4.10. For example, given a secrecy level of $\alpha_1 = 0.2$, the minimum MSE value at the intended receiver is obtained as $\epsilon_r = 0.043$ after solving (4.8). Furthermore, for a given estimation accuracy limit of $\alpha_2 = 0.043$, the maximum MSE value at the eavesdropper becomes $\epsilon_e = 0.2$ after solving (4.9). A similar relationship is also observed when the SNR of the intended receiver is lower than the SNR of the eavesdropper according to Figs. 4.6 and 4.9.

4.4.3 Large Number of Observations

In this section, the numerical examples are provided for a large number of observations. In all the examples in this section, it is assumed that the number of observations is 1000, i.e., $n = 1000$. Similarly to the previous section, it is assumed that θ is uniformly distributed in $[0, 2]$ and the SNRs are defined in the same way. Also, the stochastic generic, stochastic affine and deterministic function strategies are evaluated in a similar fashion. The stochastic generic strategy corresponds to the solution of (4.33) and alternatively (4.34). The stochastic affine strategy also solves (4.33) or (4.34) with the additional assumption that the encoding functions are affine; that is, $f_1(\theta) = k_1\theta + k_2$ and $f_2(\theta) = s_1\theta + s_2$. Based on Proposition 5, there will be no deterministic affine and deterministic generic strategies separately in this section, and the solution of the deterministic strategy is directly evaluated via Proposition 5.

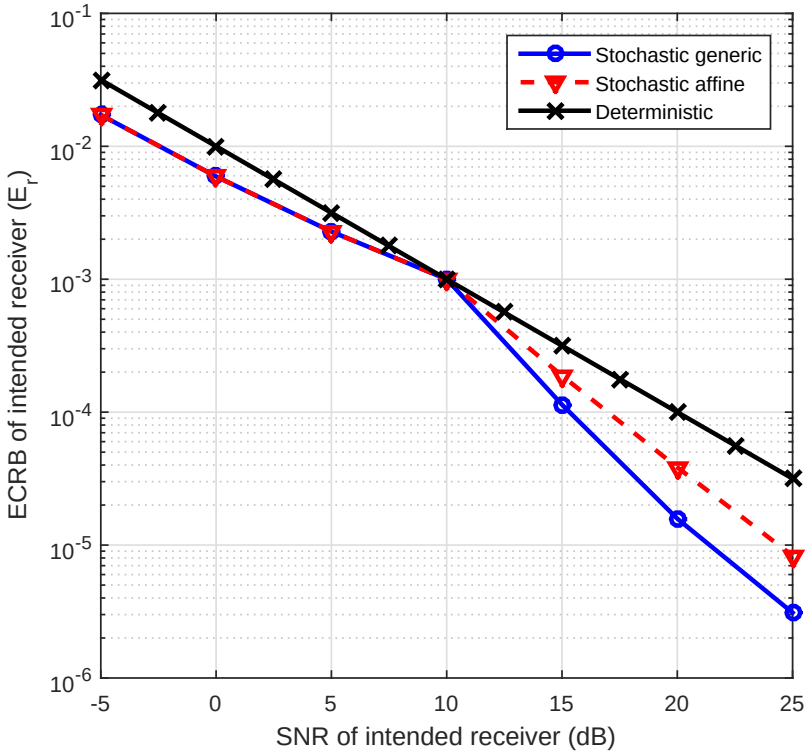

Figure 4.11: ECRB of intended receiver (E_r) versus SNR of intended receiver when SNR of eavesdropper is 10 dB, and target secrecy level η_1 is 0.001.

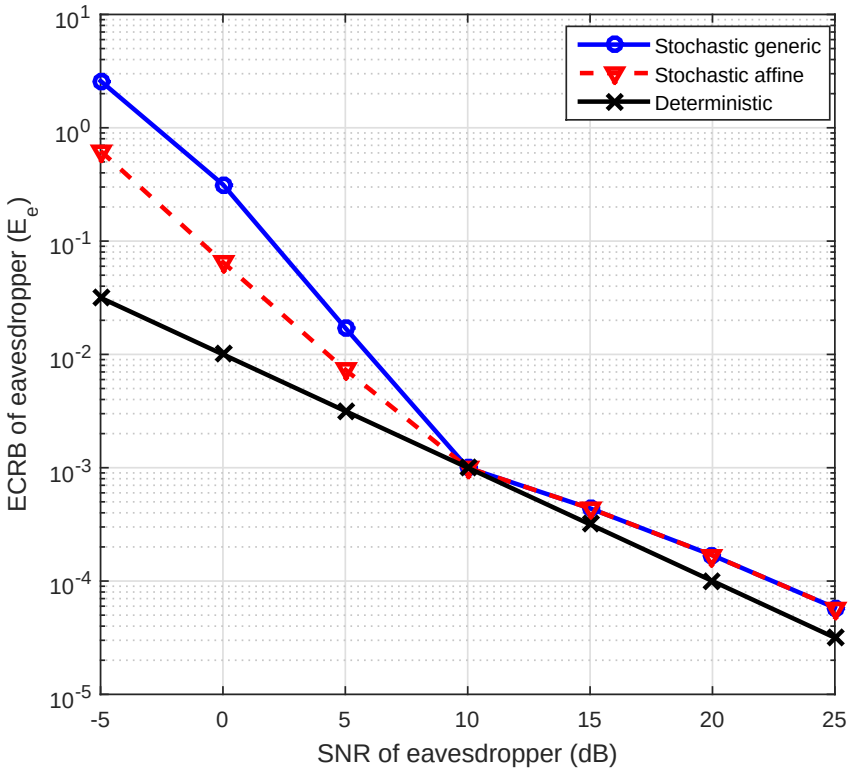

Figure 4.12: ECRB of eavesdropper (E_e) versus SNR of eavesdropper when SNR of intended receiver is 10 dB, and estimation accuracy limit η_2 is 0.001.

In this part, we consider the minimization (maximization) of the ECRB at the intended receiver (eavesdropper) for a given secrecy level (estimation accuracy limit) at the eavesdropper (intended receiver) in Figs. 4.11 and 4.13 (Figs. 4.12 and 4.14). First, the ECRB of the intended receiver (eavesdropper) is plotted versus the SNR of the intended receiver (eavesdropper) when the SNR of the eavesdropper (intended receiver) is 10 dB, and the secrecy target $\eta_1 = 0.001$ (and the estimation accuracy limit $\eta_2 = 0.001$). In Fig. 4.11 (Fig. 4.12), it is observed that the deterministic functions yield the worst performance and randomization is beneficial at all SNR values of the intended receiver (eavesdropper) for a large number of observations, which was not the case for a small number of observations. Note that the stochastic generic and affine functions have the same performance when the SNR of the intended receiver is lower than that of the eavesdropper. However, the stochastic generic functions outperform the stochastic affine functions when the SNR of intended receiver is higher than the SNR of the eavesdropper. Note that the ECRB versus SNR curve for the deterministic functions is a linear line as explained in Proposition 5. Also, the ECRB of the intended receiver (eavesdropper) is equal to η_1 (η_2) for all the strategies when the SNRs of the intended receiver and the eavesdropper are same.

Next, in Fig. 4.13 (Fig. 4.14), the ECRB of the intended receiver (eavesdropper) is plotted versus the secrecy target (estimation accuracy limit) for two different scenarios. In both scenarios, the SNR of the eavesdropper (receiver) is 10 dB and the SNR of the intended receiver (eavesdropper) is 5 and 20 dB in the first and second scenarios, respectively. In the first (second) scenario in Fig. 4.13 (Fig. 4.14), the performances of the stochastic strategies are almost the same and they are better than the deterministic solution. Furthermore, in the second (first) scenario in Fig. 4.13 (Fig. 4.14), the stochastic generic solution has better performance than the stochastic affine solution; hence it has the overall best performance. In that case, it is interesting to note that as η_1 (η_2) increases, the performance gap between the stochastic solutions and the simple deterministic solution increases, as well. This shows that randomization can bring significant performance improvements over the deterministic solution in the case of a large number of observations.

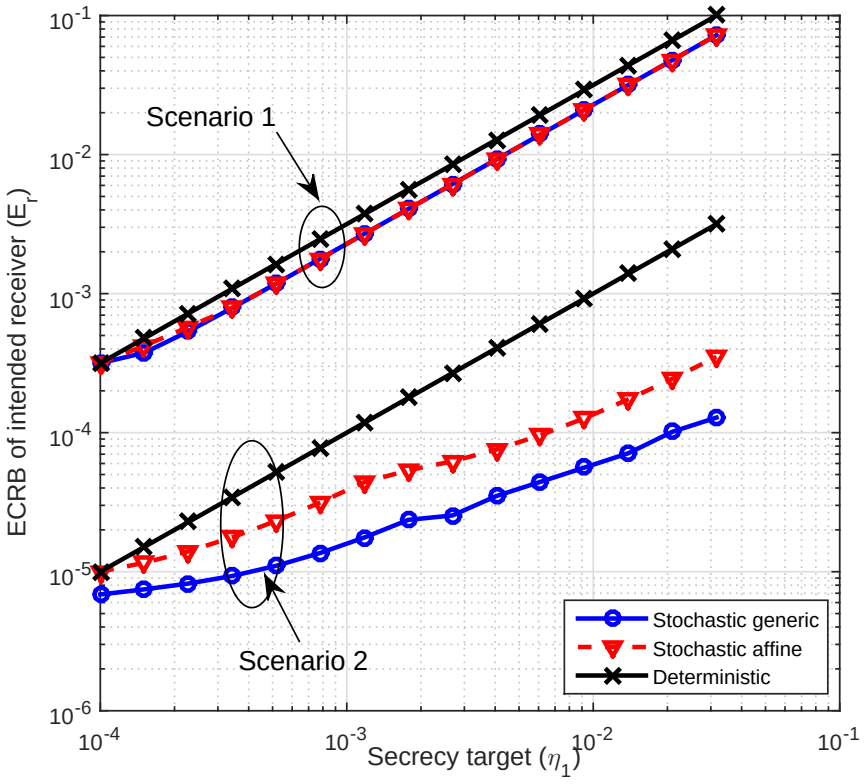

Figure 4.13: ECRB of intended receiver (E_r) versus secrecy target (η_1) for two different scenarios.

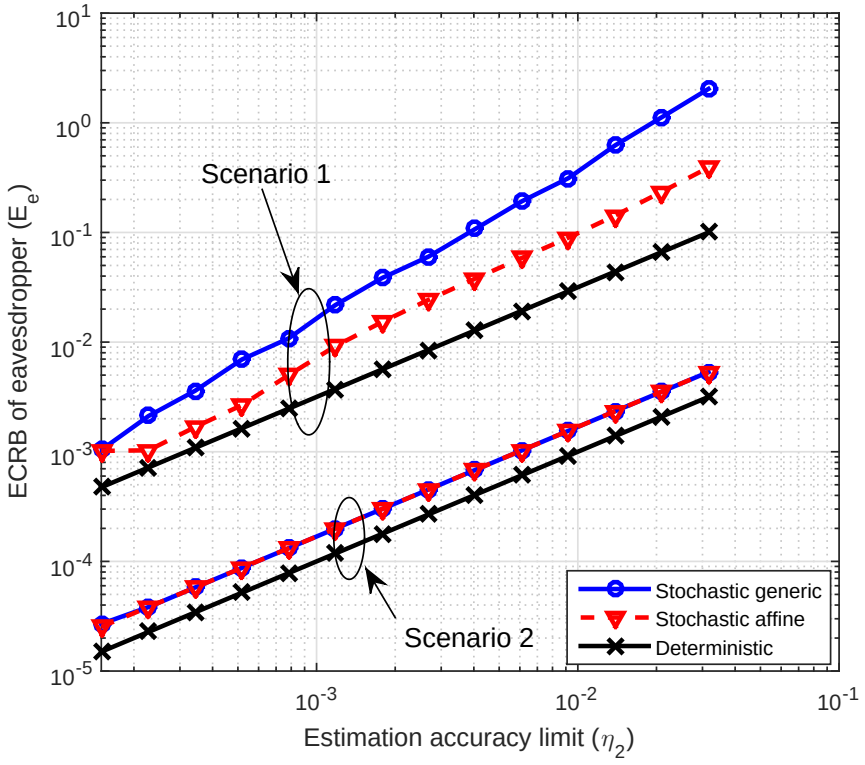

Figure 4.14: ECRB of eavesdropper (E_e) versus estimation accuracy limit (η_2) for two different scenarios.

