June 2021

COVERT COMMUNICATIONS IN CONTINUOUS-TIME SYSTEMS

Ke Li
University of Massachusetts Amherst

COVERT COMMUNICATIONS IN CONTINUOUS-TIME SYSTEMS

A Dissertation Presented

by

KE LI

Submitted to the Graduate School of the
University of Massachusetts Amherst in partial fulfillment
of the requirements for the degree of

DOCTOR OF PHILOSOPHY

May 2021

Electrical & Computer Engineering

COVERT COMMUNICATIONS IN CONTINUOUS-TIME SYSTEMS

A Dissertation Presented

by

KE LI

Approved as to style and content by:

Dennis L. Goeckel, Chair

Donald F. Towsley, Member

Hossein Pishro-Nik, Member

Marco F. Duarte, Member

Christopher V. Hollot, Department Chair
Electrical & Computer Engineering

DEDICATION

To my beloved grandparents.

ACKNOWLEDGMENTS

First and foremost, I would like to express my deepest gratitude to my advisor Prof. Dennis Goeckel for his patient guidance and support through all the years of my Ph.D. study. I sincerely appreciate his brilliant teaching, constant encouragement and thoughtful suggestions to not only help me in conducting my research, writing scientific papers and giving presentation, but also grant me knowledge that I can benefit from for a lifetime. I am very grateful for him spending so much time and efforts discussing research work with me every week and providing valuable suggestions that make this work possible. I also want to thank him for his kindness and letting me have a pleasant time during my Ph.D. study. I honestly couldn't wish for a better advisor.

Furthermore, I would like to thank Prof. Don Towsley for his guidance in my research. His brilliant ideas and critical thoughts have helped me in solving difficult problems and pushed me to keep improving my work. I would also like to thank Prof. Patrick Kelly for him guiding me to solve very technical problems and I appreciate that he is always willing to help. I also sincerely appreciate Prof. Hossein Pishro-Nik and Prof. Amir Houmansadr for their guidance and suggestions in my research, and Prof. Marco Duarte for his valuable comments and feedback on this dissertation.

Further, I would like to thank all of my current and former colleagues (in alphabetical order) who have unselfishly shared me their knowledge and experience: Alireza Bahramali, Bo Guan, Ramin Soltani, Tamara Sobers, Virat Shejwalkar, and Nazanin Takbiri. Some of them have made valuable contributions towards the work in this dissertation.

Finally, I would like to thank my beloved family. I am extremely grateful for my parents who have supported me all these years. Although that I am their only child and need to live in another country that is far away from them, they encourage me to chase my dreams and trust me in every decision that I made. I would also like to deeply thank my husband Yuxiao who is always there for me and believe in me no matter what. Thank to him that I could get through difficult times and overcome great challenges.

ABSTRACT

COVERT COMMUNICATIONS IN CONTINUOUS-TIME SYSTEMS

MAY 2021

KE LI

B.S., CHINA UNIVERSITY OF GEOSCIENCES, BEIJING

M.S., STATE UNIVERSITY OF NEW YORK AT BUFFALO

Ph.D., UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by: Professor Dennis L. Goeckel

This dissertation studies covert wireless communications where a transmitter (Alice) intends to transmit messages to a legitimate receiver (Bob) such that the presence of the message is hidden from an attentive warden (Willie). Here we consider pertinent aspects of covert communications that focus on moving such systems closer to implementation. For example, previous studies use the standard discrete-time communication model when analyzing covert communications, since this is commonly assumed without loss of generality in standard communication theory. However, it is not clear that such a model captures the salient aspects of the continuous-time covert communications problem. A power detector that is optimal for the warden in a discrete-time covert communications scenario may not be optimal on a continuous-time model. Thus, it is of interest to consider this more realistic model for physical channels. After analyzing a power optimization problem using the standard discrete-time model, we move to the key part of system implementation: the instantiation in

true continuous-time systems of the discrete-time models studied to this point in the literature. A key goal is to examine Willie's detection capability on a continuous-time model and study how the limits of covert communications change from the discrete-time case. In particular, we show that detectors for Willie can benefit from the continuous-time setting and outperform detectors based on the discrete-time model; not surprisingly, this has a significant impact on the true covert throughput of the system. Nevertheless, we establish constructions such that efficient covert communications can still be achieved in a continuous-time model, and prove the fundamental limit on the covert communication rate. After considering the continuous-time problem in detail, we then turn to addressing another limitation of previous work - the requirement for an intentional jammer to facilitate efficient covert communication. Instead, we consider how to exploit a pre-existing interference source - a radar - to achieve covert communication. We establish a covert communication scheme in such an environment, and analyze the corresponding covert rate. Finally, we consider the use of a detection technique similar to that in the covert communications problem, in the area of quantized signal detection.

TABLE OF CONTENTS

LIST OF FIGURES

Figure **Page**

CHAPTER 1

INTRODUCTION

1.1 Motivation

Secrecy in modern communications plays a crucial role in many applications. In a typical wireless communication system, secrecy is often obtained through encryption. In particular, a message is encoded before being transmitted to the legitimate receiver in order to protect it from potential eavesdroppers. A generated key is given to the legitimate receiver to decrypt the received message. In contrast, the eavesdropper has difficulty decrypting the message without possessing the key. Although modern communication technologies enable advanced algorithms for information encryption that require considerable computational resources and skills for the eavesdropper to decrypt the message, this standard cryptographic security is often not sufficient. This is due to two main reasons: 1) adversaries can enhance their decipher capabilities over time as technologies improve, putting current and especially past encrypted transmissions in danger; and 2) adversaries can apply side-channel attacks [1] to infer message content without acquiring the exact decrypted message. Thus, it is often desirable to achieve information-theoretic security that does not depend on unproven assumptions about computational hardness.

In information-theoretic secrecy, physical layer encryption exploits the wireless channel to achieve provable and unbreakable security. In the 1970s, Wyner posed the Alice-Bob-Eve problem where Alice wants to send secret messages to Bob without being decoded by Eve [2]. Wyner defined the secrecy capacity that measures the level of security of a communication system - the rate at which Alice can secretly

transmit information to Bob. He showed that secure communication is possible if the channel from Alice to Bob is statistically better than that from Alice to Eve. Later Csiszár and Körner [3] showed that secure communication is possible even if the channel from Alice to Eve is statistically better. The key of the information theoretic approach to communication security is to exploit the difference between the channel to the legitimate receiver and the channel to the adversary or eavesdropper, using the randomness of the physical medium (e.g., channel noise and fading), to benefit the legitimate receiver.

In addition to the vulnerability to developments in decipher power, in many real-life scenarios, the very existence of the transmission arouses suspicion. For example in military communications, the detection of a transmission may reveal an activity in that region, or in Internet of things (IoT) applications, the detection of a transmission can compromise a user's location privacy. A practical way of hiding the existence of a message is to conceal it within another message. This is termed steganography. The first recorded uses of steganography were mentioned by the ancient Greek historian Herodotus [4]. One example is that the spartan king Demaratus sent a warning about a forthcoming attack to Greece by writing it directly on the wooden backing of a wax tablet (commonly used then as reusable writing surfaces) before applying its beeswax surface. Nowadays, steganography is widely used in the digital world for hiding communications in document files, image files, programs or protocols. However, in the analog world, applying steganography to communication systems can be difficult due to the existence of noise. Another issue is that steganography cannot be applied when there is no cover text to hide the covert message, which is often true for many applications where the existence of the transmission needs to be hidden, such as the examples mentioned in the previous paragraph.

Covert communications can be employed to address the problem. The intuition of covert communications is simple: the transmitter must send a signal power strong

enough for the receiver to receive and decode the message, but weak enough so that the adversary cannot detect its presence. The research challenges are to determine under what conditions covert communications can be achieved, and to find the theoretical limits on the covert transmission rate. This can help the transmitter and the receiver evaluate their abilities in covert transmissions (whether it is capable to transmit covert and reliable messages and what transmission rate can be employed). Analogously, this also helps adversaries to evaluate their detection capabilities. It will eventually help us to design technologies that enable or prevent covert communications in real-life scenarios. Many studies have been done on covert communications to study the fundamental limits (as a function of the blocklength of the transmitted message) in a wide range of scenarios, which we will introduce in detail in the next section. Some works take other perspectives, such as considering an infinite blocklength to better understand the underlying mechanisms of the covert communication problem. They provide inspiration for our work on optimal power adaptation in covert communications in the second chapter where we question how to efficiently allocate power to the transmitter such that the receiver can receive the message reliably under a covertness constraint.

However, almost all previous work, including our power adaptation work, focused on a discrete-time model, which assumes an equivalence to the actual continuous-time model of interest since a discrete-time equivalent model is often sufficient for digital communications. Given a pulse-shaped baseband signal, by the Nyquist folding criterion, the receiver can extract all of the information of the signal by sampling at a certain rate, resulting in no intersymbol interference (ISI). But in covert communications, sampling at a higher rate has utility for signal detection. Since commonly generated signals are periodic and do not resemble Gaussian noise, it allows the eavesdropper to extract features in the transmitted signal that are different from the noise to detect the presence of a transmission. This may provide more information to the

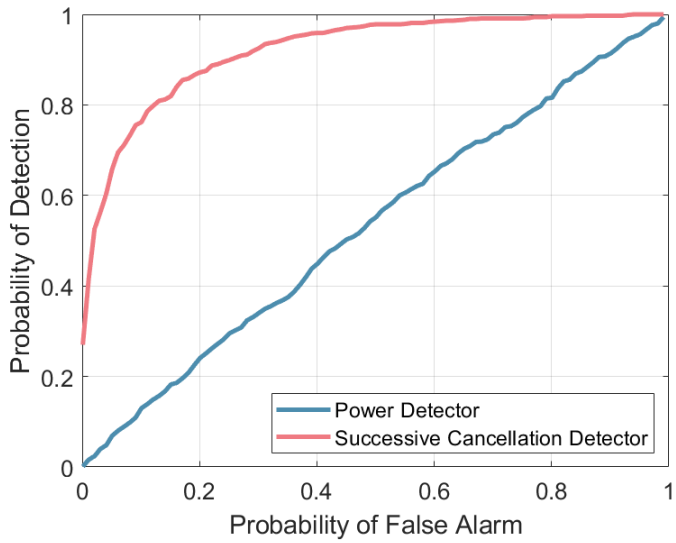


Figure 1.1: ROC curves of the standard power detector and the successive cancellation detector employed by Willie in the Alice-Bob-Willie-jammer scenario with AWGN channels.

eavesdropper in comparison to the discrete-time model. In a discrete-time model, Bash *et al.* in [5] and [6] proved that Alice can transmit at most $\mathcal{O}(\sqrt{n})$ covert bits to Bob in n channel uses of an AWGN channel. Although Lee and Baxley in [7] showed that $\mathcal{O}(n)$ covert bits can be achieved if Willie is unaware of his background noise, Goeckel *et al.* in [37] stated that Willie can estimate his noise through a large collection of observations and hence Alice can only achieve $\mathcal{O}(\sqrt{n})$ covert bits. A possible scenario to achieve $\mathcal{O}(n)$ covert bits in a discrete-time channel is provided in [9] by Sobers *et al.* where an uninformed jammer is added to the system. However, in [10] and [11], the $\mathcal{O}(n)$ result is challenged in a continuous setting. By having Willie employing a more advanced detector, the jammer's power can be cancelled and we revert to the $\mathcal{O}(\sqrt{n})$ result. Considering a continuous-time model may changes the fundamental limits of covert communications, but since some assumptions in [10] and [11] may not hold in practice, a deeper understanding of the problem is needed to help realize covert communications in real-life applications. Figure 1.1 provides an example that compares the ROC curve of another advanced detector for Willie

(developed in Chapter 3) with that of the standard power detector (which is proved optimal for a discrete-time model) in a continuous-time covert communication system. This shows that the covert rates that can be achieved in actual continuous-time channels are lower than those predicted in [9].

The above studies consider covert communications in an Alice-Bob-Willie-jammer scenario where a friendly jammer is used as an interference source to assist Alice to achieve covert communications, and in many cases, we need to have some control over the jammer in order to help Alice. However, this may not be operable in many practical scenarios. For example in military communications, introducing a friendly jammer will arouse suspicion. Therefore, it is desirable to develop covert communication systems that are able to exploit interference sources already existing in the environment. In the fourth chapter, we will study covert communications in an environment with an illuminating radar that will play the role of a jammer.

The job of a detector at the warden in a covert communication system is to decide between hypotheses that either Alice is transmitting or not transmitting. Besides covert communications, this kind of hypothesis testing is involved in many other areas of wireless communications, e.g., in detecting quantized signals. Detecting quantized signals is important since in many applications, such as network security or radar systems, we want to know whether a received signal was sent directly by a friend, or was recorded (hence quantized) by an adversary and then replayed. This problem needs advanced study to learn the theoretical limits in signal detection, which will help us evaluate the security level in such applications.

1.2 Background

1.2.1 Covert Communications

In the basic model of covert communication shown in Figure 1.2, a transmitter (Alice) attempts to reliably transmit messages to a receiver (Bob) without a warden

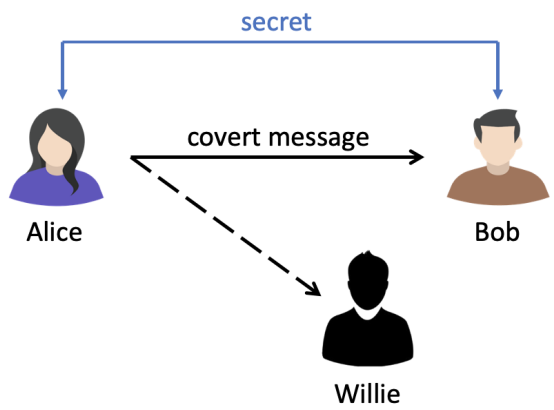


Figure 1.2: Alice-Bob-Willie model: Alice attempts to transmit reliably and covertly to Bob in the presence of a warden Willie.

(Willie) detecting her communication. The original work by Bash *et al.* in [5] and [6] demonstrates a square-root law (SRL) of the fundamental limits of covert communications over AWGN channels. They show that $\mathcal{O}(\sqrt{n})$ bits in n channel uses of a discrete-time channel can be transmitted covertly and reliably from Alice to Bob. In terms of transmission power, $\mathcal{O}(\frac{1}{\sqrt{n}})$ power can be employed by Alice for covert and reliable communications. Beyond that power, Alice's transmission will be detected by Willie with high probability. Below that power, her message cannot be successfully decoded by Bob with small probability of error. Follow-on work has studied covert communications and proved the SRL over various channel models. Che *et al.* in [13] considered covert communications on Binary Symmetric Channels (BSCs). More generic Discrete Memoryless Channels (DMCs) are considered by Wang *et al.* in [15] and [16]. This work was also extended to classical-quantum channels in [17].

Many works in covert communications are based on models assuming that Alice and Bob share a secret key that is unknown to Willie. Bloch studied the length of the secret key needed to achieve the SRL in covert communications over DMCs in [18] and [19]. He proved that the SRL can be achieved if Alice and Bob share on $\mathcal{O}(\sqrt{n})$-bit key irrespective of the quality of the channels. If the Alice-to-Bob channel

is better than the Alice-to-Willie channel, then the SRL can be achieved without a secret key. Tahmasbi and Bloch also studied first- and second-order asymptotics for covert communications with three different covertness metrics in [20] and [21]. Exact first-order asymptotics are established for all metrics.

The aforementioned works all consider covert communications when Willie has full knowledge of his channel statistics. We can also consider the case when there are uncertainties at Willie, e.g., he is not sure of the background noise or additional interference is involved. The results of many works show that a positive rate, i.e., $\mathcal{O}(n)$ bits in n channel uses, can be achieved. Lee *et al.* in [7, 22, 23] and Che *et al.* in [14] found that if Willie is uncertain of his noise power, then Alice can achieve a positive covert rate. However, the assumption that Willie's receiver does not know the noise power may not be true in real-life scenarios since Willie can estimate his noise through a collection of channel observations when Alice does not transmit [37]. To achieve positive covert rate when Willie is certain about his noise power, Sobers *et al.* introduced a scenario where an uninformed (does not have any coordination with Alice) jammer is added in the communication system in [9, 11]. This jammer randomly generates interference and creates uncertainties at Willie which allows a positive covert rate for Alice. They also investigated covert communications over fading channels in addition to AWGN channels in such a scenario. Our work in Chapters 2 and 3 will be based on this scenario where an uninformed jammer is added. One difference between this scenario and the scenario when Willie does not know his received power is that the interference generated by the jammer not only affects Willie, but also impacts Bob. Therefore, although a positive rate might be achieved such as in [9], this scenario with a jammer does not necessarily guarantee a higher covert rate than the basic Alice-Bob-Willie scenario with the same environment settings. For example, when Bob is located relatively close to the jammer, the covert rate will be degraded significantly.

1.2.2 Power Adaptation in Covert Communications

Many works in covert communications focus on obtaining the scaling limits as a function of the block length n. The authors in [24] take a different approach and consider an infinite blocklength; they then ask what is the probability that channel conditions are such that a communication is reliable and covert. We term this an "outage" approach [25]- [27] which often captures the salient aspects of the problem and can clearly illustrate the underlying mechanisms.

In [24], the authors use the outage approach to consider possible power adaptation schemes at the transmitter to achieve covert communication. In particular, they studied standard and truncated channel inversion (TCI) schemes, where Alice varies her transmit power based on the channel from herself to Bob in order to keep the received signal power at Bob a constant. The authors examine the performance in terms of the achieved effective covert throughput and show that TCI outperforms the standard channel inversion. However, [24] does not establish the optimality of the scheme, and the parameters of the scheme need to be obtained numerically. Our work in the second chapter will establish the exact optimal power adaptation schemes for different scenarios.

1.2.3 Covert Communications on a Continuous-Time Model

Although covert communications have been intensively studied for the past few years, some important issues are ignored. Almost all current research is based on a discrete-time model where transmitters sends discrete symbols. In standard communication theory, it is commonly assumed that the discrete-time model is approximately equivalent to a continuous-time model since, by sampling at the correct time instances, the receiver is able to get exactly the original sent symbols with no intersymbol interference (ISI). However, the assumption on perfect symbol synchronization may not always hold in practice. In addition, the generated continuous signals contain peri-

odic features that can be made use of by the receiver in differentiating the signal from Gaussian noise and/or interference. Based on this fact, a power detector for Willie that is optimal in the discrete-time model may not still be optimal in the continuous-time case. One example of a possible better detector is the cyclostationary detector (CSD). CSDs are signal-presence detectors that exploit the cyclostationarity of digital communication signals. It is particularly useful in detecting low signal-to-noise ratio (SNR) signals. Gardner in [28]- [31] extensively researched CSDs and the optimal CSDs for different modulation schemes. Kim *et al.* extended the study on CSDs and introduced the cyclostationary approaches to both signal detection and classification in cognitive radio in [32].