Finally, Proposition 6 and 7 can be utilized in the numerical examples to further analyze the results. For example, in Fig. 4.11, when the SNR of the eavesdropper is 10 dB, and the SNR of the intended receiver is 15 dB, with $\eta_1 = 0.001$, the solution of the optimal stochastic affine encoding strategy is found as $f_1(\theta) = 0.4824\theta + 1.0352$, $f_2(\theta) = 0.9648 - 0.4824\theta$, and $\gamma = 0.5$. Note that according to Proposition 6, this is a symmetrical mapping; therefore, the ECRB of the eavesdropper is maximized at $\gamma = 0.5$. Also, as this encoding function satisfies the secrecy constraint with equality, Proposition 6 implies that other γ values would be infeasible for this particular f_1 and f_2. Also, again in Fig. 4.11, when the SNR of the intended receiver is 5 dB, the solution of the optimal stochastic affine encoding strategy is found as $f_1(\theta) = 0.4274\,\theta + 0.2597$, $f_2(\theta) = 0.4274\,\theta + 0.8989$, and $\gamma = 0.5$. In order to employ Proposition 7, it can be shown that $d_0 < 0$ and $d_1 > 0$ for all $\theta \in [0, 2]$. Actually, $I(\theta)$ is constant for a given γ for this given f_1 and f_2, and $\gamma = 0.5$ minimizes $I(\theta)$ (as it is constant, basically for all $\theta \in [a, b]$). Therefore, the ECRB of the eavesdropper is maximized at $\gamma = 0.5$.

It is important to mention that the closed-form expressions (e.g., Proposition 1, Corollary 5, and eqns. (4.26)–(4.30)) obtained in the theoretical parts (Sections 4.2 and 4.3) are used to calculate the LMMSE and ECRB values in the numerical examples. The performance of the theoretically optimal solutions (e.g., Proposition 3 and 5) is compared with the simulations for verification and the same performance results are obtained. However, the curves are not duplicated in the figures for brevity/clarity of presentation.

4.4.4 Computational Complexity

The dimension of the search space and the number of multiplications required to calculate the constraint and objective functions are both important factors about the complexity of the proposed methods. In the case of the stochastic generic function approach, the optimization is performed over $2M + 1$ variables, where M is the number of piecewise regions. For the deterministic generic solutions,

the optimization is over M variables. The affine solutions require optimization over five and two variables for the stochastic and deterministic cases, respectively. When Proposition 3 and 5 are utilized, no search is required and the solutions can be obtained directly. Also, the intuition provided by Proposition 4 can reduce the search space to four variables for the stochastic affine solutions in the small number of observations case. For large numbers of observations, the search space for the stochastic generic solution can be reduced to $2M$ based on Proposition 6 when the conditions of the proposition hold.

For small numbers of observations, we use the expressions in Proposition 1 to calculate the MSE. In the calculations, the most costly terms are the expectation terms such as $E(f_1(\theta)\,\theta)$ and $E(f_2(\theta)\,\theta)$. To calculate these terms, which include one-dimensional integrals, one of the possible ways is to employ Riemann sums, each of which includes S terms for a given step size. Then, when the stochastic and deterministic generic functions are used, calculating the objective function requires $O(14S)$ and $O(5S)$ multiplications, respectively. For the affine solutions, we do not have any of these terms, which implies a complexity of $O(1)$. As the only difference between the objective and constraint is the noise variance term, the complexity does not double for calculating both functions. It is important to note that the computational complexity does not depend on n.

For large numbers of observations, the overall expression requires double integration and complexity of $O(14S_1S_2)$, where the Riemann sums have S_1 and S_2 terms. Even though the ECRB calculation is more complex than calculating the MSE of the LMMSE estimator, it also does not depend on n. Note that the optimal MMSE expression would require $n+1$ integrals instead of two; hence, it is possible to tightly approximate the optimal MSE performance by using the ECRB with a much lower complexity. Finally, when the conditions of Corollaries 1-4 are satisfied, it is possible to connect the optimization problems in (4.8) and (4.9) (or, (4.33) and (4.34)) so that it is sufficient to solve one of the problems to obtain the solutions of both.

4.5 Concluding Remarks

Estimation theoretic secure transmission of a random scalar parameter has been investigated in a Gaussian wiretap channel model, and various constrained optimization problems have been proposed in terms of estimation accuracy performance of the intended receiver and the eavesdropper. The results have shown that for small numbers of observations, when the SNR of the intended receiver is higher than that of the eavesdropper, the deterministic affine solution forms a class of optimal functions, which verifies the theoretical results. When the SNR of the intended receiver is lower than that of the eavesdropper, stochastic generic functions have the best performance in general; however, depending on the target secrecy/accuracy value, deterministic generic functions can provide an optimal solution, as well. Stochastic affine functions can provide significant performance gains over deterministic affine functions, and they can be an attractive alternative solution to generic functions. For large numbers of observations, deterministic generic/affine functions have worse performance than stochastic solutions at all SNRs and in all the considered scenarios. Therefore, stochastic encoding is also attractive in this region of operation. Similarly to the previous case, stochastic generic functions have the best performance in general; however, stochastic affine functions can also provide an optimal solution in certain scenarios. Intuitively, the main factor that determines whether the stochastic methods bring performance gains or not is the quality and quantity of the measurements available to the eavesdropper given the secrecy target. If the eavesdropper has a large number of observations or a small number of observations with a better SNR than the intended receiver, then it is encoder's task to make estimation more challenging for the eavesdropper; hence, stochastic encoding provides performance gains especially in such scenarios. As a relevant future work, it would be interesting to investigate the MSE-based and information theoretically optimal solutions in a common and fair framework to provide theoretical comparisons and connections.

Chapter 5

Optimal Parameter Design for Estimation Theoretic Secure Broadcast

In this chapter, optimal parameter design for estimation theoretic secure broadcast is investigated, where each receiver device employs a fixed estimator and carries a certain security risk such that its decision can be available to a malicious third party with a certain probability [46]. The main contributions of this chapter can be summarized as follows:

- The optimal parameter design problem is formulated, where the encoder at the transmitter is allowed to use a random mapping to minimize the weighted sum of the conditional Bayes risks of the estimators under secrecy and average power constraints.
- It is shown that the optimization problem can be solved individually for each parameter value and the optimal mapping at the transmitter involves a randomization among at most three signal different levels.
- Sufficient conditions for improvability and nonimprovability of the deterministic design via stochastic encoding are obtained.

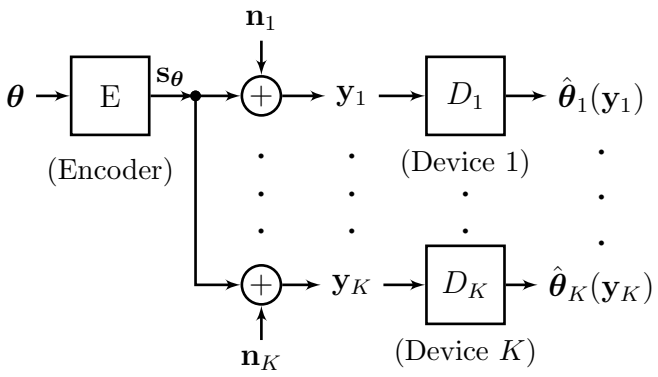

Figure 5.1: System model for the parameter encoding problem.

The rest of the chapter is organized as follows: In Section 5.1, the system model is introduced, optimal parameter design problem is formulated and its solution is analyzed. The numerical results and conclusions are provided in Section 5.2 and 5.3, respectively.

5.1 System Model and Optimal Parameter Design

Consider a system in which parameter $\boldsymbol{\theta} \in \Lambda$ is broadcasted to K different devices, where the channel for each device is modeled as an additive noise channel as in Fig. 5.1. The transmitter can send a random function of the parameter, that is, $\mathbf{s}_{\boldsymbol{\theta}}$, for each value of $\boldsymbol{\theta}$. Then, the received signal at the kth device can be written as

$$\mathbf{y}_k = \mathbf{s}_{\boldsymbol{\theta}} + \mathbf{n}_k \,, \tag{5.1}$$

where $\mathbf{n}_k$ denotes the channel noise, which has a generic probability density function (PDF) represented by $p_{\mathbf{n}_k}(\cdot)$. Also, the prior distribution of the parameter is denoted by $w(\boldsymbol{\theta})$, and $\mathbf{s}_{\boldsymbol{\theta}}$ and $\mathbf{n}_k$ are independent for all $\boldsymbol{\theta}$. It is assumed that each receiving device employs a fixed estimator $\hat{\boldsymbol{\theta}}_k(\mathbf{y}_k)$ based on their observation $\mathbf{y}_k$. (Note that the estimators of the devices can be different.) Also, each device

in the system has a certain assessed security risk probability γ_k to be compromised such that the estimate of the parameter at device k becomes available to a malicious third party with probability γ_k. It is important to emphasize that in the secrecy literature, the common assumption is that eavesdroppers employ optimal estimators/decoders to obtain the secret message since such an assumption (and knowledge of the encoding strategy at the eavesdropper in some scenarios) is required to obtain fundamental limits of secure communications. In our setting, it is assumed that the malicious parties may hijack the estimators/devices instead of designing their own, and it is assumed that the receivers in the systems are simple, low-complexity devices employing potentially suboptimal estimators. It is important to note that these two assumptions are independent. It means that the devices are not vulnerable due to their simplicity but due to possible proximity to adversarial attacks and security measures against them.