Although CSDs are efficient in extracting periodic features from a signal and hence can perform detection in basic Alice-Bob-Willie scenarios, they may not work well in the presence of a jammer in the system since the jammer also generates periodic signals. However, there are still detectors we can explore that work better than a power detector. Sobers *et al.* in [11] introduced a linear detector that exploits different timing offsets between Alice's and the jammer's pulse shaped signals. However, this detector only works for limited scenarios and requires that Willie have an accurate estimate of the timing offsets of both Alice's and the jammer's signals. In the third chapter, we will provide a different detector for Willie that works in more general scenarios and does not require knowledge of Alice's timing offset.

1.2.4 Covert communications under the cover of a radar

In covert communications, Willie attempts to determine whether he is only observing the background environment or a signal from Alice in that environment. Hence, uncertainty about the environment helps Alice hide her transmission. Sobers *et al.*, [9], introduced a model with an interference source to achieve positive covert rate: introducing an uninformed jammer to the system that randomly generates in-

terference, hence providing the required uncertainty at Willie. It shows that Alice can covertly transmit at most $\mathcal{O}(n)$ bits in n channel uses over both discrete-time AWGN and block fading channels.

In many scenarios, the jammer is assumed to be an intentional jammer that cooperates with Alice to help her achieve covert communications on purpose [33]. However, this require an active non-covert Alice-Bob teammate, which may be difficult to provide in some situations. For example, in military communications where a transmitter attempts to covertly communicate with a receiver in enemy territory. In addition, the jammer in [9] is itself not covert since the warden knows that the jammer is present and potentially trying to hide something. Hence, it is often useful to exploit an interference source that is already existing in the environment, such as radar emissions [34,35] and other existing communication sources [36,37], so that Alice and Bob can move into the area and hide under such interference.

The works [34, 35] by Blunt *et al.* introduces an intra-pulse radar-embedded communication system where the transmitter attempts to covertly send transmission symbols to the radar. In order to achieve covertness, the transmission symbols are embedded with the incident radar pulses, and are hidden behind the backscatter induced by the radar reflections. The design of intra-pulse covert symbols based on the incident radar waveform is studied in [34] such that the covert symbols are sufficiently different from the ambient radar scattering to ensure acceptable bit error rate (BER) but at the same time sufficiently similar to the scattering to avoid detection by any adversary. Our work in the fourth chapter will exploit the idea of embedding covert symbols with radar signal in a standard covert communication system, and analyze the fundamental covert rate in such system.

1.2.5 Detecting Quantized Signals

In many scenarios it is important to know whether a received signal was sent directly by a friend, or was recorded by an adversary and then replayed. Such an attack could be the replay attack (or playback attack) in network security, where an adversary records the transmitted message and replays it later to trick the receiver into unauthorized operations. This kind of attack also occurs in radar jamming and deception to protect targets from being detected by enemy radar systems. Extensive research [78]- [81] has proposed methods to efficiently detect false signals that are recorded and replayed in such fields. However, the fundamental limits of such attacks with hardware imperfections has not been explored. The fifth chapter will initiate a study on the theoretical limits in the detection of quantized signals. Learning the fundamental thresholds for the characteristics of the hardware (quantizer) will provide us with both theoretical insight and application utility.

1.3 Contributions

- **Optimal Power adaptation in Covert Communications (Chapter 2):**

  We consider a covert communication system consisting of a transmitter Alice, a legitimate receiver Bob, a warden Willie and an uninformed jammer that is not coordinated with Alice. With the information about the gain on the channel between Alice and Bob, Alice can adapt her transmit power to this gain to achieve a certain rate and meet the requirement of covertness such that Willie detects the presence of the transmission with low probability. We seek to find the optimal power adaptation that minimizes the average outage probability subject to the covertness constraint. We consider the following scenarios: 1) the jammer-to-Willie channel is AWGN; and 2) the jammer-to-Willie channel experiences Rayleigh fading. In both scenarios, we established the optimal power adaptation schemes separately under two covertness constraints: 1) an aver-

age covertness constraint; and 2) an instantaneous covertness constraint. We proved that in both scenarios, the optimal scheme under the average covertness constraint is truncated channel inversion. Under the instantaneous covertness constraint, TCI is not optimal, and we have provided the exact optimal power adaptation schemes. These schemes outperform standard approaches such as TCI by reducing the average outage probability significantly.

- **Covert Communications on a Continuous-Time Model (Chapter 3):**

  Dropping the assumption in most prior work that a discrete-time model in covert communications is equivalent to a continuous-time model, we directly study the Alice-Bob-Willie-jammer scenario on a continuous-time model in this work. We first provide an interference cancellation detector for Willie, which is inspired by co-channel interference cancellation techniques in cellular networks, that can mitigate the jammer's signal and detect Alice's presence in a continuous-time system. We show that this detector outperforms the standard power detector implemented in a continuous-time system in various circumstances, and that it does not require Willie to know Alice's timing offset. Then, we prove that covert communications can still be achieved with a continuous-time model regardless of the choice of Willie's receiver. In particular, given a time T, for a continuous-time channel with asymptotic bandwidth W as $T \to \infty$, we establish constructions such that $\mathcal{O}(WT)$ information bits can be transmitted covertly and reliably from Alice to Bob in T seconds.

- **Covert communications under the cover of a radar (Chapter 4):**

  Instead of using a friendly jammer as an interference source to assist Alice to achieve covert communications, in this work, we move Alice and Bob to an environment with a pre-existing illuminating radar. We exploit the idea of embedding covert symbols with radar signals and hide the transmission behind

the radar clutter. We provide a design of covert communication waveforms exploiting the radar signal. We also show that covert communications can be achieved with such a communication scheme, and establish the theoretical limit on the covert rate of transmission between Alice and Bob. In particular, we show that $\mathcal{O}(n)$ bits can be transmitted covertly and reliably to Bob in n samples of the radar signal.

- **Detecting Whether a Signal Has Been Quantized (Chapter 5):**

  In many areas like network security or radar jamming and deception, it is important to know whether a received signal was sent directly by a friend, or was recorded by an adversary and then replayed. In this work, we use a characteristic of the recording, namely the quantization, to study if the replayed (quantized) signal can be detected. In particular, we consider the requirements on the quantizer to keep the quantization from being detected in additive Gaussian noise. If a signal with length m is sent and a uniform quantizer with b quantization bits is employed for recording, we prove that $2^b = \omega(\sqrt{m})$ and a quantizer span of $\omega(\sqrt{\ln m})$ is sufficient for the adversary to avoid detection; that is, the probability of error P_e of the observer is bounded as $P_e \geq \frac{1}{2} - \epsilon$ for any $\epsilon > 0$. Conversely, having $2^b = \mathcal{O}(\sqrt{m})$ or a quantizer span of $o(\sqrt{\ln m})$ results in detection by the observer with high probability as $m \to \infty$.

CHAPTER 2

OPTIMAL POWER ADAPTATION IN COVERT COMMUNICATION WITH AN UNINFORMED JAMMER

2.1 Introduction

While most secure communication focuses on preventing an adversary from determining the content of a message, covert communication hides the existence of the communication between Alice and Bob, which is important for some applications. For example, in military communications the detection of a transmission may reveal activity in the region, or in IoT applications, the detection of a transmission can compromise a user's location privacy.

Recent work studied the limits of reliable covert communication. In [6] and [5], the fundamental limits of covert communications over additive white Gaussian noise (AWGN) channels were first studied. They show that $\mathcal{O}(\sqrt{n})$ bits in n channel uses of a discrete-time channel can be transmitted covertly and reliably. This is related to steganography, which is the practice of concealing a message within another message. Finite-alphabet steganographic systems have a similar square root law: at most $\mathcal{O}(\sqrt{n})$ symbols in a length n covertext may safely be modified to hide a length $\mathcal{O}(\sqrt{n}\log n)$-bit message [12]. The extra $\log n$ factor is due to the lack of noise in the steganographic context. Successive work has extended the results of [6] and [5] to binary symmetric channels (BSCs) [13, 14] and discrete memoryless channels (DMCs) [15, 16]. The work in [16]- [19] also established the constants behind the Big-$\mathcal{O}$ notation for both DMC and AWGN channels.

Since Willie detects the presence of Alice by detecting deviations from the background noise environment, uncertainty about that environment will make it harder

for Willie to determine if Alice is transmitting. Thus, when Willie is uncertain about the statistical characterization of the Alice-to-Willie channel, Alice can transmit $\mathcal{O}(n)$ bits in n channel uses reliably and covertly [7]- [23]. To remove the restrictions on Willie, [9] and [11] introduce an uninformed jammer to assist the communication by actively generating jamming signals. This might be an electronic jammer placed by Alice and Bob to enhance security, or a jammer placed by an adversary that tries to jam potential communication from Alice.

The aforementioned works focus on obtaining the scaling limits of reliable covert communication as a function of the blocklength n. The authors in [24] take a different approach and consider an infinite blocklength; they then ask what is the probability that channel conditions are such that a communication is reliable and/or covert. We term this an "outage" approach [25]- [27]. This approach often captures the salient aspects of the problem and can clearly illustrate the underlying mechanisms. We adopt such an approach, which allows Willie to have a perfect estimate of the power at his receiver. Next, we consider how such an outage formulation impacts the design of reliable communication between Alice and Bob. Since the variation on the Alice-to-Bob channel can significantly affect the outage probability, if this current channel gain is known to Alice then Alice can adapt her power to minimize outage under a covertness constraint. This motivates the derivation of an optimal power adaptation scheme.

Adapting power to enhance performance in wireless communications has long been studied. The authors in [38] prove that the optimal power adaptation to maximize channel capacity under an average power constraint is "water-pouring". To simplify designs, they also consider two suboptimal adaptation schemes: channel inversion and truncated channel inversion (TCI), which adapt the transmit power but keep the transmission rate constant. The standard channel inversion scheme is simple to implement but can exhibit a large capacity penalty in extreme fading environments.

TCI, on the other hand, only compensates for fading above a certain cutoff point, and hence avoids the capacity penalty when the channel is bad. This technique has been employed in many non-covert systems to minimize outage or to achieve good throughput under an outage constraint, such as in cellular uplink transmission [39], in cognitive radio broadcasting [40] and in hybrid free-space optical/radio-frequency transmission [41]. It can also be employed in covert communication systems to assist Alice in achieving reliable and covert transmission. In [24], the authors consider channel inversion power adaptation at the transmitter to achieve reliable covert communication, where Alice varies her transmit power based on the channel from herself to Bob in order to keep the received signal power at Bob a constant. The authors examine the performance in terms of the achieved effective covert throughput and show that TCI outperforms the standard channel inversion. However, [24] does not establish the optimality of the scheme, and the parameters of the scheme need to be obtained numerically. Here, we show that TCI is optimal in certain scenarios and provide exact optimal schemes that may or may not be TCI in different scenarios.

In this chapter, we consider a covert communication system consisting of a transmitter Alice, a legitimate receiver Bob, a warden Willie and an uninformed jammer that is not coordinated with Alice. The jammer actively sends jamming signals that interfere with reception at both Willie and Bob. With the information about the gain on the channel between Alice and Bob, Alice can adapt her transmit power to this gain to achieve a certain rate and meet the requirement on covertness such that Willie detects the presence of the transmission with low probability. We seek to find the optimal power adaptation that minimizes the average outage probability subject to the covertness constraint. We consider the following scenarios: 1) the jammer-to-Willie channel is AWGN; and 2) the jammer-to-Willie channel experiences Rayleigh fading.

We present the system model and metrics in Section 2.2. Section 2.3 considers the scenario when the jammer-to-Willie channel is AWGN, and Section 2.4 proves similar results in the scenario when there is a fading channel between the jammer and Willie. Both sections provide the optimal power adaptation in two cases: 1) long-term adaptation with an average covertness constraint; and 2) short-term adaptation with an instantaneous covertness constraint. Numerical results that examine and compare the performance of different power adaptation schemes are also presented. Finally, Section 2.5 draws the conclusions.

2.2 System Model and Metrics

2.2.1 System Model

Consider the scenario shown in Fig. 2.1 where Alice wants to transmit a message to Bob reliably and covertly without detection by a warden Willie, and a jammer assists the communication by actively sending jamming signals but without any coordination with Alice [9, 11]. We are interested in Alice's ability to transmit covertly in a time

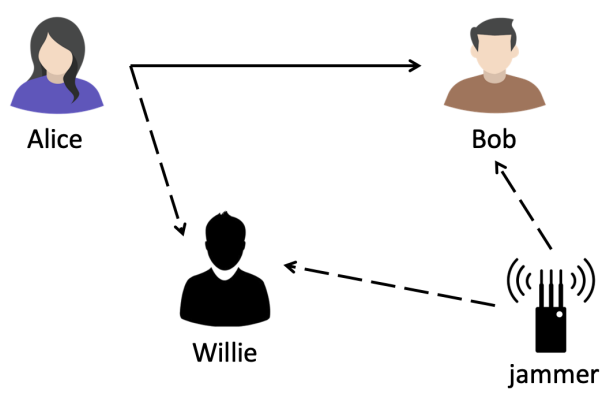


Figure 2.1: System model: With help from a jammer, Alice attempts to transmit reliably and covertly to Bob in the presence of a warden Willie.

slot equal to the codeword length n and Willie's ability to detect such a transmission in that slot. If Alice decides to transmit, she maps her message to the complex symbol sequence $\mathbf{f} = [f_1, f_2, \ldots, f_n]$. We assume that Alice transmits each symbol

of a codeword with an average power P_a. The jammer transmits complex symbol sequence $\mathbf{g} = [g_1, g_2, \ldots, g_n]$ with power P_j in the time slot, where P_j will be drawn randomly as discussed below.

We denote h_{xy} as the fading coefficient between transmitter x and receiver y, where x is either "a" (Alice) or "j" (jammer), and y is either "w" (Willie) or "b" (Bob). We assume that the fading processes do not change during one time slot and that those affecting different transmitter-receiver pairs are independent of one another. The path loss between transmitter x and receiver y is denoted as d_{xy}^{α}, where α is the path loss exponent.

Denote the vector of channel outputs observed at Willie over the time slot as $\mathbf{z} = [z_1, z_2, \ldots, z_n]$. Then,

$$z_i = \begin{cases} \frac{h_{aw}}{d_{aw}^{\alpha/2}} \cdot f_i + \frac{h_{jw}}{d_{jw}^{\alpha/2}} \cdot g_i + N_i^{(w)}, & \text{when Alice transmits} \\ \frac{h_{jw}}{d_{jw}^{\alpha/2}} \cdot g_i + N_i^{(w)}, & \text{when Alice does not transmit} \end{cases}$$

where $[N_1^{(w)}, N_2^{(w)}, \ldots, N_n^{(w)}]$ is a set of independent and identically distributed (i.i.d.) zero-mean complex Gaussian random variables, each with variance σ_w^2. Similarly, denote the vector of channel outputs observed at Bob over the time slot as $\mathbf{y} = [y_1, y_2, \ldots, y_n]$. Then.

$$y_i = \begin{cases} \frac{h_{ab}}{d_{ab}^{\alpha/2}} \cdot f_i + \frac{h_{jb}}{d_{jb}^{\alpha/2}} \cdot g_i + N_i^{(b)}, & \text{when Alice transmits} \\ \frac{h_{jb}}{d_{jb}^{\alpha/2}} \cdot g_i + N_i^{(b)}, & \text{when Alice does not transmit} \end{cases}$$

where $[N_1^{(b)}, N_2^{(b)}, \ldots, N_n^{(b)}]$ is a set of i.i.d. zero-mean complex Gaussian random variables, each with variance σ_b^2. Without loss of generality, we let the path loss $d_{xy}^{\alpha/2}$ (from each transmitter x to receiver y) to be one for our proofs to make the exposition cleaner. The results obviously extend to cases with different path losses, and we will present numerical results for the general case.

We assume that Alice has channel state information (CSI) about the channel h_{ab} from herself to Bob. For example, Bob might be allowed to transmit and thus can send a pilot signal or could send it covertly to Alice so that Alice could measure the channel condition. Since Alice learns the channel, she uses a transmit power P_a that adapts to the channel variation (i.e., as a function of $|h_{ab}|^2$). We are interested in finding the optimal function $P_a(|h_{ab}|^2)$ under the constraints and metrics defined in the next two sections. Normally, it is hard for Willie to know h_{ab}, since Willie is in a different location with different reflections. However, h_{ab} is not completely unknown to Willie. For example, Willie might be close to Alice or there could be high channel correlations between Alice's and the adversary's channel even for large spatial separations, resulting in some leakage to Willie about h_{ab} [42, 43]. Willie could also estimate h_{ab} by making a detailed record of the physical environment and the location of Alice and Bob, and then applying a ray-tracing algorithm [44]. Since we want to be pessimistic, we assume conservatively here that Willie knows h_{ab}. We also assume that he knows the function P_a. Thus, Willie knows $P_a(|h_{ab}|^2)$.

We study power adaptation under two channel models: 1) AWGN model and 2) Rayleigh fading model. In the AWGN model, we assume that the Alice-to-Willie, jammer-to-Willie and jammer-to-Bob channels are all AWGN channels, i.e., $h_{aw} = h_{jw} = h_{jb} = 1$. The only fading channel in the system is the Alice-to-Bob channel which experiences Rayleigh fading with $E[|h_{ab}|^2] = 1$ for simplicity. The jammer's power P_j is assumed to be uniformly distributed, i.e., $P_j \sim U[0, P_J]$, per [9]. In the Rayleigh fading model, both the Alice-to-Bob and the jammer-to-Willie channels are Rayleigh fading channels with $E[|h_{ab}|^2] = E[|h_{jw}|^2] = 1$. The other channels are AWGN channels, i.e., $h_{aw} = h_{jb} = 1$ which is pessimistic for Alice. In this Rayleigh fading case, the jammer is assumed to employ a constant power $P_j = P_J$ since the channel randomizes the power received at Willie from the jammer [9].

In this work, we use an outage approach where we assume that the blocklength n is sufficiently large that Willie has an accurate power measurement at his receiver. This approach has been used in recent work (e.g. [24]- [27]) to provide a clear way to understand the underlying mechanisms of the problems.

2.2.2 Receivers and Performance Metrics

1) Willie: Based on his observations over the time slot, Willie attempts to determine whether Alice transmitted or not. We define the null hypothesis (H_0) as that Alice did not transmit during the time slot and the alternative hypothesis (H_1) as that Alice transmitted a message. The optimal test for Willie to minimize the error probability is to employ the likelihood ratio test (LRT):

$$\Lambda(z) = \frac{f_{P_w}(z|H_1)}{f_{P_w}(z|H_0)}$$

where $f_{P_w}(z|H_1)$ and $f_{P_w}(z|H_0)$ are the probability density functions of Willie's received power P_w when H_1 is true and when H_0 is true, respectively. It is straightforward to show that under both the AWGN model and the Rayleigh fading model, $\Lambda(z)$ is non-decreasing as z increases. Thus, the LRT is equivalent to a threshold test on the received power. Willie's received power P_w given H_0 is true and H_1 is true are, respectively:

- H_0: $P_w = |h_{jw}|^2 P_j + \sigma_w^2$;
- H_1: $P_w = |h_{aw}|^2 P_a + |h_{jw}|^2 P_j + \sigma_w^2$.

Willie's detector compares P_w to a threshold $\sigma_w^2 + \tau$:

$$P_w \underset{H_0}{\overset{H_1}{\gtrless}} \sigma_w^2 + \tau.$$

Define $P(H_0) = 1-p$ as the probability that Alice did not transmit and $P(H_1) = p$ as the probability that Alice transmitted in the time slot, where we assume that p is

known to Willie. Willie tries to minimize his probability of error $P_{e,w} = (1-p)P_{FA} + pP_{MD}$, where P_{FA} and P_{MD} are the probabilities of false alarm and missed detection at Willie, respectively. Since $P_{e,w} \geq \min(p, 1-p)(P_{FA} + P_{MD})$ [5], we say that Alice achieves covert communication if, for a given $\epsilon > 0$, $P_{FA} + P_{MD} \geq 1 - \epsilon$ [5].

2) Bob: Bob should be able to reliably decode Alice's message. This is characterized by the probability $1 - P_{out,b}$ where $P_{out,b}$ is the outage probability at Bob, i.e., the probability that Bob's received signal-to-noise ratio (SNR) is below a certain threshold. We will seek to minimize the value of $P_{out,b}$ under power and covertness constraints.

2.2.3 Constraints

Alice wants to use as much power as possible to minimize outage, but she is bounded by average power and covertness constraints.

The average power constraint is given by:

$$E_{|h_{ab}|^2}[P_a(|h_{ab}|^2)] \leq P_A \tag{2.1}$$

where P_A is a constant power budget.

We also have an *average covertness constraint*, i.e., for a given $\epsilon_1 > 0$,

$$E_{|h_{ab}|^2}[P_{FA}(P_a(|h_{ab}|^2)) + P_{MD}(P_a(|h_{ab}|^2))] \geq 1 - \epsilon_1. \tag{2.2}$$

Some of our results will consider only the average covertness constraint, where the averaging is done over the channels that Alice will encounter on the Alice-to-Bob link. However, this constraint might be problematic operationally, as it can result in Alice sometimes sending messages knowing that she will be detected by Willie with certainty, with the idea that, since she will not be caught at other times, the long-term average of her covertness meets the constraint. Since this might not be

desirable, as Alice might want every message to be covert with some probability, we also consider cases where an instantaneous covertness constraint has been added: for a given $\epsilon_2 > 0$,

$$P_{FA}(P_a(|h_{ab}|^2)) + P_{MD}(P_a(|h_{ab}|^2)) \geq 1 - \epsilon_2 \tag{2.3}$$

for all $|h_{ab}|^2$.

2.3 AWGN Model

In this section, we consider the AWGN model, where recall that the Alice-to-Willie, jammer-to-Willie and jammer-to-Bob channels are all AWGN channels, i.e., $h_{aw} = h_{jw} = h_{jb} = 1$, and the Alice-to-Bob channel is assumed to be a Rayleigh fading channel with $E[|h_{ab}|^2] = 1$. We prove an optimal power adaptation scheme for two cases: under the average covertness constraint, and under both the average and the instantaneous covertness constraints.

2.3.1 Average Covertness Constraint

We first consider a power adaptation scheme that minimizes the outage probability at Bob and achieves covertness on average as given in (2.1). For Bob to reliably decode Alice's message, the received SNR at Bob needs to be above a threshold γ_b. Given a transmit power $P_a(|h_{ab}|^2)$, the outage probability at Bob is given by:

$$P_{out,b}(P_a(|h_{ab}|^2)) = P\left(\frac{|h_{ab}|^2 P_a(|h_{ab}|^2)}{\sigma_b^2 + P_j} < \gamma_b\right)$$

where, per Section 2.2, $\frac{|h_{ab}|^2 P_a(|h_{ab}|^2)}{\sigma_b^2 + P_j}$ is random because of the randomly chosen jamming power $P_j \sim U[0, P_J]$. Thus,

$$P_{out,b}(P_a(|h_{ab}|^2)) = \begin{cases} 1 - \frac{\sigma_b^2}{P_J} - \frac{|h_{ab}|^2 P_a(|h_{ab}|^2)}{P_J \gamma_b}, & P_a(h_{ab}) < \frac{\gamma_b(\sigma_b^2 + P_J)}{|h_{ab}|^2} \\ 0, & P_a(h_{ab}) \geq \frac{\gamma_b(\sigma_b^2 + P_J)}{|h_{ab}|^2}. \end{cases}$$

Since employing additional power beyond $\frac{\gamma_b(\sigma_b^2+P_J)}{|h_{ab}|^2}$ does not decrease the outage probability but increases Alice's transmitted power, leading uniformly not only to wasted power but a loss in covertness, we add the constraint:

$$0 \leq P_a(h_{ab}) \leq \frac{\gamma_b(\sigma_b^2 + P_J)}{|h_{ab}|^2}. \tag{2.4}$$

Taking the expectation of $P_{out,b}(P_a(|h_{ab}|^2))$ over $|h_{ab}|^2$ yields:

$$E_{|h_{ab}|^2}[P_{out,b}(P_a(|h_{ab}|^2))] = 1 - \frac{\sigma_b^2}{P_J} - \frac{1}{P_J\gamma_b}\int_0^\infty xP_a(x)e^{-x}dx.$$

Given that $|h_{ab}|^2$ is exponentially distributed with $E[|h_{ab}|^2] = 1$, the average power constraint in (2.1) can be written as:

$$\int_0^\infty P_a(x)e^{-x}dx \leq P_A. \tag{2.5}$$

The false alarm probability at Willie given $P_a(|h_{ab}|^2)$ is given by:

$$\begin{aligned} P_{FA}(P_a(|h_{ab}|^2)) &= P(P_w > \sigma_w^2 + \tau|H_0) \\ &= P(P_j > \tau) \\ &= \begin{cases} \frac{P_J-\tau}{P_J}, & \tau < P_J \\ 0, & \tau \geq P_J, \end{cases} \end{aligned}$$

and the missed detection probability at Willie is given by:

$$\begin{aligned} P_{MD}(P_a(|h_{ab}|^2)) &= P(P_w < \sigma_w^2 + \tau|H_1) \\ &= P(P_a(|h_{ab}|^2) + P_j < \tau) \\ &= \begin{cases} \frac{\tau-P_a(|h_{ab}|^2)}{P_J}, & \tau > P_a(|h_{ab}|^2) \\ 0, & \tau \leq P_a(|h_{ab}|^2). \end{cases} \end{aligned}$$

Willie picks a threshold τ that minimizes his error probability $P_{e,w}$. If $P_a(|h_{ab}|^2) \geq P_J$, then Willie can choose τ such that $P_J \leq \tau \leq P_a(|h_{ab}|^2)$ and achieves $P_{FA}(P_a(|h_{ab}|^2))+P_{MD}(P_a(|h_{ab}|^2)) = 0$. Thus,

$$P_{FA}(P_a(|h_{ab}|^2))+P_{MD}(P_a(|h_{ab}|^2)) = \begin{cases} \frac{P_J - P_a(|h_{ab}|^2)}{P_J}, & P_J > P_a(|h_{ab}|^2) \\ 0, & P_J \leq P_a(|h_{ab}|^2). \end{cases} \tag{2.6}$$

Recall that the Alice-to-Bob channel is Rayleigh fading with $E[|h_{ab}|^2] = 1$. Thus, $|h_{ab}|^2$ is exponentially distributed and taking the expectation of $P_{FA}(P_a(|h_{ab}|^2)) + P_{MD}(P_a(|h_{ab}|^2))$ over $|h_{ab}|^2$ yields:

$$E_{|h_{ab}|^2}[P_{FA}(P_a(|h_{ab}|^2)) + P_{MD}(P_a(|h_{ab}|^2))] = \int_{P_a(x)<P_J} e^{-x}dx - \frac{1}{P_J}\int_{P_a(x)<P_J} P_a(x)e^{-x}dx. \tag{2.7}$$

Then, the average covertness constraint in (2.2) requires:

$$\int_{P_a(x)\geq P_J} e^{-x}dx + \frac{1}{P_J}\int_{P_a(x)<P_J} P_a(x)e^{-x}dx \leq \epsilon_1 \tag{2.8}$$

for a given $\epsilon_1 > 0$.

We are interested in looking for an optimal $P_a(|h_{ab}|^2)$ such that the outage probability at Bob is minimized under the constraints in (2.8), (2.5) and (2.4). Therefore, we form the functional optimization problem:

$$\begin{aligned} \underset{P_a(x)}{\text{maximize:}} \quad & \int_0^\infty xP_a(x)e^{-x}dx \ \ , x \geq 0 \\ \text{subject to:} \quad & \int_{P_a(x)\geq P_J} e^{-x}dx + \frac{1}{P_J}\int_{P_a(x)<P_J} P_a(x)e^{-x}dx \leq \epsilon_1, \\ & \int_0^\infty P_a(x)e^{-x}dx \leq P_A, \\ & 0 \leq P_a(x) \leq \frac{\gamma_b(\sigma_b^2 + P_J)}{x}. \end{aligned} \tag{2.9}$$

The following lemma provides a valuable tool for the succeeding optimizations.

Lemma 2.1. For the optimization problem:

$$\begin{aligned} \underset{f(x)}{\text{maximize:}} \quad & \int_0^\infty x f(x) e^{-x} dx \ \ , x \geq 0 \\ \text{subject to:} \quad & \int_0^\infty f(x) e^{-x} dx \leq C_0 \ \text{ and } \ 0 \leq f(x) \leq g(x) \end{aligned}$$

where C_0 is a constant. If $\int_0^\infty g(x)e^{-x}dx > C_0$, the optimal solution is:

$$f^*(x) = \begin{cases} g(x), & x \geq \Delta \\ 0, & 0 \leq x < \Delta \end{cases} \tag{2.10}$$

where $\Delta > 0$ is such that $\int_\Delta^\infty g(x)e^{-x}dx = C_0$. Otherwise,

$$f^*(x) = g(x) \ \ , \ \ x \geq 0. \tag{2.11}$$

Proof. We prove the optimal solution by showing that any $f(x)$ in the constraint set other than $f^*(x)$ will not increase the objective function. For a given candidate function $f(x)$, we write $f(x) = f^*(x) + u(x)$. If $f(x)$ is in the constraint set, then we must have three conditions: (a) $\int_0^\infty u(x)e^{-x}dx \leq 0$ since $f^*(x)$ already has achieved equality in the first constraint; (b) $u(x) \geq 0$ when $0 \leq x < \Delta$ since $f(x) \geq 0$; and (c) $u(x) \leq 0$ when $\Delta \leq x$ since $f(x) \leq g(x)$. Thus, when $f^*(x)$ in given in (2.10), we have:

$$\begin{aligned} \int_0^\infty x u(x) e^{-x} dx &= \int_0^\Delta x u(x) e^{-x} dx + \int_\Delta^\infty x u(x) e^{-x} dx \\ &\leq \Delta \int_0^\Delta u(x) e^{-x} dx + \Delta \int_\Delta^\infty u(x) e^{-x} dx \qquad (2.12) \\ &\leq -\Delta \int_\Delta^\infty u(x) e^{-x} dx + \Delta \int_\Delta^\infty u(x) e^{-x} dx \qquad (2.13) \\ &= 0 \end{aligned}$$

where (2.12) is obtained from condition (b) and (c), and (2.13) is obtained from condition (a). Therefore, $\int_0^\infty xf(x)e^{-x}dx \leq \int_0^\infty xf^*(x)e^{-x}dx$. When $f^*(x)$ is given in (2.11), it is obvious that $u(x) \leq 0$ for all $x \geq 0$. Then, clearly, $\int_0^\infty xu(x)e^{-x}dx \leq 0$ and $\int_0^\infty xf(x)e^{-x}dx \leq \int_0^\infty xf^*(x)e^{-x}dx$. □

Theorem 2.1. For the optimization problem in (2.9), let $\eta = \frac{\gamma_b(\sigma_b^2+P_J)}{P_J}$, the optimal solution is given by:

$$P_a^*(x) = \begin{cases} \frac{\gamma_b(\sigma_b^2+P_J)}{x}, & x \geq \delta \\ 0, & 0 \leq x < \delta \end{cases} \tag{2.14}$$

where $\delta = \max(\Delta_1, \Delta_2)$, Δ_1 is such that $\int_{\Delta_1}^\infty \frac{\gamma_b(\sigma_b^2+P_J)}{x} e^{-x}dx = P_A$ and Δ_2 is such that:

$$\begin{cases} \int_{\Delta_2}^\eta e^{-x}dx + \frac{1}{P_J}\int_\eta^\infty \frac{\gamma_b(\sigma_b^2+P_J)}{x} e^{-x}dx = \epsilon_1, & \text{if } \int_\eta^\infty \frac{\gamma_b(\sigma_b^2+P_J)}{x} e^{-x}dx \leq P_J\epsilon_1 \\ \int_{\Delta_2}^\infty \frac{\gamma_b(\sigma_b^2+P_J)}{x} e^{-x}dx = P_J\epsilon_1, & \text{otherwise.} \end{cases} \tag{2.15}$$

Proof. The proof makes use of Lemma 2.1 and is provided in detail in the appendix in Section 2.6.1. □

From Theorem 2.1, we conclude that the optimal power adaptation under the average covertness constraint when the jammer-to-Willie channel is AWGN is truncated channel inversion, i.e., transmitting power as the inverse of the channel condition when the channel is good; and transmitting nothing when the channel is bad. Alice transmits with power that inverts the channel variation when the channel condition is good, and does not transmit when the channel condition is bad to save power and to avoid detection by Willie. The constant ($\gamma_b(\sigma_b^2 + P_J)$ in this case) of the inversion scheme is analytically provided and the cutoff point (δ in this case) can be directly computed based on the system parameters.

We notice that if $\delta < \eta$, Alice is allowed to use power greater than P_J. This will lead to an error probability of zero at Willie according to (5.1), which means

that Willie will detect Alice's transmission for certain. However, although Alice is completely exposed to Willie for channel conditions such that $\delta < |h_{ab}|^2 < \eta$, she still achieves covertness on average. We add the extra instantaneous covertness constraint that requires Alice to achieve a certain covertness for any channel condition in the next section.

2.3.2 Instantaneous Covertness Constraint

Now we consider a power adaptation scheme that minimizes the outage probability at Bob and achieves covertness both on average and instantaneously. Recall that the instantaneous covertness constraint is given in (2.3). Applying (2.6) to (2.3) we have for a given $\epsilon_2 > 0$,

$$P_a(|h_{ab}|^2) \leq P_J \epsilon_2,$$

which is equivalent to a peak power constraint.

Since $P_a(|h_{ab}|^2)$ is bounded by $P_J \epsilon_2$, we modify (5.1) in this case as:

$$P_{FA}(P_a(|h_{ab}|^2)) + P_{MD}(P_a(|h_{ab}|^2)) = \frac{P_J - P_a(|h_{ab}|^2)}{P_J},$$

and hence, the average covertness constraint becomes:

$$\int_0^\infty P_a(x) e^{-x} dx \leq P_J \epsilon_1$$

for a given $\epsilon_1 > 0$. This is equivalent to an average power constraint. Thus, the smaller of $P_J \epsilon_1$ and the power budget P_A determines the upper bound on the average power.

The average outage probability is the same and we still have the constraint in (2.4). Therefore, the optimization problem in this case is given by:

$$\begin{aligned}
\underset{P_a(x)}{\text{maximize:}} \quad & \int_0^\infty x P_a(x) e^{-x} dx \ \ , x \geq 0 \\
\text{subject to:} \quad & \int_0^\infty P_a(x) e^{-x} dx \leq \min(P_J \epsilon_1, P_A), \\
& 0 \leq P_a(x) \leq \begin{cases} \frac{\gamma_b(\sigma_b^2 + P_J)}{x}, & x > \frac{\gamma_b(\sigma_b^2 + P_J)}{P_J \epsilon_2} \\ P_J \epsilon_2, & x \leq \frac{\gamma_b(\sigma_b^2 + P_J)}{P_J \epsilon_2}. \end{cases}
\end{aligned}$$

Note that this problem is different from that encountered in standard (non-covert) communication systems: 1) with an average power constraint [38] where the optimal solution that maximizes the average capacity is "water-pouring" and the truncated channel inversion scheme is suboptimal as it only achieves a certain outage capacity; 2) with an extra peak power constraint [45]. Here we use a different metric (average outage probability), and our system employs a jammer that interferes with reception at Bob and thus affects the outage probability.

Theorem 2.2. Let

$$B_a(x) = \begin{cases} \frac{\gamma_b(\sigma_b^2 + P_J)}{x}, & x > \frac{\gamma_b(\sigma_b^2 + P_J)}{P_J \epsilon_2} \\ P_J \epsilon_2, & x < \frac{\gamma_b(\sigma_b^2 + P_J)}{P_J \epsilon_2} \end{cases} \tag{2.16}$$

and β be such that $\int_\beta^\infty B_a(x) e^{-x} dx = \min(P_J \epsilon_1, P_A)$, if $\int_0^\infty B_a(x) e^{-x} dx \geq \min(P_J \epsilon_1, P_A)$, the optimal solution to the above optimization problem is given by:

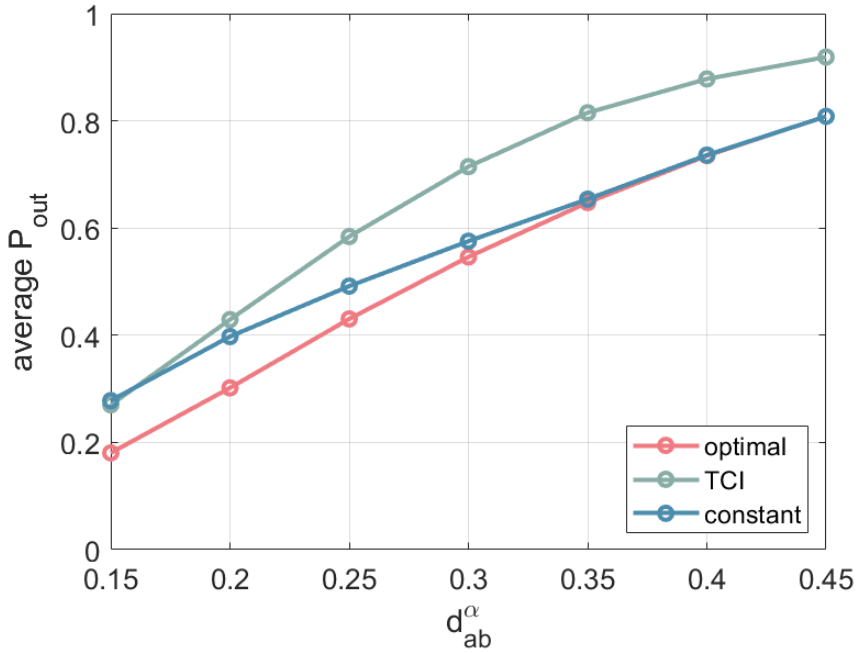
$$P_a^*(x) = \begin{cases} B_a(x), & x \geq \beta \\ 0, & 0 \leq x < \beta, \end{cases}$$

and otherwise,

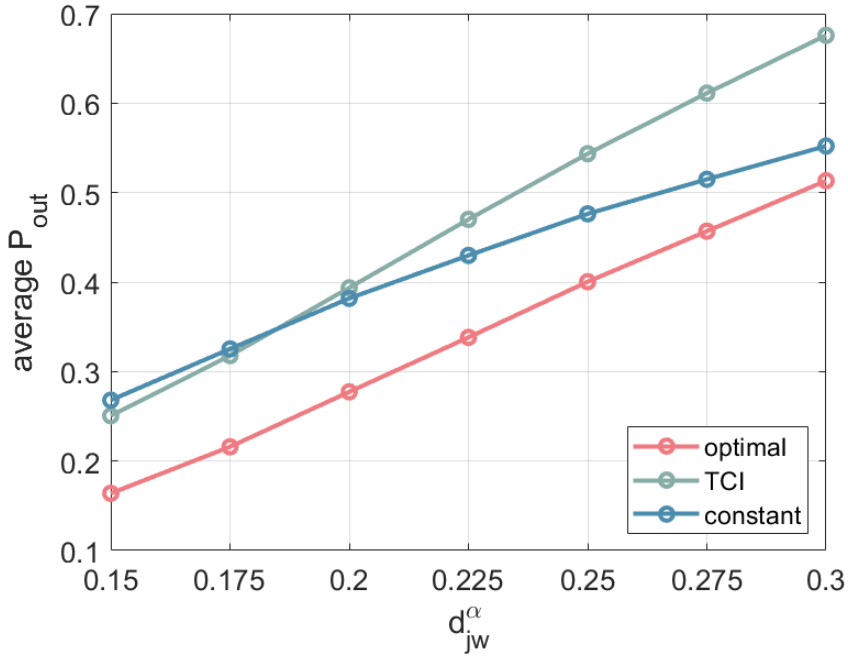
$$P_a^*(x) = B_a(x) \ , \ x \geq 0.$$

Proof. This can be shown by Lemma 2.1 with $C_0 = P_J\epsilon_1$, $g(x) = B_a(x)$ and $\Delta = \beta$ if $\int_0^\infty B_a(x)e^{-x}dx \geq \min(P_J\epsilon_1, P_A)$. $\square$

From Theorem 2.2 we see that the optimal power adaptation scheme is determined by the parameter β. If $\beta > \frac{\gamma_b(\sigma_b^2+P_J)}{P_J\epsilon_2}$, then the optimal adaptation scheme is truncated channel inversion. Otherwise, the power is limited by a constant for certain channel conditions due to the instantaneous covertness constraint.