The main goal at the transmitter is to find the optimal probability distribution of $\mathbf{s}_{\boldsymbol{\theta}}$, that is, $p_{\mathbf{s}_{\boldsymbol{\theta}}}$, for each $\boldsymbol{\theta} \in \Lambda$ in order to minimize the weighted sum of Bayes risks of the estimators in the system under a security constraint on each value of $\boldsymbol{\theta}$. For a given value of $\boldsymbol{\theta}$, the conditional Bayes risk of the estimator at the kth device, $R_{\boldsymbol{\theta}}(\hat{\boldsymbol{\theta}}_k)$, is given by

$$R_{\boldsymbol{\theta}}(\hat{\boldsymbol{\theta}}_k) = \int C[\hat{\boldsymbol{\theta}}_k(\mathbf{y}_k), \boldsymbol{\theta}] p_{\boldsymbol{\theta}}(\mathbf{y}_k) d\mathbf{y}_k \,, \tag{5.2}$$

where $C[\hat{\boldsymbol{\theta}}_k(\mathbf{y}_k), \boldsymbol{\theta}] \geq 0$ represents a cost function [47], and $p_{\boldsymbol{\theta}}(\mathbf{y}_k)$ denotes the conditional PDF of $\mathbf{y}_k$ for a given value of parameter $\boldsymbol{\theta}$. Note that $p_{\boldsymbol{\theta}}(\mathbf{y}_k)$ can be expressed in terms of the PDF of $\mathbf{n}_k$ and the probability distribution of $\mathbf{s}_{\boldsymbol{\theta}}$ as $p_{\boldsymbol{\theta}}(\mathbf{y}_k) = \int p_{\mathbf{s}_{\boldsymbol{\theta}}}(\mathbf{x}) p_{\mathbf{n}_k}(\mathbf{y}_k - \mathbf{x}) d\mathbf{x}$ since $\mathbf{s}_{\boldsymbol{\theta}}$ and $\mathbf{n}_k$ are independent. Then, (5.2) becomes

$$\begin{aligned} R_{\boldsymbol{\theta}}(\hat{\boldsymbol{\theta}}_k) &= \int C[\hat{\boldsymbol{\theta}}_k(\mathbf{y}_k), \boldsymbol{\theta}] p_{\boldsymbol{\theta}}(\mathbf{y}_k) d\mathbf{y}_k = \\ &\int p_{\mathbf{s}_{\boldsymbol{\theta}}}(\mathbf{x}) \int C[\hat{\boldsymbol{\theta}}_k(\mathbf{y}_k), \boldsymbol{\theta}] p_{\mathbf{n}_k}(\mathbf{y}_k - \mathbf{x}) d\mathbf{y}_k d\mathbf{x} = E\{f_{\boldsymbol{\theta}}^{(k)}(\mathbf{s}_{\boldsymbol{\theta}})\} \end{aligned} \tag{5.3}$$

where $f_{\boldsymbol{\theta}}^{(k)}(\mathbf{x}) \triangleq \int C[\hat{\boldsymbol{\theta}}_k(\mathbf{y}_k), \boldsymbol{\theta}] p_{\mathbf{n}_k}(\mathbf{y}_k - \mathbf{x}) d\mathbf{y}_k$ and the expectation operator in (5.3) is over the PDF of $\mathbf{s}_{\boldsymbol{\theta}}$ for a given value of $\boldsymbol{\theta}$. Also, the Bayes risk of the

estimator at the kth device is given by

$$r(\hat{\boldsymbol{\theta}}_k) = \int_{\Lambda} w(\boldsymbol{\theta}) R_{\boldsymbol{\theta}}(\hat{\boldsymbol{\theta}}_k) d\boldsymbol{\theta}\,. \tag{5.4}$$

In order to measure the estimation performance of the whole system, we consider the weighted sum of Bayes risks of the estimators at the devices, where each Bayes risk is weighted by $c_k(\gamma_k)$, which is a non-negative scalar function of γ_k. Then, the objective function becomes $\sum_{k=1}^{K} c_k(\gamma_k) r(\hat{\boldsymbol{\theta}}_k)$, which is expressed, via (5.3) and (5.4), as

$$\begin{aligned}\sum_{k=1}^{K} c_k(\gamma_k) r(\hat{\boldsymbol{\theta}}_k) &= \int_{\Lambda} w(\boldsymbol{\theta}) \sum_{k=1}^{K} c_k(\gamma_k) E\{f_{\boldsymbol{\theta}}^{(k)}(\mathbf{s}_{\boldsymbol{\theta}})\} d\boldsymbol{\theta} = \\ \int_{\Lambda} w(\boldsymbol{\theta}) E\left\{\sum_{k=1}^{K} c_k(\gamma_k) f_{\boldsymbol{\theta}}^{(k)}(\mathbf{s}_{\boldsymbol{\theta}})\right\} d\boldsymbol{\theta} &= \int_{\Lambda} w(\boldsymbol{\theta}) E\{F_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}})\} d\boldsymbol{\theta}\end{aligned} \tag{5.5}$$

where $F_{\boldsymbol{\theta}}(\mathbf{x}) \triangleq \sum_{k=1}^{K} c_k(\gamma_k) f_{\boldsymbol{\theta}}^{(k)}(\mathbf{x})$.

Furthermore, the security constraint on each value of $\boldsymbol{\theta}$ is modeled as the weighted sum of the conditional Bayes risks of the estimators in the system, where each (non-negative) weight is denoted by $b_k(\gamma_k)$,[1] that is, $\sum_{k=1}^{K} b_k(\gamma_k) R_{\boldsymbol{\theta}}(\hat{\boldsymbol{\theta}}_k)$ $= \sum_{k=1}^{K} b_k(\gamma_k) E\{f_{\boldsymbol{\theta}}^{(k)}(\mathbf{s}_{\boldsymbol{\theta}})\}$. Then, the security constraint is in the form of

$$E\{G_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}})\} \geq \eta_{\boldsymbol{\theta}}\,, \quad \forall \boldsymbol{\theta} \in \Lambda \tag{5.6}$$

where $G_{\boldsymbol{\theta}}(\mathbf{x}) \triangleq \sum_{k=1}^{K} b_k(\gamma_k) f_{\boldsymbol{\theta}}^{(k)}(\mathbf{x})$, and $\eta_{\boldsymbol{\theta}}$ is the secrecy limit for each value of $\boldsymbol{\theta}$. In $G_{\boldsymbol{\theta}}(\mathbf{x})$, the estimation performance of the devices that are more likely to be compromised is prioritized as compared to that of the safer devices via proper weighting. The physical meaning behind the constraint in (5.6) is that the total estimation accuracy of the vulnerable, high-risk devices is limited by a security target. In practical systems, there is also an average power constraint

[1]It is reasonable to select $c_k(\gamma_k)$ to be a decreasing function of γ_k and $b_k(\gamma_k)$ to be increasing with γ_k. Two example selections for $c_k(\gamma_k)$ and $b_k(\gamma_k)$ are $c_k(\gamma_k) = 1 - \gamma_k$ and $b_k(\gamma_k) = \gamma_k$ or $c_k(\gamma_k) = 1\{\gamma_k < \tau\}$ and $b_k(\gamma_k) = 1\{\gamma_k \geq \tau\}$, where τ is the risk threshold and $1\{\cdot\}$ is the indicator function.

on the encoded version of the parameter in the form of $E\{||\mathbf{s}_{\boldsymbol{\theta}}||^2\} \leq A_{\boldsymbol{\theta}}$, where $||\mathbf{s}_{\boldsymbol{\theta}}||$ is the Euclidean norm of vector $\mathbf{s}_{\boldsymbol{\theta}}$ and $A_{\boldsymbol{\theta}}$ is the average power limit for $\boldsymbol{\theta}$. Therefore, based on (5.5) and (5.6), the optimal parameter design problem can be proposed as

$$\begin{aligned} &\min_{p_{\mathbf{s}_{\boldsymbol{\theta}}}, \boldsymbol{\theta}\in\Lambda} \int_{\Lambda} w(\boldsymbol{\theta}) E\{F_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}})\} d\boldsymbol{\theta} \quad \text{s.t.} \\ &E\{G_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}})\} \geq \eta_{\boldsymbol{\theta}}\,, \;\; E\{||\mathbf{s}_{\boldsymbol{\theta}}||^2\} \leq A_{\boldsymbol{\theta}}\,, \quad \forall \boldsymbol{\theta} \in \Lambda \end{aligned} \tag{5.7}$$

where $F_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}})$ and $G_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}})$ are as defined before. Note that as the constraints in (5.7) are defined for each value of $\boldsymbol{\theta}$, the optimization problem can be solved individually for each $\boldsymbol{\theta}$; hence, the solution does not depend on the prior distribution $w(\boldsymbol{\theta})$. In particular, (5.7) becomes

$$\min_{p_{\mathbf{s}_{\boldsymbol{\theta}}}} E\{F_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}})\} \text{ s.t. } E\{G_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}})\} \geq \eta_{\boldsymbol{\theta}}, \;\; E\{||\mathbf{s}_{\boldsymbol{\theta}}||^2\} \leq A_{\boldsymbol{\theta}} \tag{5.8}$$

for $\boldsymbol{\theta} \in \Lambda$. The optimization problems in the form of (5.8) have extensively been studied in the literature [58, 72, 73]. It can be shown that if $F_{\boldsymbol{\theta}}(\mathbf{x})$ and and $G_{\boldsymbol{\theta}}(\mathbf{x})$ are continuous and each component of $\mathbf{x}$ belongs to a finite closed interval, an optimal solution of (5.8) involves randomization among at most 3 different values of $\mathbf{s}_{\boldsymbol{\theta}}$ due to Carathéodory's theorem [74]. In general, if there were N_c constraints in (5.8) involving $E\{\tilde{H}^{(i)}_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}})\}$ for $i = 1, \ldots, N_c$ with continuous functions $\tilde{H}^{(i)}_{\boldsymbol{\theta}}$, then the solution of the optimization problem would involve randomization among at most $N_c + 1$ points. A proof of the statement for $N_c = 1$ and how Carathéodory's theorem is utilized is available in [58]. Hence, the optimal parameter design problem in (5.8) can be solved via the following problem:

$$\begin{aligned} &\min_{\{\lambda_{\boldsymbol{\theta},j}, \mathbf{s}_{\boldsymbol{\theta},j},\}_{j=1}^{3}} \quad \sum_{j=1}^{3} \lambda_{\boldsymbol{\theta},j} F_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta},j}) \quad \text{s.t.} \\ &\sum_{j=1}^{3} \lambda_{\boldsymbol{\theta},j} G_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta},j}) \geq \eta_{\boldsymbol{\theta}}, \quad \sum_{j=1}^{3} \lambda_{\boldsymbol{\theta},j} ||\mathbf{s}_{\boldsymbol{\theta},j}||^2 \leq A_{\boldsymbol{\theta}}, \\ &\sum_{j=1}^{3} \lambda_{\boldsymbol{\theta},j} = 1, \quad \lambda_{\boldsymbol{\theta},j} \in [0,1], \quad j = 1,2,3. \end{aligned} \tag{5.9}$$

It is noted that the optimization problem in (5.9) is much simpler to solve compared to (5.8) as it involves optimization over 6 variables instead of PDFs. In some cases, the optimal solution may not involve randomization and a deterministic solution can be sufficient to obtain the optimal solution. However, if the deterministic solution is improvable, this result implies that it is sufficient to randomize the signal by using at most 3 different levels.

The deterministic solution corresponds to the solution of the following problem:

$$\min_{\mathbf{s}_{\boldsymbol{\theta}}} \; F_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}}) \quad \text{s.t.} \;\; G_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}}) \geq \eta_{\boldsymbol{\theta}}, \;\; ||\mathbf{s}_{\boldsymbol{\theta}}||^2 \leq A_{\boldsymbol{\theta}}. \tag{5.10}$$

The deterministic solution is improvable by the stochastic solution if there exists $p_{\mathbf{s}_{\boldsymbol{\theta}}}$ such that $E\{F_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}})\} < F_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}}^{det})$ with $E\{G_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}})\} \geq \eta_{\boldsymbol{\theta}}$ and $E\{||\mathbf{s}_{\boldsymbol{\theta}}||^2\} \leq A_{\boldsymbol{\theta}}$.

Remark 1: The optimization problem in (5.9) turns out to be non-convex in most cases, and it is required to utilize global optimization techniques such as particle swarm optimization (PSO) or approximation techniques such as convex relaxation [58]. In this work, we utilize the Global Optimization Toolbox of MATLAB to obtain the solution of the optimization problems. For some specific cost functions and noise PDFs, the optimal solution can also be obtained directly. As an example in the scalar case, when the cost function is $C[\hat{\theta}_k(\mathbf{y}_k), \theta] = (\hat{\theta}_k(\mathbf{y}_k) - \theta)^2$ and the noise component is zero-mean, i.e., $E\{n_k\} = 0$ for all $k = 1, \ldots, K$, the problem simplifies and the solution can be obtained without global optimization techniques.