(a) AWGN case: Average outage probability P_{out} in terms of the path loss d_{ab}^{α} between Alice and Bob.



(b) AWGN case: Average outage probability P_{out} in terms of the path loss d_{jw}^{α} between the jammer and Willie.

Figure 2.2: Average outage probability in the case under the extra instantaneous covertness constraint when Alice employs optimal power adaptation, TCI, and truncated constant power.

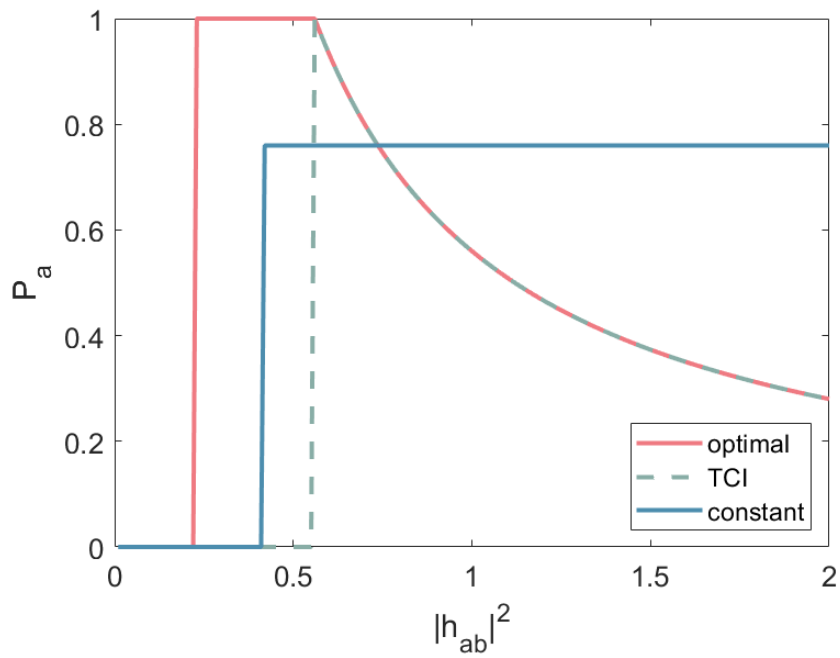


Figure 2.3: Power allocations for the optimal scheme, TCI and the truncated constant scheme.

Figure 2.2 presents the numerical results that examine the performance improvement of the optimal scheme in Theorem 2.2 over TCI and truncated constant power. The parameters in TCI and truncated constant power are obtained by numerical search such that the average outage probability is minimized. We set $P_A = 1$, $P_J = 5$, $\epsilon_1 = 0.1$, $\epsilon_2 = 0.2$, $\gamma_b = 2$, $\sigma_b^2 = 2$, and the path loss between each transmitter and receiver is one if not specified. The exact power allocations for the three schemes given the above parameters with $d_{ab}^{\alpha} = 0.2$ are shown in Figure 2.3. From both Fig. 2.2a and Fig. 2.2b, we see that there are often significant gains in achieving reliable covert communication when employing optimal power adaptation. In Fig. 2.2a, we observe that the performance of the optimal scheme and that of the constant power scheme approach each other when d_{ab}^{α} increases. This is due to the fact that the instantaneous covertness constraint limits Alice's power to a constant for channel conditions that are likely to occur, and Alice uses channel inversion only for channel conditions that

rarely occur for large d_{ab}^{α}. Mathematically, when d_{ab}^{α} increases, the channel conditions that Alice uses constant power also increases ($B_a(x) = P_J\epsilon_2$ when $x < \frac{d_{ab}^{2\alpha}\gamma_b(\sigma_b^2+P_J)}{P_J\epsilon_2}$ from Theorem 2.2) and occupies more of the probabilities.

2.4 Rayleigh Fading Model

In this section, we consider the Rayleigh fading model, where both the Alice-to-Bob and the jammer-to-Willie channels are Rayleigh fading channels with $E[|h_{ab}|^2] = E[|h_{jw}|^2] = 1$. The other channels are AWGN channels, i.e., $h_{aw} = h_{jb} = 1$. We assume conservatively that Bob is not able to cancel the jamming signal at his receiver. We prove an optimal power adaptation scheme under both the average covertness constraint and with the extra instantaneous covertness constraint.

2.4.1 Average Covertness Constraint

Given $P_a(|h_{ab}|^2)$, the outage probability at Bob is:

$$P_{out,b(P_a(|h_{ab}|^2))} = \begin{cases} 1, & |h_{ab}|^2 P_a(|h_{ab}|^2) < \gamma_b\left(\sigma_b^2 + P_J\right) \\ 0, & |h_{ab}|^2 P_a(|h_{ab}|^2) \geq \gamma_b\left(\sigma_b^2 + P_J\right). \end{cases}$$

Since letting $P_a(|h_{ab}|^2) > \frac{\gamma_b\left(\sigma_b^2+P_J\right)}{|h_{ab}|^2}$ or letting $P_a(|h_{ab}|^2) > 0$ for $P_a(|h_{ab}|^2) < \frac{\gamma_b\left(\sigma_b^2+P_J\right)}{|h_{ab}|^2}$ does not change the outage probability but increases Alice's power, $P_a(|h_{ab}|^2)$ should always take a value of either zero or $\frac{\gamma_b\left(\sigma_b^2+P_J\right)}{|h_{ab}|^2}$. Thus, taking the expectation over $|h_{ab}|^2$, the average outage probability is given by:

$$\begin{aligned} E_{|h_{ab}|^2}[P_{out,b(P_a(|h_{ab}|^2))}] &= \int_0^\infty I_{P_a(x)=0}(x)e^{-x}dx \\ &= 1 - \int_0^\infty I_{P_a(x)=\frac{\gamma_b(\sigma_b^2+P_J)}{x}}(x)e^{-x}dx \end{aligned}$$

where $I_A(x) = \begin{cases} 1, & x \in A \\ 0, & x \notin A \end{cases}$ is the indicator function of set A.

We obtain the probability of false alarm as:

$$\begin{aligned} P_{FA}(P_a(|h_{ab}|^2)) &= P\left(|h_{jw}|^2 P_J > \tau\right) \\ &= e^{-\frac{\tau}{P_J}}, \end{aligned}$$

and the probability of missed detection:

$$\begin{aligned} P_{MD}(P_a(|h_{ab}|^2)) &= P\left(P_a(|h_{ab}|^2) + |h_{jw}|^2 P_J < \tau\right) \\ &= \begin{cases} 1 - e^{-\frac{\tau - P_a(|h_{ab}|^2)}{P_J}}, & \tau > P_a(|h_{ab}|^2) \\ 0, & \tau \leq P_a(|h_{ab}|^2). \end{cases} \end{aligned}$$

Recall that we assume conservatively here that Willie knows $P_a(|h_{ab}|^2)$. Then, to minimize his error probability, Willie will choose a threshold $\tau = P_a(|h_{ab}|^2)$. Thus, we have:

$$P_{FA}(P_a(|h_{ab}|^2)) + P_{MD}(P_a(|h_{ab}|^2)) = e^{-\frac{P_a(|h_{ab}|^2)}{P_J}}, \tag{2.17}$$

and taking the expectation over $|h_{ab}|^2$ yields:

$$E_{|h_{ab}|^2}[P_{FA}(P_a(|h_{ab}|^2) + P_{MD}(P_a(|h_{ab}|^2))] = \int_0^\infty e^{-\frac{P_a(x)}{P_J}} e^{-x} dx. \tag{2.18}$$

Applying (2.18) to the average covertness constraint in (2.2) yields:

$$\int_0^\infty e^{-\frac{P_a(x)}{P_J}} e^{-x} dx \geq 1 - \epsilon_1$$

for a given $\epsilon_1 > 0$.

With the average power constraint in (2.5), we form the optimization problem as:

$$\begin{aligned} \underset{P_a(x)}{\text{maximize:}} \quad & \int_0^\infty I_{P_a(x)=\frac{\gamma_b(\sigma_b^2+P_J)}{x}}(x)e^{-x}dx \\ \text{subject to:} \quad & \int_0^\infty e^{-\frac{P_a(x)}{P_J}}e^{-x}dx \geq 1-\epsilon_1, \\ & \int_0^\infty P_a(x)e^{-x}dx \leq P_A. \end{aligned} \tag{2.19}$$

Since we focus on reasonable power schemes that could be implemented in practice, we only consider power control strategies that employ a finite collection of intervals. Therefore, the theorem below is obtained by assuming that the optimal strategy does not include singletons or an infinite number of intervals. In other words, recalling that $P_a(x)$ should always take a value of either zero or $\frac{\gamma_b(\sigma_b^2+P_J)}{x}$ since any non-zero power less than $\frac{\gamma_b(\sigma_b^2+P_J)}{x}$ is wasted as it does not change the outage probability, we assume:

$$P_a(x) = \begin{cases} \frac{\gamma_b(\sigma_b^2+P_J)}{x}, & x \in [a_1, b_1) \cup \ldots \cup [a_n, b_n) \\ 0, & \text{else} \end{cases} \tag{2.20}$$

where $n \in \mathbb{Z}^+$ and $a_n, b_n \in \mathbb{R}^+$.

Theorem 2.3. Given the structure in (2.20) of the potential strategies, the optimal solution to the problem in (2.19) is given by:

$$P_a^*(x) = \begin{cases} \frac{\gamma_b(\sigma_b^2+P_J)}{x}, & x > \mu \\ 0, & x \leq \mu \end{cases}$$

where $\mu = \max(\mu_1, \mu_2)$, μ_1 is such that $\int_0^{\mu_1} e^{-x}dx + \int_{\mu_1}^\infty e^{-\frac{\gamma_b\left(\sigma_b^2+P_J\right)}{xP_J}}e^{-x}dx = 1-\epsilon_1$ and μ_2 is such that $\int_{\mu_2}^\infty \frac{\gamma_b\left(\sigma_b^2+P_J\right)}{x}e^{-x}dx = P_A$.

Proof. Suppose that there is a better solution than $P_a^*(x)$ given in the theorem statement. Denote that solution by (2.20), where $0 \leq a_1 < \mu$, $b_n \leq \infty$, and $a_n > b_{n-1}$ if $n > 1$. $P_a(x)$ must satisfy both the average covertness and power constraints. On the other hand, $P_a^*(x)$ also satisfies the two constraints and achieves equality in at least one of the constraints by construction. Suppose that $P_a^*(x)$ achieves equality in the average power constraint; then, for $P_a(x)$ to satisfy this constraint, we must have:

$$\sum_{i=1}^{n} \int_{a_i}^{b_i} \frac{\gamma_b(\sigma_b^2 + P_J)}{x} e^{-x} dx \leq \int_{\mu}^{\infty} \frac{\gamma_b(\sigma_b^2 + P_J)}{x} e^{-x} dx. \tag{2.21}$$

However, in Appendix B, we show that in satisfying (2.21), the objective function in (2.19) cannot get larger, and hence, $P_a(x)$ is not optimal.

Now we look at the average covertness constraint, which can be modified as:

$$\int_0^{\infty} \left(1 - e^{-\frac{P_a(x)}{P_J}}\right) e^{-x} dx \leq \epsilon_1$$

for a given $\epsilon_1 > 0$. Suppose that $P_a^*(x)$ achieves equality in the average covertness constraint; then, for $P_a(x)$ to satisfy this constraint, we must have:

$$\sum_{i=1}^{n} \int_{b_i}^{a_i} \left(1 - e^{-\frac{\gamma_b(\sigma_b^2 + P_J)}{P_J x}}\right) e^{-x} dx \leq \int_{\mu}^{\infty} \left(1 - e^{-\frac{\gamma_b(\sigma_b^2 + P_J)}{P_J x}}\right) e^{-x} dx. \tag{2.22}$$

Then, following a similar argument (details provided in the appendix in Section 2.6.2) when equality in the average power constraint is achieved by $P_a^*(x)$, we can show that for $P_a(x)$ to satisfy (2.22), the objective function in (2.19) will not get larger. Therefore, $P_a(x)$ is not optimal in any case, which completes the proof.

□

We conclude from Theorem 2.3 that in the case when the jammer-to-Willie channel is faded and we are operating under the average covertness constraint, the optimal

power adaptation is truncated channel inversion. Again, the constant of the scheme is analytically provided and the cutoff point can be directly computed based on the system parameters.

2.4.2 Instantaneous Covertness Constraint

Now we consider the case that we also have the instantaneous covertness constraint (2.3) which bounds (2.17) by $1-\epsilon_2$:

$$P_a(|h_{ab}|^2) \leq P_J \ln \frac{1}{1-\epsilon_2}$$

for a given $\epsilon_2 > 0$. In this case, the optimization problem is given by:

$$\begin{aligned}
\underset{P_a(x)}{\text{maximize:}} \quad & \int_0^\infty I_{P_a(x)=\frac{\gamma_b(\sigma_b^2+P_J)}{x}}(x) e^{-x} dx \\
\text{subject to:} \quad & \int_0^\infty e^{-\frac{P_a(x)}{P_J}} e^{-x} dx \geq 1-\epsilon_1, \\
& \int_0^\infty P_a(x) e^{-x} dx \leq P_A, \\
& P_a(x) \leq P_J \ln \frac{1}{1-\epsilon_2}.
\end{aligned}$$

The optimal solution provided in Theorem 2.4 follows from Theorem 2.3.

Theorem 2.4. Let

$$Q_a(x) = \begin{cases} \frac{\gamma_b(\sigma_b^2+P_J)}{x}, & x > \frac{\gamma_b(\sigma_b^2+P_J)}{P_J \ln \frac{1}{1-\epsilon_2}} \\ P_J \ln \frac{1}{1-\epsilon_2}, & x \leq \frac{\gamma_b(\sigma_b^2+P_J)}{P_J \ln \frac{1}{1-\epsilon_2}} \end{cases} \tag{2.23}$$

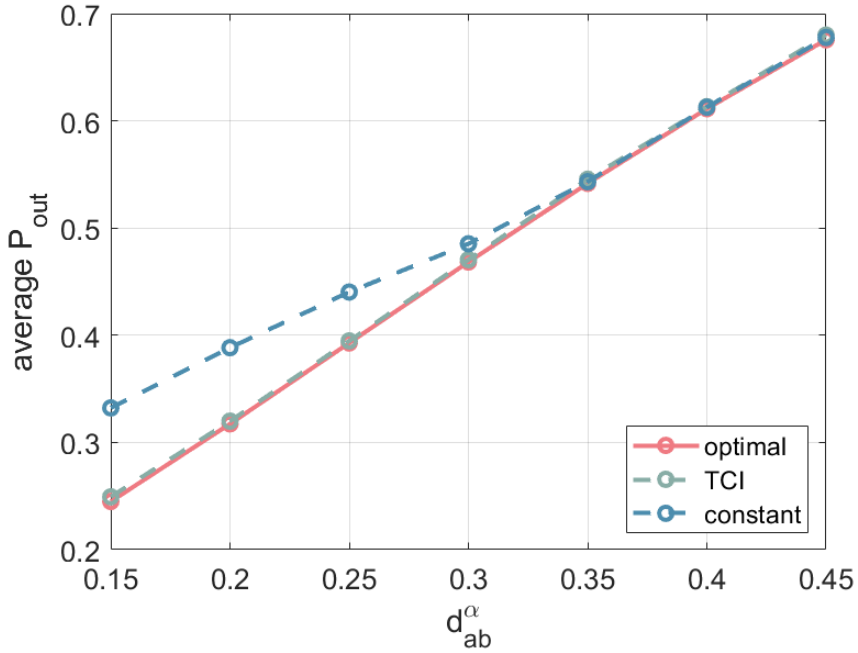
and $\lambda = \max(\lambda_1, \lambda_2)$, where λ_1 is such that $\int_0^{\lambda_1} e^{-x} dx + \int_{\lambda_1}^\infty e^{-\frac{Q_a(x)}{P_J}} e^{-x} dx = 1-\epsilon_1$ and λ_2 is such that $\int_{\lambda_2}^\infty Q_a(x) e^{-x} dx = P_A$. If either $\int_0^\infty e^{-\frac{Q_a(x)}{P_J}} e^{-x} dx \leq 1-\epsilon_1$ or $\int_0^\infty Q_a(x) e^{-x} dx \geq P_A$, the optimal solution is given by:

$$P_a^*(x) = \begin{cases} Q_a(x), & x > \lambda \\ 0, & 0 \leq x \leq \lambda, \end{cases}$$

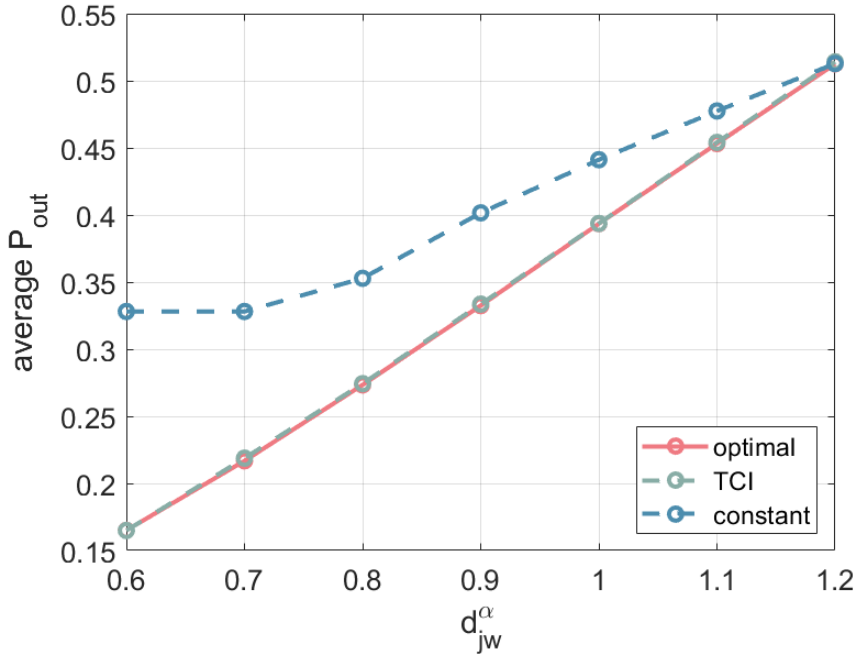
and otherwise,

$$P_a^*(x) = Q_a(x) \ , \ x \geq 0.$$

From Theorem 2.4 we conclude that the optimal power adaptation in the case that the jammer-to-Willie channel is faded and under the extra instantaneous covertness constraint is TCI if $\lambda > \frac{\gamma_b(\sigma_b^2+P_J)}{P_J \ln \frac{1}{1-\epsilon_2}}$. Otherwise, the power is limited by a constant for certain channel conditions due to the instantaneous covertness constraint.