The following proposition provides a sufficient condition for the nonimprovability of the deterministic solution.

Proposition 1: *If $F_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}})$ is a convex and $G_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}})$ is a concave function of $\mathbf{s}_{\boldsymbol{\theta}}$ for each $\boldsymbol{\theta}$, then the deterministic solution cannot be improved via the stochastic solution.*

Proof: Due to Jensen's inequality, for any $\mathbf{s}_{\boldsymbol{\theta}}$, $||E\{\mathbf{s}_{\boldsymbol{\theta}}\}||^2 \leq E\{||\mathbf{s}_{\boldsymbol{\theta}}||^2\} \leq A_{\boldsymbol{\theta}}$, where the second inequality is due to the average power constraint. Therefore, for any feasible PDF of $\mathbf{s}_{\boldsymbol{\theta}}$ ($p_{\mathbf{s}_{\boldsymbol{\theta}}}$) for the problem in (5.8), $||E\{\mathbf{s}_{\boldsymbol{\theta}}\}||^2 \leq A_{\boldsymbol{\theta}}$. Similarly,

$\eta_{\boldsymbol{\theta}} \leq E\{G_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}})\} \leq G_{\boldsymbol{\theta}}(E\{\mathbf{s}_{\boldsymbol{\theta}}\})$ due to the concavity of $G_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}})$. Let $\mathbf{s}_{\boldsymbol{\theta}}^{\dagger} = E\{\mathbf{s}_{\boldsymbol{\theta}}\}$; therefore, for any feasible $p_{\mathbf{s}_{\boldsymbol{\theta}}}$, $||\mathbf{s}_{\boldsymbol{\theta}}^{\dagger}||^2 \leq A_{\boldsymbol{\theta}}$ and $G_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}}^{\dagger}) \geq \eta_{\boldsymbol{\theta}}$. As $\mathbf{s}_{\boldsymbol{\theta}}^{\dagger}$ is a feasible deterministic point, and $F_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}})$ is convex, $F_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}}^{det}) \leq F_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}}^{\dagger}) \leq E\{F_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}})\}$, where $\mathbf{s}_{\boldsymbol{\theta}}^{det}$ denotes the optimal deterministic solution. So, when $F_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}})$ is convex and $G_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}})$ is concave, $E\{F_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}})\}$ in (5.8) cannot be lower than the optimal value of (5.10) for any feasible PDF of $\mathbf{s}_{\boldsymbol{\theta}}$. ■

The main idea behind Proposition 1 is as follows: Under the conditions in the proposition, for any candidate stochastic solution of (5.8), we can obtain the deterministic solution $\mathbf{s}_{\boldsymbol{\theta}}^{\dagger} = E\{\mathbf{s}_{\boldsymbol{\theta}}\}$, which outperforms the stochastic solution and satisfies the constraints in (5.8). Note that if both of the constraints are removed from (5.8), it is easy to see that the optimal solution to the problem becomes $p_{\mathbf{s}_{\boldsymbol{\theta}}}^{\text{opt}}(\mathbf{x}) = \delta(\mathbf{x} - \mathbf{s}_{\boldsymbol{\theta}}^{\text{unc}})$, where $\mathbf{s}_{\boldsymbol{\theta}}^{\text{unc}} = \arg\min_{\mathbf{x}} \sum_{k=1}^{K} c_k(\gamma_k) f_{\boldsymbol{\theta}}^{(k)}(\mathbf{x})$. Another simple sufficient condition for nonimprovability of the deterministic solution can be expressed as follows: If the solution of the unconstrained problem satisfies the average power and secrecy constraints, i.e., $||\mathbf{s}_{\boldsymbol{\theta}}^{\text{unc}}||^2 \geq \eta_{\boldsymbol{\theta}}$ and $G_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}}^{\text{unc}}) \geq \eta_{\boldsymbol{\theta}}$, the optimal solution of (5.7) coincides with the unconstrained solution.

Next, a sufficient condition for the improvability of the deterministic solution is provided.

Proposition 2: *The deterministic solution can be improved via the stochastic solution for a given* $\boldsymbol{\theta} \in \Lambda$*, if there exists real vectors* $\mathbf{x}$ *and* $\mathbf{z}$ *such that* $F_{\boldsymbol{\theta}}(\mathbf{x})$ *and* $G_{\boldsymbol{\theta}}(\mathbf{x})$ *are differentiable around* $\mathbf{x}$*,* $||\mathbf{x}||^2 \leq A_{\boldsymbol{\theta}}$*, and the following inequality is satisfied:*

$$F_{\boldsymbol{\theta}}(\mathbf{x}) + \eta_{\boldsymbol{\theta}} \frac{\mathbf{z}^T \mathbf{H}_f \mathbf{z} - \frac{\mathbf{z}^T \tilde{\mathbf{f}}}{\mathbf{z}^T \mathbf{x}} ||\mathbf{z}||^2}{\mathbf{z}^T \mathbf{H}_g \mathbf{z} - \frac{\mathbf{z}^T \tilde{\mathbf{g}}}{\mathbf{z}^T \mathbf{x}} ||\mathbf{z}||^2} < F_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}}^{det}) \tag{5.11}$$

where $\mathbf{s}_{\boldsymbol{\theta}}^{det}$ *is the solution of* (5.10)*,* $\tilde{\mathbf{f}}$ *and* $\tilde{\mathbf{g}}$ *denote the gradients of* $F_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}})$ *and* $G_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}})$ *at* $\mathbf{s}_{\boldsymbol{\theta}} = \mathbf{x}$*, respectively, and* $\mathbf{H}_f$ *and* $\mathbf{H}_g$ *are the Hessian matrices of* $F_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}})$ *and* $G_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}})$ *at* $\mathbf{s}_{\boldsymbol{\theta}} = \mathbf{x}$*, respectively.*

Proof: Consider a value of $\boldsymbol{\theta}$ for which the conditions in the proposition are satisfied. The main goal is to show that the randomization around $\mathbf{x}$ can achieve

a strictly lower objective value than that of the deterministic solution while satisfying the constraints. Suppose that stochastic signaling involves randomization between two values, that is, $\mathbf{x}+\boldsymbol{\epsilon}_1$ and $\mathbf{x}+\boldsymbol{\epsilon}_2$. For sufficiently small $\boldsymbol{\epsilon}_1$ and $\boldsymbol{\epsilon}_2$, the following expressions can be written by using Taylor's series expansion around $\mathbf{s}_{\boldsymbol{\theta}} = \mathbf{x}$:

$$\begin{aligned} &||\mathbf{x}+\boldsymbol{\epsilon}_i||^2 \approx ||\mathbf{x}||^2 + 2\boldsymbol{\epsilon}_i^T\mathbf{x} + ||\boldsymbol{\epsilon}_i||^2 \\ &F_{\boldsymbol{\theta}}(\mathbf{x}+\boldsymbol{\epsilon}_i) \approx F_{\boldsymbol{\theta}}(\mathbf{x}) + 2\boldsymbol{\epsilon}_i^T\tilde{\mathbf{f}} + \boldsymbol{\epsilon}_i^T\mathbf{H}_f\boldsymbol{\epsilon}_i \\ &G_{\boldsymbol{\theta}}(\mathbf{x}+\boldsymbol{\epsilon}_i) \approx G_{\boldsymbol{\theta}}(\mathbf{x}) + 2\boldsymbol{\epsilon}_i^T\tilde{\mathbf{g}} + \boldsymbol{\epsilon}_i^T\mathbf{H}_g\boldsymbol{\epsilon}_i \end{aligned} \tag{5.12}$$

for $i = 1, 2$. In order to show that the stochastic solution with PDF $p_{\mathbf{s}_{\boldsymbol{\theta}}}(\mathbf{s}_{\boldsymbol{\theta}}) = \lambda\delta(\mathbf{s}_{\boldsymbol{\theta}} - (\mathbf{x}+\boldsymbol{\epsilon}_1)) + (1-\lambda)\delta(\mathbf{s}_{\boldsymbol{\theta}} - (\mathbf{x}+\boldsymbol{\epsilon}_2))$ improves the deterministic solution, it is sufficient to satisfy the following conditions:

$$\begin{aligned} &\lambda||\mathbf{x}+\boldsymbol{\epsilon}_1||^2 + (1-\lambda)||\mathbf{x}+\boldsymbol{\epsilon}_2||^2 = ||\mathbf{x}||^2 \leq A_{\boldsymbol{\theta}} \\ &\lambda F_{\boldsymbol{\theta}}(\mathbf{x}+\boldsymbol{\epsilon}_1) + (1-\lambda)F_{\boldsymbol{\theta}}(\mathbf{x}+\boldsymbol{\epsilon}_2) < F_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}}^{det}) \\ &\lambda G_{\boldsymbol{\theta}}(\mathbf{x}+\boldsymbol{\epsilon}_1) + (1-\lambda)G_{\boldsymbol{\theta}}(\mathbf{x}+\boldsymbol{\epsilon}_2) = \eta_{\boldsymbol{\theta}} \end{aligned} \tag{5.13}$$

If we insert the relations in (5.12) into (5.13) and let $\boldsymbol{\epsilon}_1 = \alpha\mathbf{z}$ and $\boldsymbol{\epsilon}_2 = \beta\mathbf{z}$, the sufficient conditions in (5.13) can be expressed as

$$\begin{aligned} &k\mathbf{z}^T\mathbf{x} = -||\mathbf{z}||^2 \\ &k\mathbf{z}^T\tilde{\mathbf{f}} + \mathbf{z}^T\mathbf{H}_f\mathbf{z} < (F_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}}^{det}) - F_{\boldsymbol{\theta}}(\mathbf{x}))/k_2 \\ &k\mathbf{z}^T\tilde{\mathbf{g}} + \mathbf{z}^T\mathbf{H}_g\mathbf{z} = \eta_{\boldsymbol{\theta}}/k_2 \end{aligned} \tag{5.14}$$

where $k = k_1/k_2$ with $k_1 = 2(\lambda\alpha + (1-\lambda)\beta)$ and $k_2 = \lambda\alpha^2 + (1-\lambda)\beta^2$. Also note that $k = -||\mathbf{z}||^2/\mathbf{z}^T\mathbf{x}$ and $k_2 = \eta_{\boldsymbol{\theta}}/(k\mathbf{z}^T\tilde{\mathbf{g}} + \mathbf{z}^T\mathbf{H}_g\mathbf{z})$ due to the first and the third equalities in (5.14). If they are inserted in the second inequality, the sufficient condition corresponds to one given in (5.11). ■

The idea in Proposition 2 is to provide conditions under which randomization around a real vector leads to an improvement over the optimal deterministic solution. To illustrate this idea, consider a scenario in which $F_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}})$ and

$G_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}})$ are both convex functions. In this scenario, randomization increases the value of the objective and secrecy functions due to Jensen's inequality. Therefore, suppose that there exists a point $\mathbf{x}$ satisfying the average power constraint with $F_{\boldsymbol{\theta}}(\mathbf{x}) < F_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}}^{det})$ and $G_{\boldsymbol{\theta}}(\mathbf{x}) < \eta_{\boldsymbol{\theta}}$. The condition in Proposition 2 implies that randomization around $\mathbf{x}$ ensures that the security constraint is satisfied, i.e., $E\{G_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}})\} = \eta_{\boldsymbol{\theta}}$, while the increase in the objective value due to randomization is still sufficiently small to improve the deterministic solution, i.e., $E\{F_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}})\} < F_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}}^{det})$. Also, even though the derivation of Proposition 2 is based on the similar idea and techniques presented in [58], we manage to reduce the number of equations in the sufficient condition compared to [58], by allowing the randomization to be around any feasible point $\mathbf{x}$ with $||\mathbf{x}||^2 \leq A_{\boldsymbol{\theta}}$ and letting the candidate stochastic solution satisfy the secrecy constraint with equality.

As Proposition 2 does not provide a necessary condition, it is possible that the improvement over the deterministic solution via the stochastic solution can be observed even if the condition in the proposition is not satisfied. However, if the condition is satisfied, then it is guaranteed that the solution of (5.10) is not optimal and the problem has to be solved based on (5.9).

It is also interesting to obtain the solution of the problem with only the secrecy constraint. In the following, the special case without the power constraint is investigated.

Special Case with No Average Power Constraint: As a special case, we consider the problem in (5.8) with only the secrecy constraint. (This case has practical relevance as the secrecy constraint also implies a certain limit on the average transmit power in practice.) In this case, the optimization problem can be expressed as

$$\min_{p_{\mathbf{s}_{\boldsymbol{\theta}}}} E\{F_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}})\} \text{ s.t. } E\{G_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta}})\} \geq \eta_{\boldsymbol{\theta}} \tag{5.15}$$

The solution of (5.15) involves randomization between at most 2 different values;

hence, the simplified problem becomes

$$\min_{\{\lambda,\mathbf{s}_{\boldsymbol{\theta},1},\mathbf{s}_{\boldsymbol{\theta},2}\}} \lambda F_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta},1}) + (1-\lambda)F_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta},2})$$
$$\text{s.t.} \quad \lambda G_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta},1}) + (1-\lambda)G_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta},2}) \geq \eta_{\boldsymbol{\theta}}, \quad \lambda \in [0,1]. \tag{5.16}$$

Note that for given $(\mathbf{s}_{\boldsymbol{\theta},1}, \mathbf{s}_{\boldsymbol{\theta},2})$, finding the optimal λ, λ^{opt}, from (5.16) is straightforward as all the functions can be calculated directly and have scalar real values. If $\max\{G_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta},1}), G_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta},2})\} \leq \eta_{\boldsymbol{\theta}}$, then $(\mathbf{s}_{\boldsymbol{\theta},1}, \mathbf{s}_{\boldsymbol{\theta},2})$ is not a feasible pair. If $\min\{G_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta},1}), G_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta},2})\} \geq \eta_{\boldsymbol{\theta}}$, then $\lambda^{\text{opt}} = 1$ if $F_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta},1}) < F_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta},2})$ and $\lambda^{\text{opt}} = 0$ otherwise. If $G_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta},2}) \leq \eta_{\boldsymbol{\theta}} \leq G_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta},1})$, then $\lambda^{\text{opt}} = 1$ if $F_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta},1}) < F_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta},2})$ and $\lambda^{\text{opt}} = (\eta_{\boldsymbol{\theta}} - G_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta},2}))/(G_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta},1}) - G_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta},2}))$ otherwise. The result for the case of $G_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta},1}) \leq \eta_{\boldsymbol{\theta}} \leq G_{\boldsymbol{\theta}}(\mathbf{s}_{\boldsymbol{\theta},2})$ can be obtained similarly. There are important implications of these possibilities. First, it is observed that as the optimal λ can theoretically be found for given $(\mathbf{s}_{\boldsymbol{\theta},1}, \mathbf{s}_{\boldsymbol{\theta},2})$, the search for the optimal solution of the problem in (5.16) can be performed over two variables, i.e., $\mathbf{s}_{\boldsymbol{\theta},1}$ and $\mathbf{s}_{\boldsymbol{\theta},2}$. Second, if the optimal solution involves randomization, that is, $0 < \lambda^{\text{opt}} < 1$, then the secrecy constraint is satisfied with equality.[2] This implies that any stochastic solution, which is in the interior region of the feasible set, can be improved by pushing it to the boundary and meeting at least one of the constraints.

Remark 2: The problem without the secrecy constraint but with average power constraint is studied in [58], and it is shown that the optimal solution involves a randomization among at most 2 different signal values and if stochastic design is optimal, then the solution operates at the average power limit. So by combining this and the result given previously, it is possible to claim that if the optimal solution to original problem given in (5.8) involves randomization, then at least one of the constraints (average power or secrecy) should be satisfied with equality. This implies that any stochastic solution, which is in the interior region of the feasible set, can be improved by pushing it to the boundary and meeting

[2]The problem in (5.8) without the secrecy constraint is studied in [58], which shows that the optimal solution involves randomization between at most two different signal values and that the solution operates at the average power limit if the stochastic design is optimal. Combining this result with the one in this chapter, we claim that if the optimal solution to the original problem in (5.8) involves randomization, then at least one of the constraints (average power or secrecy) should be satisfied with equality.

at least one of the constraints.

5.2 Numerical Results

In the numerical examples, we consider the transmission of a scalar parameter θ to $K = 5$ devices, all of which employ the fixed estimator given by $\hat{\theta}_k(y_k) = y_k$ and the cost function is selected as the squared error function, i.e., $C[\hat{\theta}_k(\mathbf{y}_k), \theta] = (\hat{\theta}_k(\mathbf{y}_k) - \theta)^2$ for $k = 1, \ldots, K$. The security risk probabilities, γ_k's, of the devices are assessed as $(1, 0.75, 0.5, 0.25, 0)$ and $c_k(\gamma_k) = 1 - \gamma_k$ and $b_k(\gamma_k) = \gamma_k$ are used. The noise of the kth device (user) is modeled by Gaussian mixture noise with two mass points such that its PDF is given by $p_{n_k}(x) = \nu_k e^{-\frac{(x-\mu_k)^2}{\sigma^2}}/\sqrt{2\pi\sigma^2} + (1-\nu_k)e^{-\frac{(x+\mu_k)^2}{\sigma^2}}/\sqrt{2\pi\sigma^2}$, and the noise parameters are taken as $\boldsymbol{\nu} = [0.2,\ 0.25,\ 0.3,\ 0.4,\ 0.9]$ and $\boldsymbol{\mu} = [0,\ 0.4,\ 0.8,\ 1.2,\ 1.6]$, where $\boldsymbol{\nu} = [\nu_1,\ \nu_2,\ \nu_3,\ \nu_4,\ \nu_5]$ and $\boldsymbol{\mu} = [\mu_1,\ \mu_2,\ \mu_3,\ \mu_4,\ \mu_5]$. In the examples, the stochastic and deterministic solutions are considered for the original problem with the average power and secrecy constraints in (5.9) and (5.10), respectively. Also, the performance results are presented when there exists only the average power constraint, only the secrecy constraint and no constraints for comparison purposes as they yield various lower bounds for the original problem.

In Fig. 5.2, the weighted sum of the conditional Bayes risks (i.e., $E\{F_\theta(s_\theta)\}$) is plotted versus $1/\sigma^2$ for $\theta = 1$, $\eta_\theta = 2$, and $A_\theta = 1$. It is observed that the stochastic solution improves the deterministic solution especially for lower values of σ^2. It is also noted that the stochastic and deterministic parameter designs have the same performance when one of the constraints is removed, and their performance is close to the unconstrained solution (which is the solution of (5.8) in the absence of the constraints) in this particular scenario. However, when both of the constraints are imposed, the performance of the deterministic design starts to deteriorate severely in the low σ^2 region due to the interference components present in the Gaussian mixture noise. The randomization via stochastic signaling alleviates the effects of the interference resulting in an improved solution

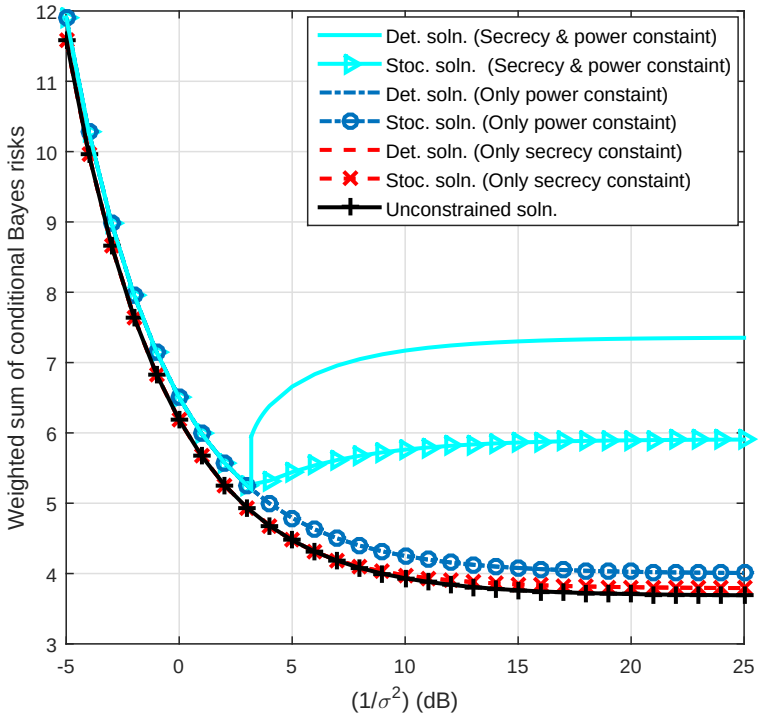

Figure 5.2: Weighted sum of conditional Bayes risks versus $1/\sigma^2$.

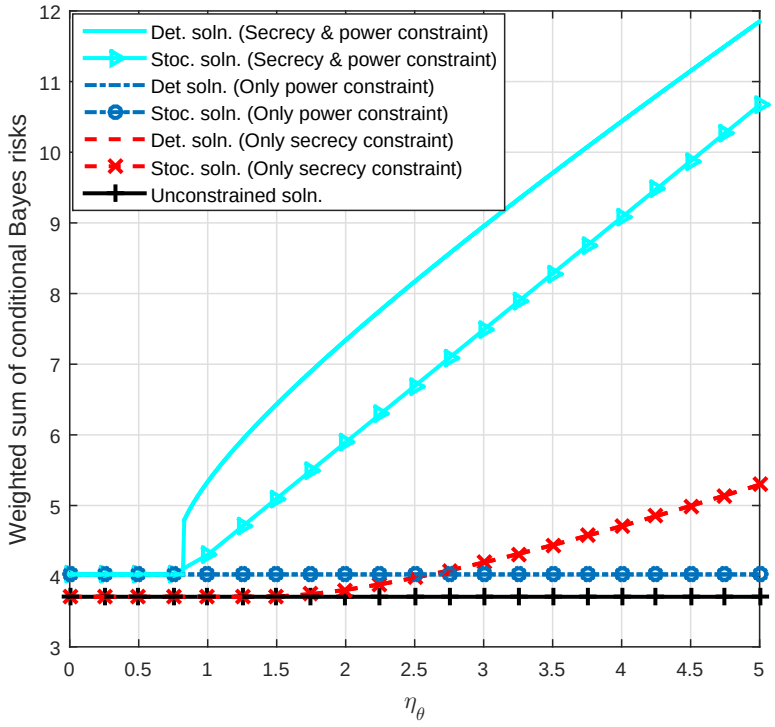

Figure 5.3: Weighted sum of conditional Bayes risks versus η_θ.