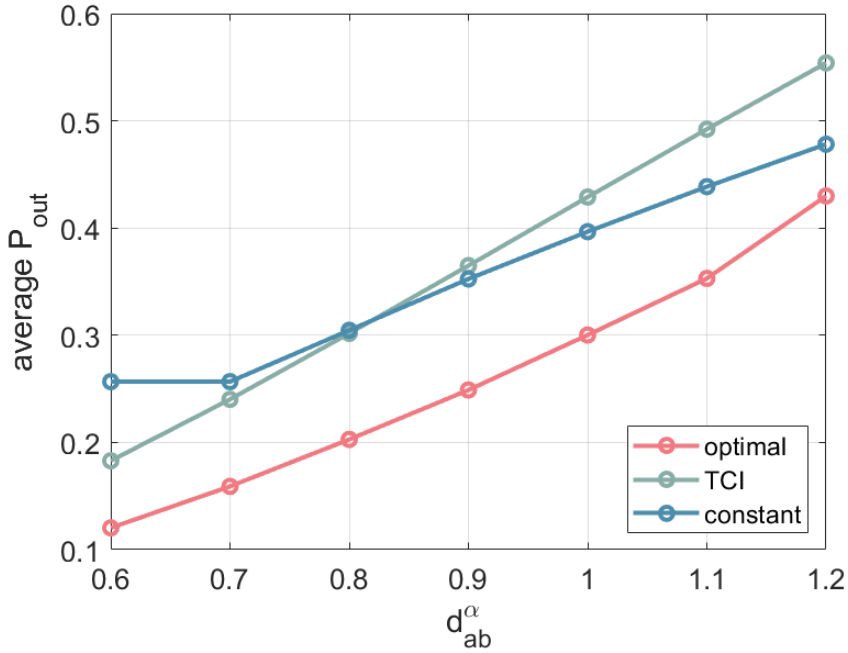


(a) Rayleigh fading case: Average outage probability P_{out} in terms of the path loss d_{ab}^{α} between Alice and Bob.

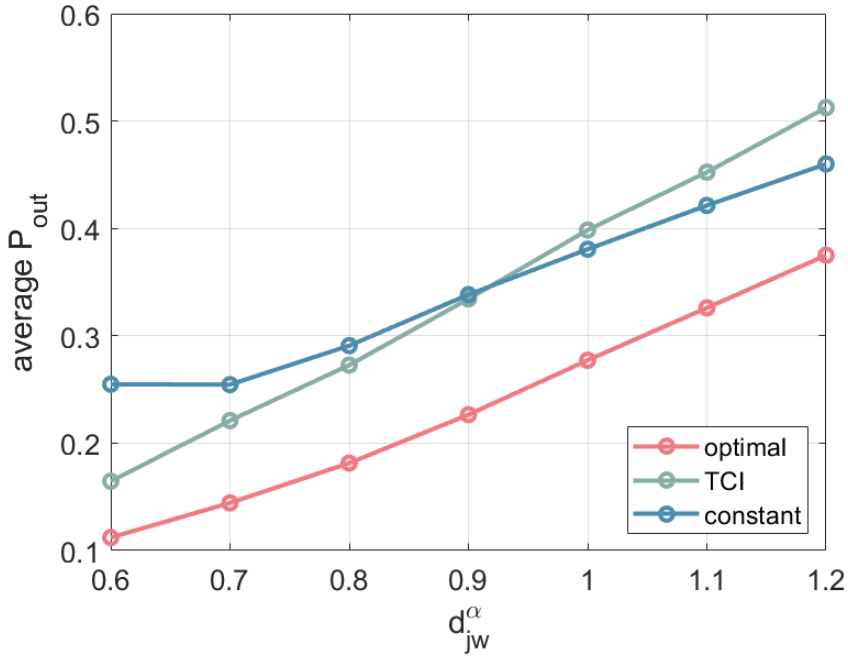


(b) Rayleigh fading case: Average outage probability P_{out} in terms of the path loss d_{jw}^{α} between the jammer and Willie.

Figure 2.4: Average outage probability under both the average and the instantaneous covertness constraints when Alice employs optimal power adaptation, TCI, and truncated constant power The jammer employs constant power.



(a) Rayleigh fading case: Average outage probability P_{out} in terms of the path loss d_{ab}^{α} between Alice and Bob.



(b) Rayleigh fading case: Average outage probability P_{out} in terms of the path loss d_{jw}^{α} between the jammer and Willie.

Figure 2.5: Average outage probability under both the average and the instantaneous covertness constraints when Alice employs optimal power adaptation, TCI, and truncated constant power the jammer employs uniformly distributed power.

Figure 2.4 presents the numerical results that examine the performance of the optimal scheme in Theorem 2.4 and that of TCI and truncated constant power. The parameters in TCI and truncated constant power are obtained by numerical search such that the average outage probability of the scheme is minimized. We set $P_A = 1$, $P_J = 5$, $\epsilon_1 = 0.1$, $\epsilon_2 = 0.2$, $\gamma_b = 2$, $\sigma_b^2 = 2$ and the pathloss between each transmitter and receiver as one if not specified. The jammer's power is set to be a constant. We see from both Fig. 2.4a and Fig. 2.4b that, unlike the AWGN case, the optimal schemes do not have a significant performance gain over TCI. Note that this is for the case when the jammer's power is set to be a constant as suggested in [9], since the Rayleigh fading channel randomizes the jammer's power and hence makes it unknown at Willie. However, it is very unlikely that the jammer power would be set to a constant in practice, because such a setting would be fragile: the nature of the potential operating environment is generally uncertain and, if the environment encountered were AWGN rather than Rayleigh fading, Willie would know the jamming plus noise power exactly when Alice is not transmitting and covertness would be lost for any positive rate [5]. Therefore, it is of interest to consider a random power at the jammer in the Rayleigh fading case. If the jammer employs a uniformly distributed power and we adopt the optimal power adaptation scheme from Theorem 2.4, we can observe from Fig. 2.5 that the optimal scheme significantly outperforms TCI and the truncated constant power scheme.

2.5 Conclusion

In this chapter, we have considered covert communication with help from an uninformed jammer. We have taken an outage approach that considers an infinite blocklength, which allows Willie to have an accurate estimate of his received power, and established optimal power adaptation schemes in different scenarios. We proved that in the case of an AWGN channel or a Rayleigh fading channel between the

jammer and Willie, the optimal scheme under the average covertness constraint is truncated channel inversion. Under both the average and the instantaneous covertness constraint, TCI is not optimal, and we have provided exact optimal power adaptation schemes. These schemes outperform standard approaches such as TCI by reducing the average outage probability significantly in many scenarios.

2.6 Appendix

2.6.1 Proof of Theorem 2.1

Here we privide the detailed proof of Theorem 2.1. Let us denote the first constraint (average covertness constraint) in (2.9) as CSTR1, the second constraint (average power constraint) in (2.9) as CSTR2 and the last constraint in (2.9) as CSTR3. We first remove CSTR1 while keep CSTR2 and CSTR3 in the optimization problem. Then, we note that there always exists $\Delta_1 > 0$ such that $\int_{\Delta_1}^{\infty} \frac{\gamma_b(\sigma_b^2+P_J)}{x} e^{-x} dx = P_A$ is satisfied. By Lemma 2.1, the optimal solution can be obtained as in (2.14) with $\delta = \Delta_1$. Next, we remove CSTR2 but keep CSTR1 and CSTR3. We show in the appendix that the optimal solution can also be obtained as in (2.14) using similar ideas in the proof of Lemma 2.1.

From the discussion above, we see that if $\Delta_1 \geq \Delta_2$, then the optimal solution that satisfies the two constraints CSTR2 and CSTR3 also satisfies all three constraints and hence is optimal when all three constraints are enforced; when $\Delta_1 < \Delta_2$, the optimal solution that satisfies CSTR1 and CSTR3 also satisfies all three constraints and hence is optimal under all three constraints. Moreover, since we pick $\delta = \max(\Delta_1, \Delta_2)$ and $\Delta_1 > 0$, we will never have the case that $P_a^*(x) = \frac{\gamma_b(\sigma_b^2+P_J)}{x}$ for all $x > 0$ (having $\delta = 0$). Therefore, the optimal solution under the three constraints in (2.9) is obtained as in (2.14) with the cutoff point δ determined by either Δ_1 or Δ_2, i.e., $\delta = \max(\Delta_1, \Delta_2)$.

Now we show why (2.14) is the optimal solution to the problem in (2.9) with the second constraint (CSTR2) removed. Letting $J(P_a(x)) = \int_0^{\infty} x P_a(x) e^{-x} dx$, we

take the functional derivative of J (which is the left hand side of the Euler-Lagrange equation [46]):

$$\frac{\partial J(P_a(x))}{\partial P_a(x)} = xe^{-x} \ , \ \ x \geq 0.$$

This cannot equal zero, which shows that there are no stationary points. Thus, if there exists an optimal solution, it should be on the boundary of the set determined by the constraints. From the third constraint, we note that the optimal $P_a(x)$ must be either zero or $\frac{\gamma_b(\sigma_b^2+P_J)}{x}$ for all $x \geq 0$ to lie on the boundary of the constraint set.

In the case when there exists $\Delta_2 > 0$ such that $\int_{\Delta_2}^{\eta} e^{-x}dx + \frac{1}{P_J}\int_{\eta}^{\infty} \frac{\gamma_b(\sigma_b^2+P_J)}{x} e^{-x}dx = \epsilon_1$ is satisfied, suppose $\int_{\eta}^{\infty} \frac{\gamma_b(\sigma_b^2+P_J)}{x} e^{-x}dx \leq P_J\epsilon_1$, then $\Delta_2 \leq \eta$. For a given function $P_a(x)$ other than $P_a^*(x)$, we write $P_a(x) = P_a^*(x) + u(x)$. If $P_a(x)$ is in the constraint set (satisfies CSTR1 and CSTR2) and is a possible optimal solution (is either zero or $\frac{\gamma_b(\sigma_b^2+P_J)}{x}$ for all $x \geq 0$), then we must have three conditions (note that $\frac{\gamma_b(\sigma_b^2+P_J)}{x} < P_J$ when $x > \eta$, and $\frac{\gamma_b(\sigma_b^2+P_J)}{x} \geq P_J$ when $x \leq \eta$):

- (a) $u(x)$ is zero or $\frac{\gamma_b(\sigma_b^2+P_J)}{x}$ for $0 \leq x < \Delta_2$ since $P_a(x) \geq 0$;
- (b) $u(x)$ is either zero or $-\frac{\gamma_b(\sigma_b^2+P_J)}{x}$ for $x \geq \Delta_2$ since $P_a(x) \leq \frac{\gamma_b(\sigma_b^2+P_J)}{x}$;
- (c) $\int_0^{\Delta_2} I_{u(x)=\frac{\gamma_b(\sigma_b^2+P_J)}{x}}(x)e^{-x}dx - \int_{\Delta_2}^{\eta} I_{u(x)=-\frac{\gamma_b(\sigma_b^2+P_J)}{x}} e^{-x}dx + \frac{1}{P_J}\int_{\eta}^{\infty} u(x)e^{-x}dx \leq 0$ in order for $P_a(x)$ to satisfy the average covertness constraint since the equality is already reached for $P_a^*(x)$. Here, $I_A(x) = \begin{cases} 1, & x \in A \\ 0, & x \notin A \end{cases}$ is the indicator function of set A.

Therefore, we have:

$$
\begin{aligned}
\int_0^\infty xu(x)e^{-x}dx &= \int_0^{\Delta_2} xu(x)e^{-x}dx + \int_{\Delta_2}^{\eta} xu(x)e^{-x}dx + \int_{\eta}^{\infty} xu(x)e^{-x}dx \\
&= \int_0^{\Delta_2} x \cdot \frac{\gamma_b(\sigma_b^2 + P_J)}{x} I_{u(x)=\frac{\gamma_b(\sigma_b^2+P_J)}{x}}(x)e^{-x}dx \\
&\quad - \int_{\Delta_2}^{\eta} x \cdot \frac{\gamma_b(\sigma_b^2 + P_J)}{x} I_{u(x)=-\frac{\gamma_b(\sigma_b^2+P_J)}{x}}(x)e^{-x}dx + \int_{\eta}^{\infty} xu(x)e^{-x}dx \qquad (2.24) \\
&\leq -\frac{\gamma_b(\sigma_b^2 + P_J)}{P_J} \int_{\eta}^{\infty} u(x)e^{-x}dx + \int_{\eta}^{\infty} xu(x)e^{-x}dx \qquad (2.25) \\
&= \int_{\eta}^{\infty} \left(x - \frac{\gamma_b(\sigma_b^2 + P_J)}{P_J} \right) u(x)e^{-x}dx \\
&\leq 0 \qquad (2.26)
\end{aligned}
$$

where (2.24) is obtained from condition (a) and (b), (2.25) is obtained from condition (c), and (2.26) is obtained from the fact that $\frac{\gamma_b(\sigma_b^2+P_J)}{x} \leq P_J$ and $u(x) \leq 0$ for $x \geq \eta$.

Suppose $\int_\eta^\infty \frac{\gamma_b(\sigma_b^2+P_J)}{x} e^{-x}dx > P_J\epsilon_1$; then, $\Delta_2 > \eta$. Similarly, we can show that $\int_0^\infty xu(x)e^{-x}dx \leq 0$ in this case by changing condition (c) to: $\int_0^\eta I_{u(x)=\frac{\gamma_b(\sigma_b^2+P_J)}{x}}(x)e^{-x}dx + \frac{1}{P_J}\int_{\Delta_2}^\infty u(x)e^{-x}dx \leq 0$ and prove accordingly.

In the case when there does not exist $\Delta_2 > 0$ such that

$$
\int_{\Delta_2}^{\eta} e^{-x}dx + \frac{1}{P_J}\int_{\eta}^{\infty} \frac{\gamma_b(\sigma_b^2 + P_J)}{x} e^{-x}dx = \epsilon_1 \,,
$$

the optimal solution should be $P_a^*(x) = \frac{\gamma_b(\sigma_b^2+P_J)}{x}$ for all $x > 0$. In this case, we know that $u(x)$ must be $-\frac{\gamma_b(\sigma_b^2+P_J)}{x}$ for all $x > 0$ for $P_a(x)$ to stay in the constraint set. Then, it is clear that $\int_0^\infty xu(x)e^{-x}dx \leq 0$, which shows the optimality of $P_a^*(x)$. This is equivalent to having $\Delta_2 = 0$ in (2.14).

2.6.2 Proof of Theorem 2.3

Here we prove that the power function $P_a(x)$ given in (2.20) is not optimal since it fails to satisfy (2.21). With $b_n < \infty$, (2.21) implies that:

$$\frac{\gamma_b(\sigma_b^2+P_J)}{b_n}\sum_{i=1}^{n}\int_{a_i}^{b_1} e^{-x}dx \le \frac{\gamma_b(\sigma_b^2+P_J)}{\mu}\int_{\mu}^{\infty} e^{-x}dx.$$

When $b_n \le \mu$,

$$\sum_{i=1}^{n}\int_{a_i}^{b_1} e^{-x}dx \le \frac{b_n}{\mu}\int_{\mu}^{\infty} e^{-x}dx \le \int_{\mu}^{\infty} e^{-x}dx,$$

which shows:

$$\int_0^{\infty} I_{P_a(x)=\frac{\gamma_b(\sigma_b^2+P_J)}{x}}(x)e^{-x}dx < \int_0^{\infty} I_{P_a^*(x)=\frac{\gamma_b(\sigma_b^2+P_J)}{x}}(x)e^{-x}dx. \tag{2.27}$$

When $b_n > \mu$, first suppose that $a_m \le \mu < b_m$ and $1 \le m \le n$, from (2.21) we have:

$$\begin{aligned}&\sum_{i=1}^{m-1}\int_{a_i}^{b_i}\frac{\gamma_b(\sigma_b^2+P_J)}{x}e^{-x}dx + \int_{a_m}^{\mu}\frac{\gamma_b(\sigma_b^2+P_J)}{x}e^{-x}dx\\ &\le \sum_{i=m}^{n-1}\int_{b_i}^{a_{i+1}}\frac{\gamma_b(\sigma_b^2+P_J)}{x}e^{-x}dx + \int_{b_n}^{\infty}\frac{\gamma_b(\sigma_b^2+P_J)}{x}e^{-x}dx\end{aligned}$$

where $\sum_{i=m}^{n-1}\int_{b_i}^{a_{i+1}}\frac{\gamma_b(\sigma_b^2+P_J)}{x}e^{-x}dx$ is replaced by zero when $m=n$. This implies that:

$$\begin{aligned}&\frac{\gamma_b(\sigma_b^2+P_J)}{\mu}\left(\sum_{i=1}^{m-1}\int_{a_1}^{b_i} e^{-x}dx + \int_{a_m}^{\mu} e^{-x}dx\right)\\ &\le \frac{\gamma_b(\sigma_b^2+P_J)}{b_m}\sum_{i=m}^{n-1}\int_{b_i}^{a_{i+1}} e^{-x}dx + \frac{\gamma_b(\sigma_b^2+P_J)}{b_n}\int_{b_n}^{\infty} e^{-x}dx,\end{aligned}$$

and then by the assumption that $\mu < b_m \le b_n$, we have:

$$\sum_{i=1}^{m-1}\int_{a_i}^{b_i} e^{-x}dx + \int_{a_m}^{\mu} e^{-x}dx \le \sum_{i=m}^{n-1}\int_{b_i}^{a_{i+1}} e^{-x}dx + \int_{b_n}^{\infty} e^{-x}dx.$$

Thus,

$$\begin{aligned}\sum_{i=1}^{n}\int_{a_i}^{b_i} e^{-x}dx &= \sum_{i=1}^{m-1}\int_{a_i}^{b_i} e^{-x}dx + \int_{a_m}^{\mu} e^{-x}dx - \sum_{i=m}^{n-1}\int_{b_i}^{a_{i+1}} e^{-x}dx + \int_{\mu}^{b_n} e^{-x}dx \\ &\leq \int_{b_n}^{\infty} e^{-x}dx + \int_{\mu}^{b_n} e^{-x}dx \\ &= \int_{\mu}^{\infty} e^{-x}dx\end{aligned}$$

which establishes (2.27).