$1/\sigma^2$	λ_θ	$s_{\theta,1}$	$s_{\theta,2}$	s_θ^{det}	s_θ^{unc}	s_θ^{pow}	s_θ^{sec}
-5dB	1	1	N/A	1	1.356	1	1.356
0dB	1	1	N/A	1	1.356	1	1.356
5dB	0.421	0.403	1.269	0.421	1.356	1	1.356
15dB	0.761	1.116	-0.468	0.166	1.356	1	1.538
25dB	0.701	1.164	-0.365	0.146	1.356	1	1.558

Table 5.1: The solutions for various approaches when $\eta_\theta = 2$.

compared to the deterministic one. Although the deterministic solution is an attractive alternative due to its simplicity and achieves the optimal solution when the noise variance is large, there is no guarantee that the deterministic solution is a good approximation to the optimal stochastic solution in general.

In Fig. 5.3, the weighted sum of the conditional Bayes risks is plotted versus η_θ for $\theta = 1$, $1/\sigma^2 = 20$dB, and $A_\theta = 1$. When both of the constraints are considered, it is observed that the Bayes risks start to rise as the security demand increases, especially for $\eta_\theta \geq 0.75$, and the stochastic solution improves the deterministic solution similarly to Fig. 5.2. Note that the unconstrained solution and the solution of the problem with only the average power constraint is constant as they do not consider the secrecy constraint. The solution of the unconstrained problem (s_θ^{unc}) satisfies the secrecy constraint until a certain point ($\eta_\theta \approx 1.75$); however, for larger η_θ, the secrecy constraint becomes effective leading to a slight increase in the Bayes risks. When one or both of the constraints are removed, the deterministic and stochastic designs have the same performance similarly to Fig. 5.2.

In Table 5.1, the solutions of the problems for the stochastic and deterministic approaches with both constraints, with only the power constraint, and with only the secrecy constraint are presented together with the unconstrained solution. In this example, the stochastic and deterministic solutions coincide when there is only the power (or secrecy) constraint (see the last two columns in the table). In addition, for the stochastic solution with both constraints, it is observed that the optimal solution can be attained by using two signal levels; hence, it is not always necessary to randomize among three signal levels. Moreover, it is noticed that the sudden increases (jumps) in the Bayes risks for the deterministic solution in

both Figs. 5.2 and 5.3 occur exactly when $s_\theta^{det} = 1$ becomes infeasible.

The improvement via stochastic signaling can theoretically be justified based on Proposition 2 when both constraints are considered. For example, when $\eta_\theta = 2$ and $1/\sigma^2 = 20$ dB, the optimal deterministic solution, s_θ^{det} is 0.151 yielding 7.340 as the weighted sum of Bayes risks. The inequality condition given in (5.11) can be expressed as $F_{cond}(x) < F_\theta(s_\theta^{det}) = 7.340$ where $F_{cond}(x) = F_\theta(x) + \eta_\theta \frac{xF_\theta^{''}(x) - F_\theta^{'}(x)}{xH_\theta^{''}(x) - H_\theta^{'}(x)}$ and $F_\theta^{''}(x) = \frac{dF_\theta^2(x)}{dx^2}$, $F_\theta^{'}(x) = \frac{dF_\theta(x)}{dx}$, $H_\theta^{''}(x) = \frac{dH_\theta^2(x)}{dx^2}$, and $H_\theta^{'}(x) = \frac{dH_\theta(x)}{dx}$. It is noted that the resulting condition is independent of z, hence we can simply select $z = 1$. It can be observed in Fig. 5.4 that $F_{cond}(x) < F_\theta(s_\theta^{det})$ for any $x \in [0.548, 1]$; therefore, the inequality condition given in Proposition 2 is satisfied. Since this is a sufficient condition for improvability, it is known that the deterministic solution can be improved via the stochastic solution as stated in Proposition 2. In fact, the stochastic solution represented by $p_{s_\theta}(s_\theta) = \lambda_\theta\, \delta(s_\theta - s_{\theta,1}) + (1 - \lambda_\theta)\, \delta(s_\theta - s_{\theta,2})$ with $\lambda_\theta = 0.693$, $s_{\theta,1} = 1.196$, and $s_{\theta,2} = -0.168$ yields 5.762 as the weighted sum of Bayes risks. Note that for this solution $E\{s_\theta\} = 0.777$; hence, randomization around this point improves the deterministic solution as predicted by Proposition 2. In general, one possible way to check the sufficient condition in Proposition 2 is to fix $\mathbf{z}$, and then perform the search over $\mathbf{x}$ in the closed ball $||\mathbf{x}||^2 \leq A_{\boldsymbol{\theta}}$ while checking the inequality in (5.11). If there is no $\mathbf{x}$ satisfying the condition for a given $\mathbf{z}$, then another $\mathbf{z}$ can be selected until a preset number of maximum trials is reached.

The second example is provided when the cost function is modeled as the uniform cost function, that is, $C[\hat{\theta}_k(\mathbf{y}_k), \theta] = 1$ if $|\hat{\theta}_k(\mathbf{y}_k) - \theta| > \Delta$ and it is 0 otherwise, where $\Delta = 0.25$ is used in the example. The noise parameters are set to $\mu = [0,\ 1,\ 1.25,\ 1.5,\ 2.5]$ and $\nu_k = 0.5$ for $k = 1, \ldots, 5$; hence, the received signal has zero-mean total noise and interference terms for all users. The rest of the parameters are the same as in the first example. In Fig. 5.5, the weighted sum of conditional Bayes risks is plotted versus $1/\sigma^2$ when $\theta = 1$, $\eta_\theta = 1.1$ and $A_\theta = 1$. It is observed that all the solutions have the same performance when $1/\sigma^2 < 5$ dB. After that point, the stochastic solution improves the performance of the deterministic one for the case with both constraints and with only the secrecy

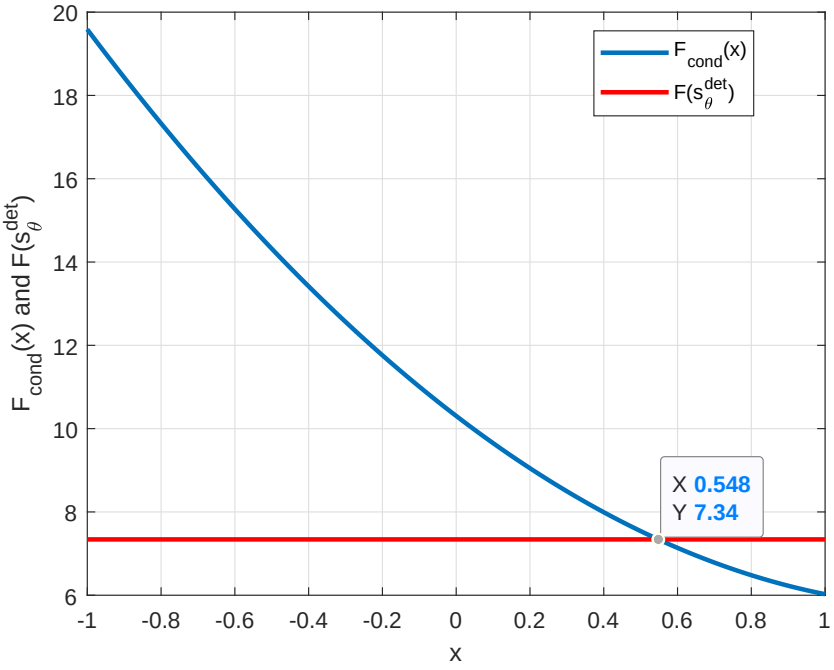

Figure 5.4: For $x \in [0.548, 1]$, $F_{cond}(x) < F_\theta(s_\theta^{det}) = 7.340$.

constraint. It is observed that the unconstrained solution also satisfies the power constraints at all points. In Fig. 5.6, the weighted sum of conditional Bayes risks is plotted versus $\theta \in [0.2, 2]$, when $1/\sigma^2 = 20$ dB and $A_\theta = \theta^2$. It is noted that the stochastic solution provides better performance than the deterministic one for all θ values considered in the figure, when both constraints are taken into account or when there is only the secrecy constraint. For the case with only the power constraint, the stochastic solution also improves the performance of the deterministic one when θ is small ($\theta < 0.5$). As θ increases, the solution for the case with the power constraint converges to that of the unconstrained case, and the solution for the case with both constraints converges to that of the case with only the secrecy constraint. This is due to the fact that as θ increases A_θ also increases and the power constraint starts to be relaxed.

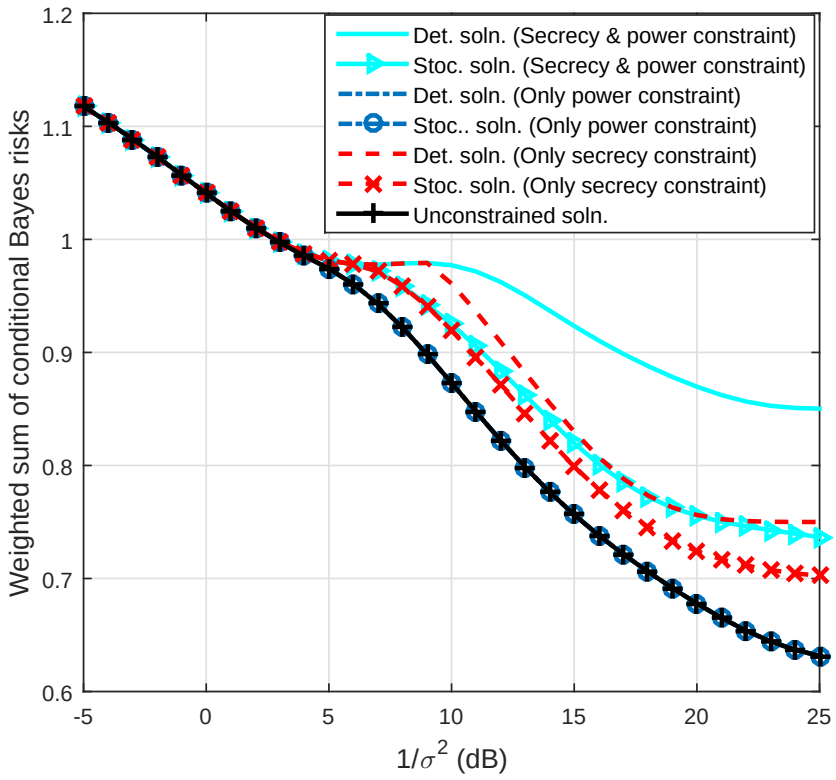

Figure 5.5: Weighted sum of conditional Bayes risks versus $1/\sigma^2$ for different scenarios.

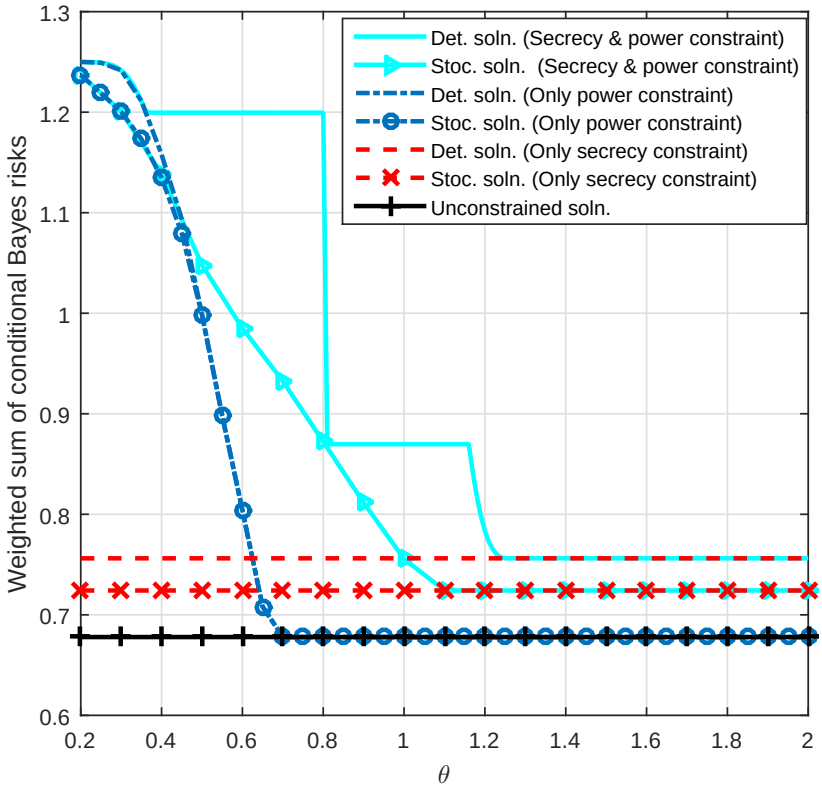

Figure 5.6: Weighted sum of conditional Bayes risks versus θ for different scenarios.