Now, supposing that $b_m \leq \mu < a_{m+1}$ and $1 \leq m < n$ when $b_n > \mu$, the proof follows similar ideas. From (2.21) we have:

$$\begin{aligned}&\sum_{i=1}^{m}\int_{a_i}^{b_i} \frac{\gamma_b(\sigma_b^2 + P_J)}{x} e^{-x}dx \\ &\leq \int_{\mu}^{a_{m+1}} \frac{\gamma_b(\sigma_b^2 + P_J)}{x} e^{-x}dx + \sum_{i=m}^{n-2}\int_{b_{i+1}}^{a_{i+2}} \frac{\gamma_b(\sigma_b^2 + P_J)}{x} e^{-x}dx + \int_{b_n}^{\infty} \frac{\gamma_b(\sigma_b^2 + P_J)}{x} e^{-x}dx\end{aligned}$$

where $\sum_{i=m}^{n-2}\int_{b_{i+1}}^{a_{i+2}} \frac{\gamma_b(\sigma_b^2+P_J)}{x} e^{-x}dx$ is replaced by zero when $m = n-1$. This implies that:

$$\begin{aligned}&\frac{\gamma_b(\sigma_b^2 + P_J)}{b_m}\sum_{i=1}^{m}\int_{a_i}^{b_i} e^{-x}dx \\ &\leq \frac{\gamma_b(\sigma_b^2 + P_J)}{\mu}\int_{\mu}^{a_{m+1}} e^{-x}dx + \frac{\gamma_b(\sigma_b^2 + P_J)}{b_{m+1}}\sum_{i=m}^{n-2}\int_{b_{i+1}}^{a_{i+2}} e^{-x}dx + \frac{\gamma_b(\sigma_b^2 + P_J)}{b_n}\int_{b_n}^{\infty} e^{-x}dx,\end{aligned}$$

and then by the assumption that $b_m \leq \mu < b_{m+1} \leq b_n$, we have:

$$\sum_{i=1}^{m}\int_{a_i}^{b_i} e^{-x}dx \leq \int_{\mu}^{a_{m+1}} e^{-x}dx + \sum_{i=m}^{n-2}\int_{b_{i+1}}^{a_{i+2}} e^{-x}dx + \int_{b_n}^{\infty} e^{-x}dx.$$

Thus,

$$
\begin{aligned}
\sum_{i=1}^{n} \int_{a_i}^{b_i} e^{-x} dx &= \sum_{i=1}^{m-1} \int_{a_i}^{b_i} e^{-x} dx + \int_{\mu}^{b_n} e^{-x} dx - \sum_{i=m}^{n-2} \int_{b_i+1}^{a_{i+2}} e^{-x} dx - \int_{\mu}^{a_{m+1}} e^{-x} dx \\
&\leq \int_{b_n}^{\infty} e^{-x} dx + \int_{\mu}^{b_n} e^{-x} dx \\
&= \int_{\mu}^{\infty} e^{-x} dx,
\end{aligned}
$$

which establishes (2.27).

Note that when $b_n = \infty$ (n must be larger than one since $a_1 < \mu$), we can subtract the integration from a_n to b_n on both sides of (2.21) and then prove that (2.27) is true for both $b_{n-1} \leq \mu$ and $b_{n-1} > \mu$ using the same argument as above.

CHAPTER 3

COVERT COMMUNICATIONS ON A CONTINUOUS-TIME MODEL WITH AN UNINFORMED JAMMER

3.1 Introduction

Security is a major concern in modern wireless communications, where it is often obtained by encryption. However, this is not sufficient in applications where the very existence of transmission arouses suspicion. For example, in military communications, the detection of a transmission may reveal activity in the region. Thus, it is important to study covert communication: a transmitter (Alice) reliably sending messages to a legitimate receiver (Bob) without being detected by an attentive warden (Willie). Previous work studied the limits of reliable covert communications. Bash *et al.* first studied such limits over discrete-time AWGN channels in [5], establishing a square-root law (SRL): Alice can transmit at most $\mathcal{O}(\sqrt{n})$ covert bits to Bob in n channel uses of a discrete-time AWGN channel. This SRL was then established in successive work over binary symmetric channels (BSCs) by Che *et al.* in [13], over discrete memoryless channels (DMCs) by Wang *et al.*, [16] and Bloch, [19], and over multiple-access channels [47] by Arumugam *et al.*. The length of the secret key needed to achieve the SRL in covert communications over DMCs was established in [19]. The work in [16] and [19] also established scaling constants for the covert throughput. These works provide a thorough study of covert communications in common discrete-time channel models when Willie has an accurate statistical characterization of Alice's channel to him.

Recently, covert communication has been further studied in various applications such as the Internet of Things (IoT) [48]- [50], relay and ad-hoc networks [51]- [55], Unmanned Aerial Vehicle (UAV) networks [56, 57], and D2D underlaying cellular networks [58, 59]. In particular, in IoT networks, covert communication technology is emerging as a crucial security technique, since it hides the very existence of the transmission and hence prevents the detection of the IoT users' presence. For example, it can allow patients to privately wear medical devices unknown to others in public places. Hence, the study of the theory and application of covert communication is well-motivated.

In covert communications, Willie attempts to determine whether he is only observing the background environment or a signal from Alice in that environment. Hence, uncertainty about the environment helps Alice to hide her transmission. Lee *et al.*, [23] and Che *et al.*, [14] show that $\mathcal{O}(n)$ covert bits in n channel uses can be reliably transmitted from Alice to Bob if Willie is unsure of the variance of the noise at his receiver. However, Goeckel *et al.*, [37] shows that Willie's lack of knowledge of his noise statistics can be compensated for by estimation through a collection of channel observations when Alice does not transmit. Thus, the limit of covert communications in this case reverts to the SRL. Sobers *et al.*, [9], introduced another model to achieve positive covert rate: introducing an uninformed jammer to the system that randomly generates interference, hence providing the required uncertainty for Willie. The work of [9] also established the optimality for Willie of a power detector in the presence of the uninformed jammer in the discrete-time model. It then shows that Alice can covertly transmit $\mathcal{O}(n)$ bits in n channel uses over both AWGN and block fading channels when Willie uses the optimal detector.

The works mentioned above are all based on a discrete-time model and thus implicitly assume that analogous results can be obtained on the corresponding continuous-time model. Bash *et al.* first mentioned the potential fragility of such an assump-

tion [5]: ideal sinc($\cdot$) pulse shapes are not feasible for implementation, perfect symbol synchronization might not always hold true, and sampling at higher rates sometimes has utility for signal detection at Willie even if the Nyquist intersymbol interference (Nyquist ISI) criterion is satisfied. Considering covert communications in a continuous-time model, Wang in [60, 61] shows that covert channel capacity is positive over an AWGN channel when there is no bandwidth constraint. Wang [62] also shows that the covert channel capacity is infinite over a continuous-time, infinite-bandwidth Poisson channel. Zhang *et al.* in [63] consider a similar problem as in [60], but under a spectral mask constraint. They show that without any jamming, despite the fragilities mentioned in [5], the converted discrete-time signal forms a sufficient statistic for Willie to perform detection, and hence, it suffices to apply standard techniques for discrete-time models to the continuous-time model. In particular, they prove that given a time T and a spectral mask with bandwidth W, Alice can transmit $\mathcal{O}(\sqrt{WT})$ information bits covertly and reliably to Bob.

However, the above works only study the standard Alice-Bob-Willie scenario, and the suitability of a discrete-time model has not been considered in the important case where an uninformed jammer is present to assist Alice. In fact, we will show later that this study leads to very different conclusions than the case without a jammer. In [11], Sobers *et al.* introduced a linear detector for warden Willie that exploits the difference in the timing offsets of Alice's and the jammer's signals. This detector outperforms the standard power detector implemented in the continuous-time system in some limited scenarios. Therefore, a major tenet of [9] that facilitated the establishment of positive rate covert communications in the discrete-time case does not hold in the continuous-time case. Rather, Willie's detection capability benefits from the continuous-time setting, hence raising questions on the covert limits in true continuous-time channels in the case with a jammer. For general scenarios, we introduce an interference cancellation detector in Section 3.3 inspired by co-channel interference cancellation

techniques in cellular networks [70], and show that this detector outperforms the standard power detector implemented in continuous-time covert communication systems. Traditional co-channel interference cancellation requires that the receiver be able to decode the entire stream of data hidden behind the interference. However, in covert communication, Willie only needs to detect the existence of Alice's transmission, i.e., the receiver only needs to resolve a single bit of uncertainty about what is behind the interference, hence suggesting the difficult challenge of achieving positive rate covert communication in such an environment.

In this chapter, we will establish constructions for Alice and an uninformed jammer such that positive covert rate is achievable over continuous-time band-limited AWGN channels. In contrast to what is observed in [63] for the case without a jammer, the continuous-time covert communication problem with a jammer is very different than the discrete-time problem considered in [9]. In particular, the reader will note how the constructions provided here are quite different from those in [9]. In addition, we will also consider the case where there is no frame synchronization between Alice and the jammer, which also extends the work in [9]. The contributions of the work can be summarized as follows:

- *Continuous-Time Covert Communications in the Presence of a Jammer Requires Different Approaches than Discrete-Time:* In contrast to [9], we show that a straightforward implementation of the standard power detector (which is optimal in the discrete-time setting) in a continuous-time system in the presence of a jammer is not the optimal detector at Willie. We introduce an interference cancellation detector for Willie that outperforms the standard power detector.

- *Technique for Alice to Achieve Continuous-Time Covert Communication with Positive Rate in the Presence of a Jammer:* For a continuous-time channel with asymptotic bandwidth W as $T \to \infty$, as defined precisely in Section 3.2, we

establish a construction for Alice and the jammer such that $\mathcal{O}(WT)$ information bits can be transmitted covertly and reliably in T seconds.

- *No Frame Synchronization Between Alice and the Jammer:* The work in [9] assumes perfect frame synchronization between Alice and the jammer in a discrete-time system. We argue that the construction in [9] that achieves $\mathcal{O}(n)$ covert bits in n channel uses may not achieve such if there is no frame synchronization. For a continuous-time system, we extend our construction to show that $\mathcal{O}(WT)$ bits can be transmitted covertly and reliably when there is no frame synchronization.

The rest of this chapter is organized as follows: Section 3.2 provides the system model and metrics. Section 3.3 introduces the interference cancellation detector and presents a performance analysis that shows Willie's advantages in a continuous-time system over a discrete-time system. In Section 3.4, we establish constructions that enable covert communications when there is perfect frame synchronization between Alice and the jammer. In particular, for a continuous-time channel with asymptotic bandwidth W, $\mathcal{O}(WT)$ information bits can be transmitted covertly and reliably in T seconds. Section 3.5 extends the construction by relaxing the requirement of frame synchronization between Alice and the jammer. Finally, Section 3.6 draws the conclusions.

3.2 System Model and Metrics

3.2.1 System Model

Consider the scenario shown in Fig. 3.1 where transmitter Alice ("a") wants to transmit a message to intended recipient Bob ("b") reliably without being detected by a warden Willie ("w"). A jammer ("j") assists the communication by actively sending jamming signals, but without any coordination with Alice.

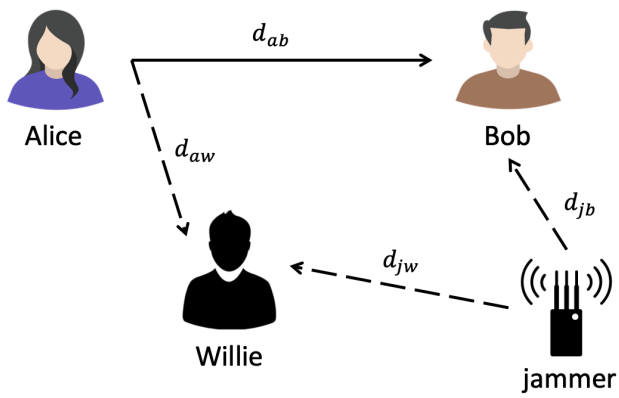


Figure 3.1: System model: With help from a jammer, Alice attempts to transmit reliably and covertly to intended recipient Bob in the presence of a warden Willie.

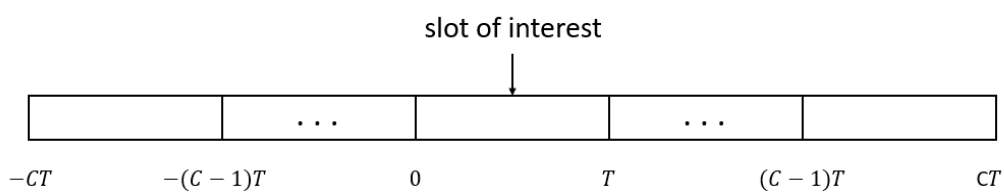


Figure 3.2: Illustration of the time slots, each of length T. Alice may (or may not) transmit in slot $[0, T]$, and Willie attempts to detect her presence in that slot.

For an integer C, we consider a continuous-time channel with $2C$ time slots, each of length T, as shown in Fig. 3.2. We focus on slot $[0, T]$; i.e., Alice may (or may not) transmit in this slot, and assume the jammer transmits in all time slots, and does not know if Alice decides to transmit. If Alice decides to transmit, she will use the slot $[0, T]$. Since any signal restricted in a finite time interval cannot have finite bandwidth, [63] employs a spectral mask that restricts excessive radiation beyond the bandwidth of interest, resulting in an approximate finite bandwidth. For our work, with $T \to \infty$, we use the following definition of asymptotic bandwidth:

Definition 3.1 (Asymptotic Bandwidth)**.** We say that a signal $x(t)$ with $t \in [0, T]$ has an asymptotic bandwidth W in the limit of large T, if $\lim_{T\to\infty} \frac{1}{T} \int_W^\infty E\left[|X(f)|^2\right] df = 0$,

where $X(f)$ is the Fourier transform of $x(t)$, and $X(f)$ depends on T since $x(t)$ is restricted to $[0, T]$.

Alice sends message signal $x_a(t)$ (if she decides to transmit) that is restricted to asymptotic bandwidth W under an average power constraint of P_A. The jammer sends interference signal $x_j(t)$ that is restricted to asymptotic bandwidth W under an average power constraint of P_J. The channels between each transmitter and receiver pair are assumed to be AWGN, and thus the signal observed by Willie is given by:

$$z(t) = \begin{cases} \frac{x_a(t-\tau_a)}{d_{\text{aw}}^{r/2}} + \frac{x_j(t-\tau_j)}{d_{\text{jw}}^{r/2}} + N^{(w)}(t), & \text{Alice transmits and } t \in [0, T] \\ \frac{x_j(t-\tau_j)}{d_{\text{jw}}^{r/2}} + N^{(w)}(t), & \text{else} \end{cases}, \tag{3.1}$$

where d_{xy} is the distance between a transmitter x and a receiver y, r is the path-loss exponent, τ_a and τ_j are timing offsets of Alice's and the jammer's signal, respectively, and $N^{(w)}(t)$ is the noise observed at Willie's receiver, which is a zero-mean stationary Gaussian random process with power spectral density $N_0^{(w)}/2$. Bob observes the channel output $y(t)$ at time t, which is analogous to $z(t)$ but with the substitution of the distance d_{xb} for d_{xw}, where transmitter x is either Alice or the jammer, and the substitution of the noise $N^{(b)}(t)$ for $N^{(w)}(t)$, where $N^{(b)}(t)$ is a zero-mean stationary Gaussian random process with power spectral density $N_0^{(b)}/2$.

We consider two cases: 1) When there is frame synchronization between Alice's and the jammer's signals; and 2) When there is no frame synchronization. In both cases, we assume that the path-loss between Alice and Willie is unknown, but there is an upper and lower bound on the received power at Willie from Alice that is known to the jammer.

3.2.2 Metrics

3.2.2.1 Willie

Based on his observations, Willie attempts to determine whether Alice transmitted or not. We define the null hypothesis (H_0) to be that Alice did not transmit during the time interval and the alternative hypothesis (H_1) to be that Alice transmitted a message. We denote $P(H_0)$ and $P(H_1)$ as the probability that hypothesis H_0 or H_1 is true, respectively. Willie tries to minimize his probability of error $P_e^{(w)} = P(H_0)P_{FA} + P(H_1)P_{MD}$, where P_{FA} and P_{MD} are the probabilities of false alarm and missed detection at Willie, respectively. We assume pessimistically that $P(H_0)$ and $P(H_1)$ are known by Willie. Since $P_e^{(w)} \geq \min(P(H_0), P(H_1))(P_{FA} + P_{MD})$ [5], we say that Alice achieves covert communication if, for a given $\epsilon > 0$, $P_{FA} + P_{MD} \geq 1 - \epsilon$ [5].

We assume that Willie has full knowledge of the statistical model: the time slot $[0, T]$, the parameters for Alice's codebook generation, the parameters for the jammer's interference generation, and the noise power of his channel. Willie does not know the secret key shared between Alice and Bob, or the instantiation of the random jamming.

3.2.2.2 Bob

Bob should be able to reliably decode Alice's message. This is characterized by the probability $1 - P_e^{(b)}$ where $P_e^{(b)}$ is the probability of error at Bob. We say that Alice achieves reliable communication if, for a given $\delta > 0$, $P_e^{(b)} < \delta$ [5].

3.3 Interference Cancellation Detection by Willie

Here we show that in contrast to [9], a straightforward implementation of the standard power detector at Willie is not optimal in the continuous-time model. This means that the achievability result for the covert limit cannot assume a power detector at Willie, and hence, the positive covert rate established in [9] may not hold

true. The work in [70] introduced a co-channel interference cancellation technique that exploits the excess bandwidth in any realizable continuous-time system for initial signal separation when the signals have different timing offsets. We show that similar techniques can be applied by Willie in continuous-time covert communication systems and that these techniques outperform the power detector. Different from traditional co-channel interference cancellation, where the receiver wants to decode the information in a signal that is mixed with another signal, here Willie only needs to detect the existence of Alice's transmission – a single bit decision indicating H_0 or H_1. Without loss of generality, we assume that $d^r_{\text{xy}} = 1$ for all transmitters x (Alice or the jammer) and receivers y (Willie or Bob). However, this assumption is not necessary, and is not required to be known by Willie.

3.3.1 Construction

Here we introduce the natural extension of the work in [9] to continuous-time to demonstrate its limitation. We employ random coding and generate codewords by independently drawing symbols from a zero-mean complex Gaussian distribution with variance σ_a^2 ($\sigma_a^2 \leq P_A$). The codebook is shared between Alice and Bob, and is unknown to Willie. If Alice decides to transmit, she first selects the codeword corresponding to her message, sets f_i to the ith symbol of that codeword ($E[|f_i|^2] = \sigma_a^2$), and transmits the sequence $\{f_i\}_{i=1}^n$, where n is the length of the codeword. The jammer, with knowledge of the slot boundaries but not whether Alice transmits in a given slot (or at all), transmits the zero-mean complex Gaussian symbol sequence $\{v_i^{(\xi)}\}_{i=1}^n$ in time slot $[\xi T, (\xi+1)T]$ with variance $E[|v_i^{(\xi)}|^2] = \sigma_{j,\xi}^2$ ($\sigma_{j,\xi}^2$ varies between slots and $\sigma_{j,\xi}^2 \leq P_J$), where $\xi \in \{-C, -(C-1), \ldots, C-1, \}$.

Over time interval $[0, T]$, Alice (if she decides to transmit) sends her codeword with pulse-shaped waveform:

$$x_a(t) = \sum_{i=1}^{n} f_i p(t - iT_s), \quad 0 \le t \le T$$

where $T_s = T/n$ is the symbol period, and $p(t)$ is a square-root raised cosine (SRRC) pulse shaping filter with bandwidth $(1+\beta)/2T_s < W$, where β is the excess bandwidth or roll-off factor. The jammer sends waveform:

$$x_j(t) = \sum_{\xi=-C}^{C-1} \sum_{i=1}^{n} v_i^{(\xi)} p(t - \xi T - iT_s), \quad -CT \le t \le CT$$

3.3.2 Willie's Receiver

Willie employs a matched filter with impulse response $p(-t)$ at the front-end of his receiver. Define $z_{mf}(t) = z(t) * p(-t)$ as the output of the matched filter, where $*$ denotes convolution. Then, when Alice does not transmit and $t \in [0, T]$, $z_{mf}(t)|H_0$ is given by:

$$z_{mf}(t)|H_0 = \sum_{i=1}^{n} v_i^{(0)} q(t - iT_s - \tau_j) + N^{(w)}(t) * p(-t), \quad 0 \le t \le T \tag{3.2}$$

where $q(t) = p(t) * p(-t)$ is the zero-ISI raised cosine (RC) pulse. When Alice transmits and $t \in [0, T]$, $z_{mf}(t)|H_1$ is given by:

$$z_{mf}(t)|H_1 = \sum_{i=1}^{n} f_i q(t - iT_s - \tau_a) + \sum_{i=1}^{n} v_i^{(0)} q(t - iT_s - \tau_j) + N^{(w)}(t) * p(-t), \quad 0 \le t \le T \tag{3.3}$$

We assume that Willie knows the timing offset τ_j of the jammer's signal, as he can accurately estimate it prior to Alice's transmission [24]. Unlike [11], we do not require that Willie know Alice's timing offset τ_a. Since the power of Alice's signal is much smaller than that of the jammer and Alice might transmit just once, obtaining τ_a could be challenging for Willie; hence, removing the requirement that τ_a is available

at Willie is a significant strength of the converse results of this section. As shown in Fig. 3.3, Willie first samples the signal at $\tau_j + kT_s, k = 1, 2, \ldots, n$ at Branch J. This sampled signal is then used to reconstruct an estimate of the jammer's signal. Next, this estimated interference signal is subtracted from the received signal, and Willie thus obtains an approximation of Alice's signal as the output, which he can sample (above Nyquist rate) and then employ a standard power detector.

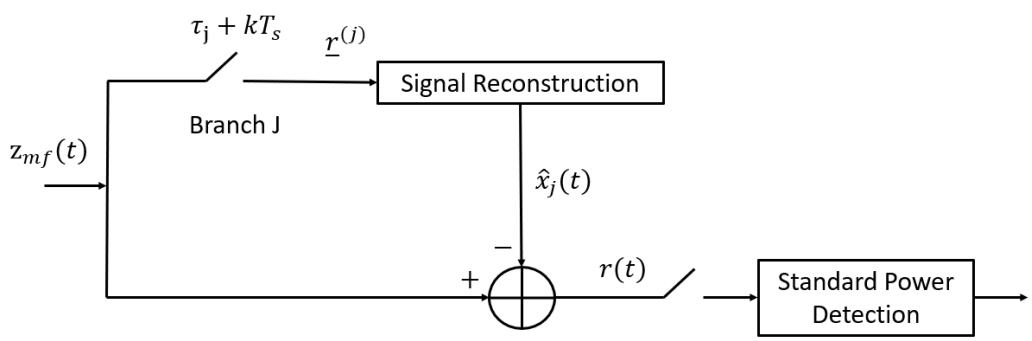


Figure 3.3: Model of interference cancellation at Willie in a covert communication system in the presence of a jammer.

Let $\underline{r}^{(j)}$ denote the vector of samples at Branch J. The kth element of the sample vector $\underline{r}^{(j)}$ is given by:

$$r_k^{(j)} = v_k^{(0)} + \sum_{i=1}^{n} f_i q\left((k-i)T_s + \tau_j - \tau_a\right) + N_k^{(j)}, \quad k = 1, 2, \ldots, n$$

where $N_k^{(j)} = N^{(w)} * p(-kT_s - \tau_j)$ is the sampled noise at Branch J. Willie then reconstructs the jammer's interference signal using pulse shape function $q(t)$:

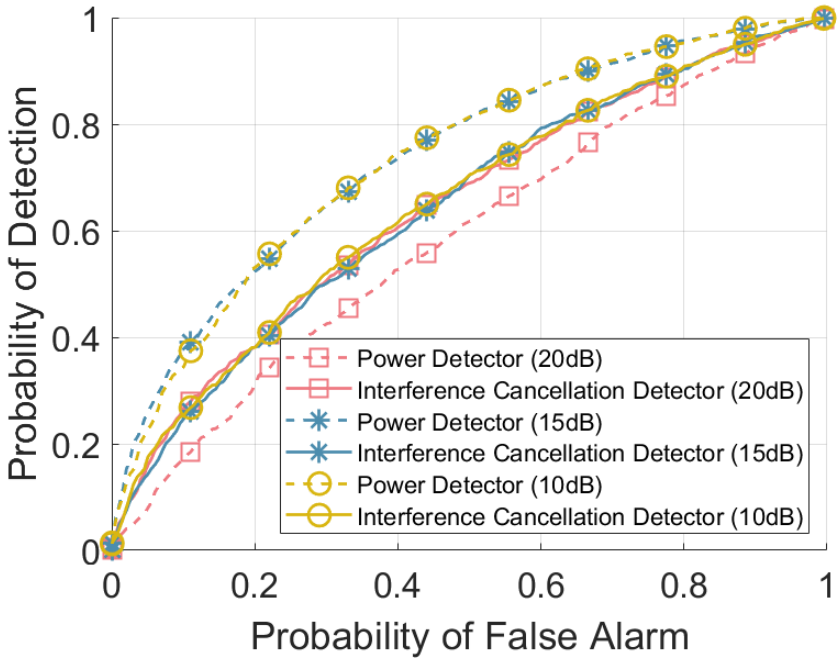
$$\hat{x}_j(t) = \sum_{k=1}^{n} r_k^{(j)} q(t - kT_s - \tau_j)\,. \tag{3.4}$$

Then, $\hat{x}_j(t)$ is subtracted from $z_{mf}(t)$:

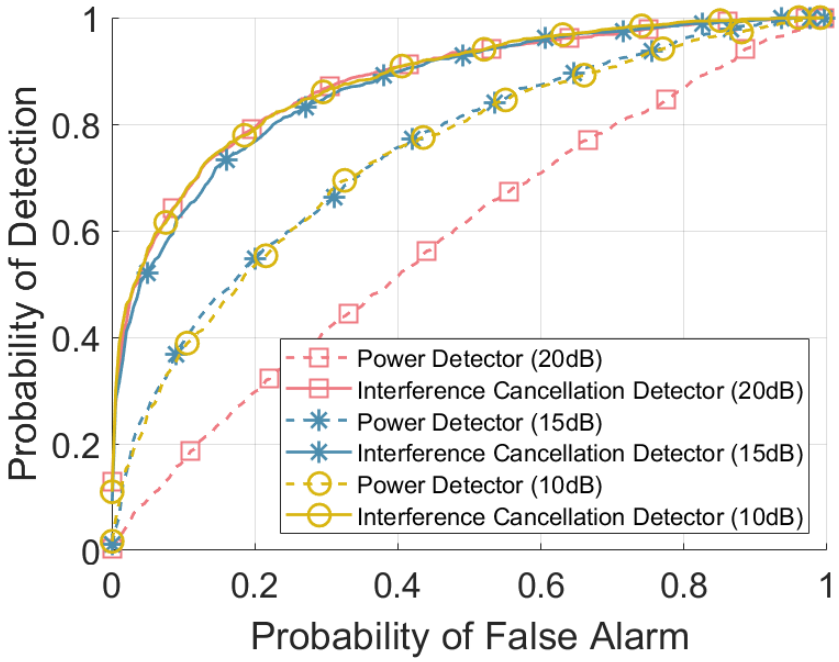
$$r(t) = \sum_{i=1}^{n} f_i q(t - iT_s - \tau_a) - \sum_{k=1}^{n} \sum_{i=1}^{n} f_i q((k-i)T_s + \tau_j - \tau_a) q(t - kT_s - \tau_j)$$
$$+ N^{(w)}(t) * p(-t) - N_k^{(j)} q(t - kT_s - \tau_j) \tag{3.5}$$

which has no components due to the jammer. Rather, since the signal is projected on the null space of the jammer's interference, it only experiences a small reduction in the signal-to-noise ratio (SNR) of Alice's signal due to noise enhancement, as will be observed in the simulation results in the next section. This implies that when Willie employs this interference cancellation detector, adding a jammer will not change the order of the covert throughput as $n \to \infty$. Thus, in contrast to [9], where the authors prove that $\mathcal{O}(n)$ bits in n channel uses can be transmitted reliably and covertly on a discrete-time channel in the presence of a jammer, the natural extension of the techniques in [9] to continuous-time channels is not effective in the presence of interference cancellation detection at Willie. Rather, the covert throughput obeys the SRL proved in [5] for the case without a jammer. Obviously, the same upper bound on covert throughput then holds true for an optimal detector at Willie.

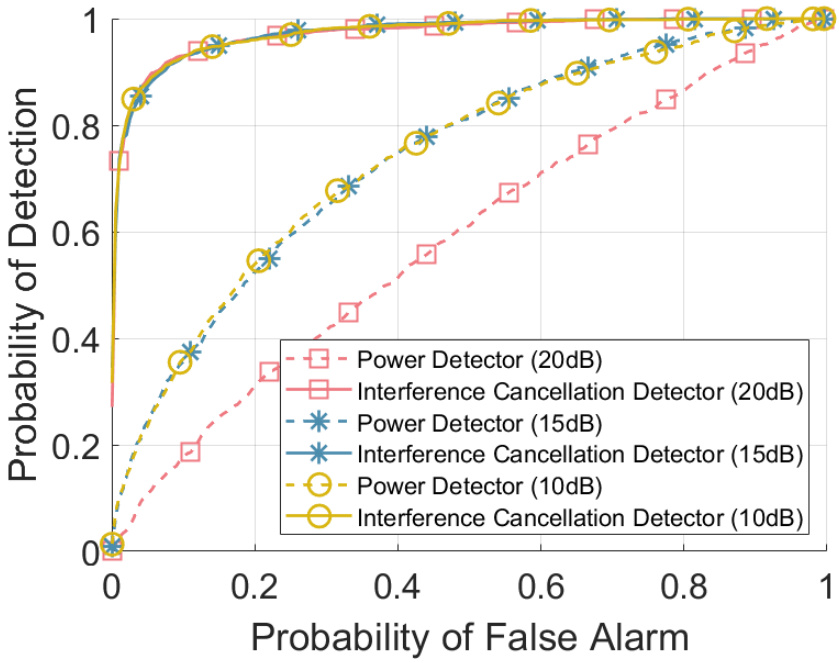
3.3.3 Simulation



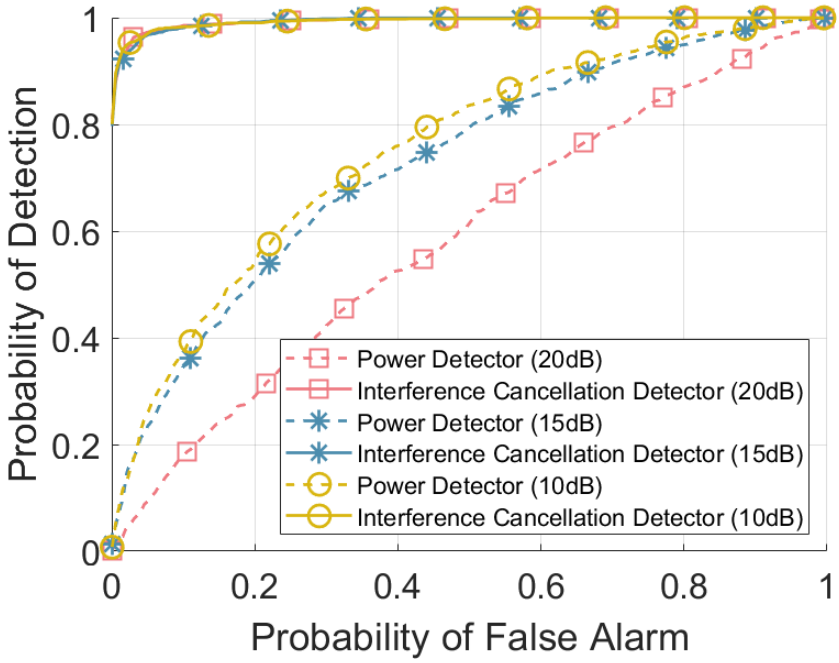
(a) $\tau_a = 4$ samples



(b) $\tau_a = 8$ samples



(c) $\tau_a = 12$ samples



(d) $\tau_a = 16$ samples

Figure 3.4: Receiver operating characteristic of the interference cancellation detector and the standard power detector (implemented in a continuous-time covert communication system) when the jammer's SNR is 20, 15 and 10 dB.

Fig. 3.4 compares the performance of the interference cancellation detector and the standard power detector implemented at Willie in the continuous-time system. In the simulation, we set the number of trials to 2000. For each trial, Alice and the jammer each send 200 independent and identically distributed (i.i.d.) zero-mean Gaussian symbols with pulse-shaped waveforms using a square-root raised cosine pulse shaping filter with roll-off factor 0.2. The two signals have symbol period $T_s = 48$ discrete-time samples. Alice has a timing offset τ_a of $4, 8, 12$ and 16 discrete-time samples, respectively, and the jammer has a timing offset of zero. Alice's SNR is set to 5 dB, and the jammer's SNR is set to 20, 15, and 10 dB. For the interference cancellation detector, a standard power detector is applied after interference cancellation to detect Alice's presence. The standard power detector employs a sample rate of $2/T_s$.

From Fig. 3.4, we observe that the performance of the interference cancellation detector does not change with respect to the jammer's SNR, as expected, since the jammer's signal is canceled. When $\tau_a = 4$, the difference between the timing offsets of Alice's and the jammer's signals $|\tau_a - \tau_j| = T_s/12$ is small. As shown in Fig. 3.4 (a), when the jammer's SNR is relatively low (10 and 15 dB), the interference cancellation detector does not perform as well as the standard power detector implemented in the continuous-time system due to the significant noise enhancement in this case. Note that the noise enhancement is independent of n, so even in this case (small $|\tau_a - \tau_j|$), the interference cancellation detector at Willie would limit the order of the covert throughput that Alice could achieve. When τ_a is large enough ($\tau_a = 8, 10$ or 12 samples), the interference cancellation detector significantly outperforms the standard power detector. This shows that while Alice can reliably transmit $\mathcal{O}(n)$ bits in n channel uses when the standard power detector is employed at Willie, she can only transmit $\mathcal{O}(\sqrt{n})$ bits if Willie employs the interference cancellation detector since the interference cancellation detector completely cancels out the jammer's signal with only a reduction to the SNR that is independent of n. Hence, it is important

for us to analyze covert communications using a continuous-time model to establish achievability results for the covert communications. In the next two sections, we will show that positive rate covert communications can be achieved in continuous-time systems, using a very different construction than that employed in [9].

3.4 Achievable Covert Communications: Perfect Frame Synchronization between Alice and the Jammer

In this section, we provide a construction for Alice and the jammer in the case that they both agree (or know) on the codeword slot timing, as is plausible because Alice could listen before transmitting to obtain such. A challenge to hiding Alice's transmission in the jammer's interference will be that Alice and the jammer will have different and unknown pathloss d^r_{aw} and d^r_{jw}, respectively, to Willie. Hence, without loss of generality, we assume $d^r_{\text{xy}} = 1$ for all transmitter and receiver pairs (x, y), but that d^r_{aw} may not. The construction consists of Alice and the jammer sending randomly located pulses, and hence, Willie will not be able to detect the presence of Alice by exploiting the difference between the timing offsets of Alice and the jammer's signals, as was done in Section 3.3. Also, to thwart Willie detecting the presence of Alice by looking for a pulse power distribution that is the combination of two distributions, we have the jammer send pulses at multiple power levels that cover a wide range of the power spectrum. Thus, if Alice uses an average power resulting in the pulses arriving at Willie with power within that range, she can hide herself in the jammer's interference. We demonstrate that, under this construction, the number of power levels Willie observes is a sufficient statistic for a genie-aided Willie to detect Alice's presence. The ability for Alice to covertly send $\mathcal{O}(WT)$ bits in $[0, T]$ is then established against an optimal genie-aided Willie, which guarantees the achievability against the optimal true Willie. This shows that positive rate covert communications can be achieved in a continuous-time systems when aided by a jammer:

Theorem 3.1. Given the system model in Section 3.2 with frame synchronization between Alice and the jammer, there exists a construction for Alice and the jammer to achieve $\mathcal{O}(WT)$-bit covert and reliable transmission on a continuous-time channel employing asymptotic bandwidth W Hz for T seconds as $T \to \infty$.

3.4.1 Construction

Alice: We employ random coding and generate an i.i.d. Gaussian codebook. If Alice decides to transmit, she sends the codeword corresponding to her message, an i.i.d. complex zero-mean Gaussian symbol sequence $\{f_i\}_{i=1}^{M_n}$ with variance $E[|f_i|^2] = \sigma_a^2$, where $M_n = \lfloor \alpha WT \rfloor$ with a constant $0 < \alpha < 1$ to be chosen later. Alice's transmission power is random: she chooses a power level uniformly at random from $[P_a, P_a + \Delta_{P_a}]$, i.e., $\sigma_a^2 \sim U[P_a, P_a + \Delta_{P_a}]$, where P_a and Δ_{P_a} are constants such that $P_a + \Delta_{P_a} \leq P_A$, and then transmits symbols with this average power σ_a^2 for each pulse she sends in the time interval $[0, T]$. Her waveform within $[0, T]$ is given by:

$$x_a(t) = \sum_{i=1}^{M_n} f_i p(t - \tau_i) \tag{3.6}$$

where $p(t)$ is a unit-energy square-root raised cosine pulse with roll-off factor β and bandwidth W, and $\tau_i, i = 1, 2, \ldots, M_n$ is a sequence of i.i.d. pulse delays that are uniformly distributed in $[0, T]$. Alice's signal $x_a(t)$ with $t \in [0, T]$ has an asymptotic bandwidth W, as shown in Appendix 3.7.1. Since Alice sends M_n pulses, and sends them at any time in the continuous-time interval $[0, T]$, she and Bob share an infinite length key [19] encoding those symbols and time locations unknown to Willie.