5.3 Concluding Remarks

In this chapter, we have proposed an optimal stochastic parameter design problem in the presence of both secrecy and average power constraints by considering a broadcast scenario. We have characterized the optimal solution and provided sufficient conditions for improvability and non-improvability of the deterministic approach via stochastic signaling. Numerical examples have illustrated the benefits of the proposed approach in various scenarios.

Chapter 6

Conclusion and Future Work

Estimation theoretic security has attracted a lot of interest with the advances in the wireless communications technology as smart grids, Internet of Things and wireless sensor networks have turned into practical realities. Theoretical investigation of such a framework is not only important to understand and obtain the achievable estimation performance under secrecy constraints in parameter estimation problems, but it also provides a practical and low-complexity approach for providing an additional/alternative layer of security for communication systems. Based on this motivation, we have investigated optimal encoding strategies to ensure estimation theoretic secure communications under different assumptions and transmission scenarios in this dissertation.

In Chapter 2, we have investigated the optimal deterministic encoding of a random scalar parameter in the presence of an eavesdropper, which is modeled to be unaware of the encoding operation. The optimization problems have been formulated to optimize the estimation accuracy based on ECRB and alternatively worst-case Fisher information at the intended receiver, while keeping the MSE at the eavesdropper above a certain level. For both design approaches, we have obtained the optimal encoding functions in closed-form when the secrecy constraint is removed. The theoretical results on the form of the optimal encoding

functions have been derived to gain intuitive understanding of the parameter encoding problem, when the secrecy constraint is activated. Furthermore, we have provided various practical solution approaches and algorithms to obtain encoding functions.

In Chapter 3, optimal deterministic encoding strategies for a random vector parameter have been considered under secrecy constraints. The main goal has been specified as the minimization of the ECRB at the intended receiver, while satisfying an individual secrecy constraint on the MSE of estimating each parameter at the eavesdropper. A generic optimization problem has been formulated as a constrained optimization problem in the space of vector-valued functions. Two practical solution strategies have been developed based on nonlinear individual encoding and affine joint encoding of parameters. We have provided theoretical results on the solutions of the proposed strategies for various channel conditions and parameter distributions. It has been observed that the proposed approaches can provide a certain estimation quality at the receiver and create large estimation errors at the eavesdropper, which could not have been possible if there were no encoding utilized.

In Chapter 4, we have investigated estimation theoretic secure transmission of a random scalar parameter via stochastic encoding in a Gaussian wiretap channel model. In particular, the encoder at the transmitter has been modeled to perform randomization between two one-to-one and continuous encoding functions, which should also be designed. We have investigated the minimization of estimation error at the intended receiver under a secrecy constraint at the eavesdropper, and the maximization of the estimation error at the eavesdropper under an estimation accuracy limit at the receiver. It has been assumed that both the intended receiver and the eavesdropper have the exact knowledge of the encoding strategy in the transmitter. We have argued that the MSE of the LMMSE estimator and the ECRB are good approximations to the MSE of the optimal MMSE estimator for small and large numbers of observations; hence, the optimization problems have been formulated based on these metrics depending on the available number of observations. We have provided theoretical results on the effects of randomization and the closed-form solutions for specific scenarios for both cases. It has been

observed that randomization can bring significant performance gains compared to deterministic encoding.

In Chapter 5, we have focused on the optimal stochastic parameter design for secure broadcast of a random parameter to multiple receivers with fixed estimators. It has been assumed that each receiver can be compromised with a certain risk probability, and the optimization problem has been formulated to minimize the weighted sum of conditional Bayes risks of the estimators under secrecy and average power constraints. We have showed that the optimization problem decouples into individual optimization problems, where the optimal parameter design can be performed for each parameter value. We have also argued that the optimal mapping for each value of the parameter involves randomization among at most three different signal levels based on Carathéodory's theorem. We have provided sufficient conditions for the improbability and nonimprovabilility of the deterministic encoding via the stochastic one.

As future work, the optimal parameter encoding for estimation theoretic security problem can be formulated in a game-theoretic framework. In Chapter 2 and 3, the eavesdropper is modeled to have no knowledge of encoding, and in Chapter 4, it has the full knowledge of the encoding strategies. In both cases, the transmitter is also aware of how much information the eavesdropper has about the encoding operation. Alternatively, the eavesdropper can have partial information about the encoding strategy, and the transmitter can adapt its encoding strategy based on the assumption about eavesdropper's knowledge. Such a framework will also cover the possible ambiguities over how much and how certain the transmitter and the eavesdropper know about each other's strategies. Another future work is to investigate robust encoding strategies under imperfect channel state information (CSI) at the transmitter. According to our model, the transmitter has perfect CSI for the channels of both the intended receiver and the eavesdropper. In practice, the transmitter may have a quantized information or a rough estimate about CSI information; hence, it is useful to develop encoding strategies under CSI uncertainty as well. Another important future work is the application of encoding ideas presented in this dissertation to wireless localization problems with anchor, target and eavesdropper nodes [75]-[77].

It is also interesting to find and develop necessary tools to provide a direct comparison between information and estimation theoretic approaches. In this dissertation, we have investigated the scenarios with finite numbers of observations. Even though most of the literature in information theoretic security focuses on asymptotic results, in some of the recent studies, new achievability results and converse bounds are provided for the maximal secret communication rate for wiretap channels in the finite regime [70]. Therefore, for a given blocklength n, the secrecy level, and the block error probability, the upper and lower bounds are available; hence, they may prove useful in practical comparisons with our setting. However, in order to build the bridge between estimation theoretic secrecy, which is based on the MSE constraint, and information theoretical secrecy, it is necessary to define a common framework in which these two considerably different approaches and their implications can be compared analytically. This seems as an open problem (to the best of our knowledge), which is definitely worthwhile to investigate in the future.

Bibliography

[1] W. Diffie and M. Hellman, "New directions in cryptography," *IEEE Trans. Inf. Theory*, vol. 22, pp. 644–654, Nov. 1976.

[2] R. L. Rivest, A. Shamir, and L. Adleman, "A method for obtaining digital signatures and public-key cryptosystems," *Commun. ACM*, vol. 21, pp. 120–126, Feb. 1978.

[3] C. E. Shannon, "Communication theory of secrecy systems," *Bell Syst. Tech. J.*, vol. 28, pp. 656–715, Oct. 1949.

[4] J. Yang, L. Wang, G. Geraci, M. Elkashlan, J. Yuan, and M. Renzo, "Safeguarding 5G wireless communication networks using physical layer security," *IEEE Commun. Mag.*, vol. 53, pp. 20–27, Apr. 2015.

[5] A. Mukherjee, "Physical-layer security in the Internet of Things: Sensing and communication confidentiality under resource constraints," *Proc. IEEE*, vol. 103, pp. 1747–1761, Oct. 2015.

[6] Y. Wang, G. Attebury, and B. Ramamurthy, "A survey of security issues in wireless sensor networks," *IEEE Commun. Surveys Tuts.*, vol. 8, pp. 2–23, Apr.–June 2006.

[7] A. Yener and S. Ulukus, "Wireless physical-layer security:Lessons learned from information theory," *Proc. IEEE*, vol. 103, pp. 1814–1825, Oct. 2015.

[8] A. D. Wyner, "The wire-tap channel," *Bell Syst. Tech. J.*, vol. 54, pp. 1355–1387, Oct. 1975.

[9] Y. Liang, H. V. Poor, and S. Shamai, "Secure communication over fading channels," *IEEE Trans. Inf. Theory*, vol. 54, pp. 2470–2492, June 2008.

[10] A. Khisti, G. Wornell, A. Wiesel, and Y. Eldar, "On the Gaussian MIMO wiretap channel," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, pp. 2471–2475, June 2007.

[11] P. K. Gopala, L. Lai, and H. E. Gamal, "On the secrecy capacity of fading channels," *IEEE Trans. Inf. Theory*, vol. 54, pp. 4687–4698, Oct. 2008.

[12] I. Csiszár and J. Körner, "Broadcast channels with confidential messages," *IEEE Trans. Inf. Theory*, vol. 24, pp. 339–348, May 1978.

[13] H. Weingarten, Y. Steinberg, and S. S. Shamai, "The capacity region of the Gaussian multiple-input multiple-output broadcast channel," *IEEE Trans. Inf. Theory*, vol. 52, pp. 3936–3964, Sept. 2006.

[14] Y. Liang, A. Somekh-Baruch, H. V. Poor, S. Shamai, and S. Verdu, "Capacity of cognitive interference channels with and without secrecy," *IEEE Trans. Inf. Theory*, vol. 55, pp. 604–619, Feb. 2009.

[15] P. Xu, Z. Ding, X. Dai, and K. K. Leung, "A general framework of wiretap channel with helping interference and state information," *IEEE Trans. Inf. Forensics Security*, vol. 9, pp. 182–195, Feb. 2014.

[16] O. O. Koyluoglu, H. E. Gamal, L. Lai, and H. V. Poor, "Interference alignment for secrecy," *IEEE Trans. Inf. Theory*, vol. 57, pp. 3323–3332, June 2011.

[17] J. Zhu, J. Mo, and M. Tao, "Cooperative secret communication with artificial noise in symmetric interference channel," *IEEE Commun. Lett.*, vol. 14, pp. 885–887, Oct. 2010.

[18] S. Leung-Yan-Cheong and M. E. Hellman, "The Gaussian wire-tap channel," *IEEE Trans. Inf. Theory*, vol. 24, pp. 451–456, July 1978.

[19] Y. W. P. Hong, P. Lan, and C. C. J. Kuo, "Enhancing physical-layer secrecy in multiantenna wireless systems: An overview of signal processing approaches," *IEEE Signal Process. Mag.*, vol. 30, pp. 29–40, Sept. 2013.

[20] R. Bassily, E. Ekrem, X. He, E. Tekin, J. Xie, M. R. Bloch, S. Ulukus, and A. Yener, "Cooperative security at the physical layer: A summary of recent advances," *IEEE Signal Process. Mag.*, vol. 30, pp. 16–28, Sept. 2013.

[21] L. Lai and H. E. Gamal, "The relay-eavesdropper channel: Cooperation for secrecy," *IEEE Trans. Inf. Theory*, vol. 54, pp. 4005–4019, Sept. 2008.

[22] Z. Chu, K. Cumanan, Z. Ding, M. Johnston, and S. Y. L. Goff, "Secrecy rate optimizations for a MIMO secrecy channel with a cooperative jammer," *IEEE Trans. Veh. Technol.*, vol. 64, pp. 1833–1847, May 2015.

[23] Z. Chu, K. Cumanan, Z. Ding, M. Johnston, and S. L. Goff, "Secrecy rate optimization for a MIMO secrecy channel based on Stackelberg game," in *22nd European Signal Process. Conf.(EUSIPCO)*, pp. 126–130, Sept. 2014.

[24] H. Shen, W. Xu, and C. Zhao, "QoS constrained optimization for multi-antenna af relaying with multiple eavesdroppers," *IEEE Signal Process. Lett.*, vol. 22, pp. 2224–2228, Dec. 2015.

[25] H. Shen, Y. Deng, W. Xu, and C. Zhao, "Secrecy-oriented transmitter optimization for visible light communication systems," *IEEE Photonics Journal*, vol. 8, pp. 1–14, Oct. 2016.

[26] W. Liao, T. Chang, W. Ma, and C. Chi, "Qos-based transmit beamforming in the presence of eavesdroppers: An optimized artificial-noise-aided approach," *IEEE Trans. Signal Process.*, vol. 59, pp. 1202–1216, Mar. 2011.

[27] J. Guo, U. Rogers, X. Li, and H. Chen, "Secrecy constrained distributed detection in sensor networks," *IEEE Trans. Signal Inf. Process. Netw.*, vol. 4, pp. 378–391, June 2018.

[28] S. Asoodeh, F. Alajaji, and T. Linder, "Privacy-aware MMSE estimation," in *IEEE Int. Symp. Inf. Theory (ISIT)*, pp. 1989–1993, July 2016.

[29] A. Ozcelikkale and T. M. Duman, "Cooperative precoding and artificial noise design for security over interference channels," *IEEE Signal Process. Lett.*, vol. 22, pp. 2234–2238, Dec. 2015.

[30] B. Kailkhura, V. S. S. Nadendla, and P. K. Varshney, "Distributed inference in the presence of eavesdroppers: A survey," *IEEE Commun. Mag.*, vol. 53, pp. 40–46, June 2015.

[31] X. Guo, A. S. Leong, and S. Dey, "Estimation in wireless sensor networks with security constraints," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 53, pp. 544–561, Apr. 2017.

[32] J. Guo, H. Chen, and U. Rogers, "Asymptotic perfect secrecy in distributed estimation for large sensor networks," in *Proc. IEEE Int. Conf. on Acoust., Speech, Signal Process. (ICASSP)*, pp. 3336–3340, Mar. 2017.

[33] X. Guo, A. S. Leong, and S. Dey, "Distortion outage minimization in distributed estimation with estimation secrecy outage constraints," *IEEE Trans. Signal Inf. Process. Netw.*, vol. 3, pp. 12–28, Mar. 2017.

[34] J. Zhang, R. S. Blum, and H. V. Poor, "Approaches to secure inference in the Internet of Things: Performance bounds, algorithms, and effective attacks on IoT sensor networks," *IEEE Signal Proc. Mag.*, vol. 35, pp. 50–63, Sept. 2018.

[35] F. Farokhi and H. Sandberg, "Fisher information as a measure of privacy: Preserving privacy of households with smart meters using batteries," *IEEE Trans. Smart Grid*, vol. 9, pp. 4726–4734, Sept. 2018.

[36] H. Reboredo, J. Xavier, and M. R. D. Rodrigues, "Filter design with secrecy constraints: The MIMO Gaussian wiretap channel," *IEEE Trans. Signal Process.*, vol. 61, pp. 3799–3814, Aug. 2013.

[37] M. Pei, J. Wei, K. Wong, and X. Wang, "Masked beamforming for multiuser MIMO wiretap channels with imperfect CSI," *IEEE Trans. Wireless Commun.*, vol. 11, pp. 544–549, Feb. 2012.

[38] Y. Chen, S. Kar, and J. M. F. Moura, "The internet of things: Secure distributed inference," *IEEE Signal Process. Mag.*, vol. 35, pp. 64–75, Sept. 2018.

[39] T. C. Aysal and K. E. Barner, "Sensor data cryptography in wireless sensor networks," *IEEE Trans. Inf. Forensics Security*, vol. 3, pp. 273–289, June 2008.

[40] A. N. Samudrala and R. S. Blum, "Asymptotic analysis of a new low complexity encryption approach for Internet of Things, smart cities and smart grid," in *Proc. IEEE Int. Conf. on Smart Grid and Smart Cities (ICSGSCS)*, pp. 200–204, 2017.

[41] A. N. Samudrala and R. S. Blum, "On the estimation and secrecy capabilities of stochastic encryption for parameter estimation in IoT," in *52nd Annual Conf. on Inf. Sci. and Syst. (CISS)*, pp. 1–6, March 2018.

[42] C. Goken and S. Gezici, "ECRB-based optimal parameter encoding under secrecy constraints," *IEEE Trans. Signal Process.*, vol. 66, pp. 3556–3570, July 2018.

[43] C. Goken and S. Gezici, "Optimal parameter encoding based on worst case Fisher information under a secrecy constraint," *IEEE Signal Process. Lett.*, vol. 24, pp. 1611–1615, Nov. 2017.

[44] C. Goken, S. Gezici, and O. Arikan, "Estimation theoretic optimal encoding design for secure transmission of multiple parameters," *IEEE Trans. Signal Process.*, vol. 67, pp. 4302–4316, Aug. 2019.

[45] C. Goken and S. Gezici, "Estimation theoretic secure communication via encoder randomization," *IEEE Trans. Signal Process.*, vol. 67, pp. 6105–6120, Dec. 2019.

[46] C. Goken and S. Gezici, "Optimal parameter design for estimation theoretic secure broadcast," *IEEE Signal Process. Lett.*, 2019 (under review).

[47] H. V. Poor, *An Introduction to Signal Detection and Estimation*. New York: Springer-Verlag, 1994.

[48] H. L. V. Trees and K. L. B. (editors), *Bayesian Bounds for Parameter Estimation and Nonlinear Filtering/Tracking*. New Jersey: WileyInterscience, 2007.

[49] E. Gonendik and S. Gezici, "Fundamental limits on RSS based range estimation in visible light positioning systems," *IEEE Commun. Lett.*, vol. 19, pp. 2138–2141, Dec. 2015.

[50] A. A. Nasir, H. Mehrpouyan, S. Durrani, S. D. Blostein, R. A. Kennedy, and B. Ottersten, "Optimal training sequences for joint timing synchronization and channel estimation in distributed communication networks," *IEEE Trans. Commun.*, vol. 61, pp. 3002–3015, July 2013.

[51] A. Shahmansoori, R. Montalban, and G. Seco-Granados, "Effect of channel variability on pilot design for joint communications and positioning in OFDM," in *11th Int. Symp. Wireless Commun. Syst. (ISWCS)*, pp. 292–296, Aug. 2014.

[52] H. Gazzah and J. Delmas, "Direction finding antenna arrays for the randomly located source," *IEEE Trans. Signal Process.*, vol. 60, pp. 6063–6068, Nov. 2012.

[53] R. Montalban, J. A. Lopez-Salcedo, G. Seco-Granados, and A. L. Swindlehurst, "Power allocation method based on the channel statistics for combined positioning and communications OFDM systems," in *Proc. IEEE Int. Conf. on Acoust., Speech, Signal Process. (ICASSP)*, pp. 4384–4388, May 2013.

[54] M. F. Keskin, E. Gonendik, and S. Gezici, "Improved lower bounds for ranging in synchronous visible light positioning systems," *J. Lightw. Technol*, vol. 34, pp. 5496–5504, Dec. 2016.

[55] R. Liu, T. Liu, H. V. Poor, and S. Shamai, "Multiple-input multiple-output Gaussian broadcast channels with confidential messages," *IEEE Trans. Inf. Theory*, vol. 56, pp. 4215–4227, Sept. 2010.

[56] S. A. A. Fakoorian and A. L. Swindlehurst, "On the optimality of linear precoding for secrecy in the MIMO broadcast channel," *IEEE J. Sel. Areas Commun.*, vol. 31, pp. 1701–1713, Sept. 2013.

[57] G. Geraci, S. Singh, J. G. Andrews, J. Yuan, and I. B. Collings, "Secrecy rates in broadcast channels with confidential messages and external eavesdroppers," *IEEE Trans. Wireless Commun.*, vol. 13, pp. 2931–2943, May 2014.

[58] H. Soganci, S. Gezici, and O. Arikan, "Optimal stochastic parameter design for estimation problems," *IEEE Trans. Signal Process.*, vol. 60, pp. 4950–4956, Sept. 2012.

[59] H. Soganci, S. Gezici, and O. Arikan, "Optimal signal design for multiparameter estimation problems," *IEEE Trans. Signal Process.*, vol. 63, pp. 6074–6085, Nov. 2015.

[60] A. G. i Fabregas and G. Caire, "Coded modulation in the block-fading channel: Coding theorems and code construction," *IEEE Trans. Inf. Theory*, vol. 52, pp. 91–114, Jan. 2006.

[61] R. Knopp and P. A. Humblet, "On coding for block fading channels," *IEEE Trans. Inf. Theory*, vol. 46, pp. 189–205, Jan. 2000.

[62] I. M. Gelfand and S. V. Fomin, *Calculus of Variations*. Prentice-Hall, 1963.

[63] D. G. Luenberger, *Optimization by Vector Space Methods*. John Wiley & Sons, Inc., 1997.

[64] E. W. Karmen and J. K. Su, *Introduction to Optimal Estimation*. London, U.K.: Springer-Verlag, 1999.

[65] K. Atkinson, *An Introduction to Numerical Analysis*. John Wiley, 2 ed., 1989.

[66] T. J. Rivlin, *Approximation of Functions*. New York: Chelsea, 1986.

[67] H. Jeffreys and B. S. Jeffreys, "Weierstrass's theorem on approximation by polynomials," *Methods of Mathematical Physics*, pp. 446–448, 1988.

[68] S. M. Kay, *Fundamentals of Statistical Signal Processing:, Vol I: Estimation Theory*. Upper Saddle River, New Jersey: Prentice Hall, 1993.

[69] A. J. Goldsmith, *Wireless Communications*. New York: Cambridge University Press, 2005.

[70] W. Yang, R. F. Schaefer, and H. V. Poor, "Wiretap channels: Nonasymptotic fundamental limits," *IEEE Trans. Inf. Theory*, vol. 65, July 2019.

[71] M. Cohen, "The Fisher information and convexity (corresp.)," *IEEE Trans. Inf. Theory*, vol. 14, pp. 591–592, July 1968.

[72] A. Patel and B. Kosko, "Optimal noise benefits in Neyman-Pearson and inequality-constrained signal detection," *IEEE Trans. Signal Process.*, vol. 57, pp. 1655–1669, May 2009.

[73] S. Bayram, S. Gezici, and H. V. Poor, "Noise enhanced hypothesis-testing in the restricted Bayesian framework"," *IEEE Trans. Commun.*, vol. 58, pp. 3972–3989, Aug. 2010.

[74] R. T. Rockafellar, *Convex Analysis*. Princeton, NJ: Princeton University Press, 1968.

[75] W. Dai and M. Z. Win, "A theoretical foundation for location secrecy," in *Proc. IEEE Int. Conf. on Commun. (ICC)*, pp. 1–6, May 2017.

[76] W. Dai and M. Z. Win, "On protecting location secrecy," in *Proc. Int. Symp. on Wireless Commun. Syst. (ISWCS)*, pp. 31–36, Aug 2017.

[77] T. Zhang, X. Li, and Q. Zhang, "Location privacy protection: A power allocation approach," *IEEE Trans. Commun.*, vol. 67, pp. 748–761, Jan. 2019.