Jammer: The jammer transmits an i.i.d. zero-mean complex Gaussian symbol sequence $\{v_i^{(\xi)}\}_{i=1}^{M_n}$ in time slot $[\xi T, (\xi+1)T]$, where $\xi = -C, -(C-1), \ldots, C-1$. For time slot $[\xi T, (\xi + t)T]$, it first selects a number of power levels, K_ξ, according to a Poisson distribution, i.e., $K_\xi \sim Pois\,(\lambda_j)$, where λ_j is a constant to be chosen later. The jammer then chooses each of the K_ξ power levels uniformly in $[P_j, P_j + \Delta_{P_j}]$,

where P_j and Δ_{P_j} are constants and $P_j + \Delta_{P_j} \leq P_J$, to transmit its symbols. Note that the range of the jammer's power received at Willie needs to cover the range of all possible values of Alice's power at Willie. Then, P_j and Δ_{P_j} are chosen such that $\left[\frac{P_a}{d_{\text{aw}}^r}, \frac{P_a+\Delta_{P_a}}{d_{\text{aw}}^r}\right] \subset \left[P_j, P_j + \Delta_{P_j}\right]$. The jammer transmits M_n pulses at each randomly chosen power level in each time slot. Hence, it transmits a total of $K_\xi M_n$ pulses in slot $[\xi T, (\xi+1)T]$, and its waveform is given by:

$$x_j(t) = \sum_{\xi=-C}^{\xi=C-1} \sum_{k=1}^{K_\xi} \sum_{i=1}^{M_n} v_{k,i}^{(\xi)} p(t - \xi T - \tau_{k,i}^{(\xi)}), \quad -CT \leq t \leq CT \tag{3.7}$$

where $\{v_{k,i}^{(\xi)}\}_{i=1}^{M_n}$ is a sequence of i.i.d. zero-mean complex Gaussian symbols with variance being the kth power level randomly chosen by the jammer in slot $[\xi T, (\xi+1)T]$, and $\{\tau_{k,i}^{(\xi)}\}_{i=1}^{M_n}$ is a sequence of M_n i.i.d. pulse delays that are uniformly distributed in $[\xi T, (\xi+1)T]$ for each k. Note that although the above form of $x_j(t)$ is complicated, the outer sum will go away when we focus on the single slot of interest in the next section.

3.4.2 Analysis

Since observations outside of $[0, T]$ do not help Willie detect Alice under our construction, Willie makes his decision on Alice's transmission based on his observation in the slot $[0, T]$. When Alice does not transmit, his received signal $z(t)|H_0$ when $t \in [0, T]$ is given by:

$$z(t) \mid H_0 = \sum_{k=1}^{K_0} \sum_{i=1}^{M_n} v_{k,i}^{(0)} p(t - \tau_{k,i}^{(0)} - \tau_j) + N^{(w)}(t)\,.$$

When Alice transmits, Willie's received signal $z(t) \mid H_1$ for $t \in [0, T]$ is given by:

$$z(t) \mid H_1 = \sum_{i=1}^{M_n} \frac{f_i}{d_{\text{aw}}^{r/2}} p(t - \tau_i - \tau_a) + \sum_{k=1}^{K_0} \sum_{i=1}^{M_n} v_{k,i}^{(0)} p(t - \tau_{k,i}^{(0)} - \tau_j) + N^{(w)}(t)\,.$$

To obtain an achievability result for covert communications between transmitter Alice and intended recipient Bob, as assumed the adversary Willie employs an optimal detector. We will upper bound the performance of the optimal detector by assuming a genie provides Willie additional information; in particular, we provide Willie with knowledge of the exact power range $\left[\frac{P_a}{d_{\text{aw}}^2}, \frac{P_a+\Delta_{P_a}}{d_{\text{aw}}^r}\right]$ that may contain a signal from Alice, the distribution of the number of the jammer's power levels, the locations of the pulses, and values of all power levels employed by the jammer and Alice (if she decides to transmit), but not which power level is employed by whom. We then prove our achievability result against an optimal Willie who possesses this extra information, which guarantees achievability against the optimal Willie under the assumptions of Section 3.2.

3.4.3 Optimal Hypothesis Test

In this section, we show that the number of power levels in the range $\left[\frac{P_a}{d_{\text{aw}}^2}, \frac{P_a+\Delta_{P_a}}{d_{\text{aw}}^r}\right]$, which we term the detection region, is a sufficient statistic for the genie-aided Willie in deciding between hypothesis H_0 or H_1. Fig. 3.5 illustrates the power levels observed by Willie.

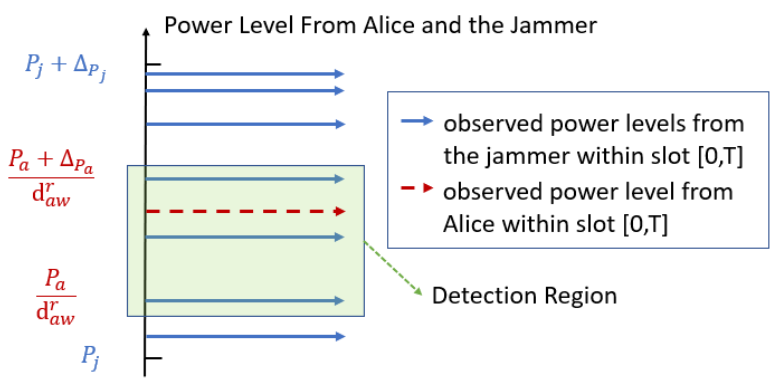


Figure 3.5: Willie's received power levels from Alice and the jammer. An impulse means a power level chosen by either Alice or the jammer within slot $[0, T]$.

Let $K_0^{(1)}$ be the number of power levels inside the detection region, $K_0^{(2)}$ be the number of power levels outside the detection region, i.e. $K_0 = K_0^{(1)} + K_0^{(2)}$. By construction, the power levels sent by the jammer form a Poisson point process on $[P_j, P_j+\Delta_{P_j}]$ with $K_0 \sim Pois\,(\lambda_j)$, under H_0; $K_0-1 \sim Pois\,(\lambda_j)$, under H_1. Note that for a Poisson point process, intervals are independent, and thus generating $K_0^{(1)}+K_0^{(2)}$ power levels with mean λ_j and placing them uniformly over $[P_j, P_j+\Delta_{P_j}]$ is equivalent to generating $K_0^{(1)}$ power levels with mean $\frac{\Delta_{P_a}}{\Delta_{P_j} d_{\text{aw}}^r}\lambda_j$ and placing them uniformly inside the detection region, and generating $K_0^{(2)}$ power levels with mean $\left(1 - \frac{\Delta_{P_a}}{\Delta_{P_j} d_{\text{aw}}^r}\right)\lambda_j$ and placing them uniformly outside the detection region. This is critical in the proof below.

Here we use a Markov chain to prove that the number of power levels in the detection region is a sufficient statistic for Willie to detect Alice's presence. An alternative derivation via the likelihood ratio test (LRT) is provided in Appendix 3.7.2. We denote Alice's decision on transmission as D (which corresponds to hypothesis H_1 when Alice decides to transmit, or H_0 when she decides not to); the locations over $[0, T]$ of pulses as a vector $\mathbf{L}$, i.e., $\mathbf{L}$ is the vector of pulse delays of both the jammer and Alice (if she decides to transmit); the values of all power levels (within and outside the detection region) as a vector $\mathbf{V}$; and the original complex symbols sent as $\mathbf{S}$, i.e., $\mathbf{S}$ consists of the elements in both the vector $\mathbf{v}$ and $\mathbf{f}$ of symbols (if Alice decides to transmit). The random variables D, $K_0^{(1)}$, $\mathbf{V}$, $\mathbf{L}$ and $\mathbf{S}$ form a Markov chain shown in Fig. 3.6, which illustrates the transition from Alice's state D to Willie's received signal $z(t)$ in $[0, T]$. The transitions of the Markov chain are:

- $D \longrightarrow K_0^{(1)}$: $K_0^{(1)}$ and $K_0^{(1)}-1$ are characterized by a Poisson process with mean $\frac{\Delta_{P_a}\lambda_j}{\Delta_{P_j} d_{\text{aw}}^r}$ when Alice does not transmit ($D = H_0$) and when she does transmit ($D = H_1$), respectively.

- $K_0^{(1)} \longrightarrow \mathbf{V}, \mathbf{L}$: Let $\{V_k : k = 1, 2, \ldots, K_0^{(1)}\}$, be the values of power levels within the detection region, and $\{V_k : k = K_0^{(1)} + 1, K_0^{(1)} + 2, \ldots, K_0\}$, be

the values of the power levels outside the detection region. Given $K_0^{(1)}$, the conditional distribution of $V_k, k = 1, 2, \ldots, K_0^{(1)}$, is uniform within the detection region. Note that $K_0^{(2)}$ is independent of D since the pulses sent with power levels outside the detection region can only come from the jammer, no matter if Alice transmits or not. Given $K_0^{(2)}$ (Poisson with mean $\left(1 - \frac{\Delta_{P_a}}{\Delta_{P_j} d_{\text{aw}}^r}\right) \lambda_j$), the distribution of $V_k, k = K_0^{(1)} + 1, K_0^{(1)} + 2, \ldots, K_0$, is uniform outside the detection region. Let $\{L_{k,m} : k = 1, \ldots, K_0^{(1)}, m = 1, \ldots, M_n\}$ denote the locations in $[0, T]$ of pulses sent with power within the detection region, and $\{L_{k,m} : k = K_0^{(1)} + 1, \ldots, K_0, m = 1, \ldots, M_n\}$ denote the locations of pulses sent with power outside the detection region. Given $K_0^{(1)}$, the distribution of $L_{k,m}$ for $k = 1, 2, \ldots, K_0^{(1)}$ and all m is uniform over $[0, T]$. Given $K_0^{(2)}$, the distribution of L_k for $k = K_0^{(1)} + 1, K_0^{(1)} + 2, \ldots, K_0$ and all m is also uniform over $[0, T]$, which is independent from D.

- $\mathbf{V}, \mathbf{L} \longrightarrow \mathbf{S}, \mathbf{L}$: The conditional distribution of $S_{k,m}$, for $k = 1, 2, \ldots, K_0, m = 1, 2, \ldots, M_n$, given V_k, is a zero-mean complex Gaussian random variable with variance $E[|S_{k,m}|^2] = V_k$.

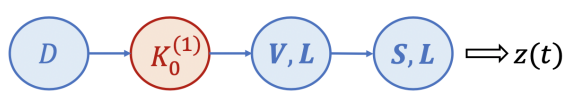


Figure 3.6: Markov chain illustrating the transition from Alice's decision D on transmission, to Willie's observed signal $z(t)$.

Given the pulse locations and the original complex symbols sent, the signal $z(t)$ can be constructed from $p(t)$ and the AWGN of Willie's channel. From the Markov chain shown in Fig. 3.6, we see that $z(t)$ conditioned on $K_0^{(1)}$ is independent of D. Therefore, $K_0^{(1)}$ is a sufficient statistic for Willie to decide between hypotheses H_0 and H_1.

In particular, hypotheses H_0 and H_1 can be stated as:

- H_0: the number of power levels within the detection region follows $Pois\left(\frac{\lambda_j \Delta_{P_a}}{\Delta_{P_j} d^r_{\text{aw}}}\right)$;
- H_1: the number of power levels within the detection region follows $Pois\left(\frac{\lambda_j \Delta_{P_a}}{\Delta_{P_j} d^r_{\text{aw}}}\right) + 1$.

3.4.4 Covertness

Let P_0 and P_1 denote the distributions of the number of power levels observed by Willie within the detection region given H_0 and H_1, respectively:

$$P_0(k) = \frac{\lambda^k e^{-\lambda}}{k!}, \quad k \geq 0 \tag{3.8}$$

and

$$P_1(k) = \frac{\lambda^{k-1} e^{-\lambda}}{(k-1)!}, \quad k \geq 1 \tag{3.9}$$

where $\lambda = \frac{\lambda_j \Delta_{P_a}}{\Delta_{P_j} d^r_{\text{aw}}}$. Theorem 13.1.1 in [85] shows that for the optimal hypothesis test,

$$P_{FA} + P_{MD} = 1 - \mathcal{V}_T(P_0, P_1)$$

where

$$\mathcal{V}_T(P_0, P_1) = \frac{1}{2} \sum_k |P_0(k) - P_1(k)|$$

is the total variation distance between P_0 and P_1, where the sum is over all k in the support of $P_0 \cup P_1$. Therefore, by the definition of covertness, if

$$\mathcal{V}_T(P_0, P_1) \leq \epsilon, \tag{3.10}$$

Alice achieves covert communications. Given (3.8) and (3.9):

$$
\begin{aligned}
\mathcal{V}_T(P_0, P_1) &= \frac{1}{2}\sum_{k=1}^{\infty} |P_0(k) - P_1(k)| + \frac{1}{2}P_0(0) \\
&= \frac{1}{2}\sum_{k=1}^{\infty} \frac{\lambda^{k-1}e^{-\lambda}}{(k-1)!}\left|\frac{\lambda}{k} - 1\right| + \frac{1}{2}e^{-\lambda} \\
&= \frac{1}{2}\left[\sum_{k=1}^{\lfloor\lambda\rfloor} \frac{\lambda^{k-1}e^{-\lambda}}{(k-1)!}\left(\frac{\lambda}{k} - 1\right) + \sum_{k=\lfloor\lambda\rfloor+1}^{\infty} \frac{\lambda^{k-1}e^{-\lambda}}{(k-1)!}\left(1 - \frac{\lambda}{k}\right) + e^{-\lambda}\right] \\
&= \frac{1}{2}\left[\sum_{k=1}^{\lfloor\lambda\rfloor} \frac{\lambda^{k}e^{-\lambda}}{k!} - \sum_{k=1}^{\lfloor\lambda\rfloor} \frac{\lambda^{k-1}e^{-\lambda}}{(k-1)!} + \sum_{k=\lfloor\lambda\rfloor+1}^{\infty} \frac{\lambda^{k-1}e^{-\lambda}}{(k-1)!} - \sum_{k=\lfloor\lambda\rfloor+1}^{\infty} \frac{\lambda^{k}e^{-\lambda}}{k!} + e^{-\lambda}\right] \\
&= \frac{1}{2}\left[\sum_{k=1}^{\lfloor\lambda\rfloor} \frac{\lambda^{k}e^{-\lambda}}{k!} + \sum_{k=\lfloor\lambda\rfloor}^{\infty} \frac{\lambda^{k}e^{-\lambda}}{k!} - \sum_{k=0}^{\lfloor\lambda\rfloor-1} \frac{\lambda^{k}e^{-\lambda}}{k!} - \sum_{k=\lfloor\lambda\rfloor+1}^{\infty} \frac{\lambda^{k}e^{-\lambda}}{k!} + e^{-\lambda}\right] \\
&= \frac{1}{2}\left(\sum_{k=1}^{\infty} \frac{\lambda^{k}e^{-\lambda}}{k!} + \frac{\lambda^{\lfloor\lambda\rfloor}e^{-\lambda}}{\lambda!} - \sum_{k=0}^{\infty} \frac{\lambda^{k}e^{-\lambda}}{k!} + \frac{\lambda^{\lfloor\lambda\rfloor}e^{-\lambda}}{\lambda!} + e^{-\lambda}\right) \\
&= \frac{1}{2}\left(2\frac{\lambda^{\lfloor\lambda\rfloor}e^{-\lambda}}{\lambda!} - e^{-\lambda} + e^{-\lambda}\right) \\
&= \frac{\lambda^{\lfloor\lambda\rfloor}e^{-\lambda}}{\lambda!}
\end{aligned}
$$

Using Stirling's approach, this can be upper bounded as:

$$
\frac{\lambda^{\lfloor\lambda\rfloor}e^{-\lambda}}{\lambda!} \leq \frac{\lambda^{\lambda}e^{-\lambda}}{\sqrt{2\pi}\lambda^{\lambda+1/2}e^{-\lambda}} = \frac{1}{\sqrt{2\pi\lambda}}\,.
$$

Thus, if

$$
\lambda \geq \frac{1}{2\pi\epsilon^2},
$$

i.e.,

$$
\lambda_j \geq \frac{\Delta_{P_j} d^r_{\text{aw}}}{2\pi\Delta_{P_a}\epsilon^2}\,, \tag{3.11}
$$

covertness is achieved. This implies that one of two strategies can be employed: 1) Alice chooses a Δ_{P_a} and the jammer can use an upper bound on d^r_{aw} to choose λ_j; 2)

the jammer chooses a λ_j and Alice can use d_{aw}^r to choose Δ_{P_a}. Since under the above construction, Alice can send $M_n = \lfloor \alpha WT \rfloor$ pulses with a constant power (which does not decrease with WT), $\mathcal{O}(WT)$ bits can be transmitted covertly from Alice to Bob.

3.4.5 Reliability

Since Alice and Bob share a secret key indicating the locations (in $[0, T]$) of the pulses sent by Alice, then Bob can sample at the correct time instances and obtain an M_n-length vector $\mathbf{Y}^{(a)}$ with each element corresponding to a transmitted symbol from Alice, plus noise and interference. The interference is due to pulses sent by the jammer and Alice's intersymbol interference (ISI), since Alice's pulses are sent at random. We upper bound the power of the interference by a constant.

Recall that $x_a(t)$ and $x_j(t)$ are given by (3.6) and (3.7), respectively. Bob employs a matched filter with impulse response $p(-t)$. Let $y_{mf}(t) = \frac{x_a(t)}{d_{\text{ab}}^{r/2}} * p(-t) + \frac{x_j(t)}{d_{\text{jb}}^{r/2}} * p(-t) + N^{(b)}(t) * p(-t)$ be the received signal at Bob for $t \in [0, T]$ after passing through his matched filter:

$$y_{mf}(t) = \sum_{i=1}^{M_n} \frac{f_i}{d_{\text{ab}}^{r/2}} q(t - \tau_i) + \sum_{k=1}^{K} \sum_{i=1}^{M_n} v_{k,i}^{(0)} q(t - \tau_{k,i}^{(0)}) + N^{(b)}(t) * p(-t) \tag{3.12}$$

where $q(t) = p(t) * p(-t)$ is the zero-ISI pulse, e.g., raised cosine pulse. Per above, Bob can sample at $\{\tau_i\}_{i=1}^{M_n}$ and obtain an M_n-length vector $\mathbf{Y}^{(a)}$. For any $m = 1, 2, \ldots, M_n$, the sample attempting to capture symbol f_m is given by:

$$Y_m^{(a)} = f_m + \sum_{i=1, i \neq m}^{M_n} \frac{f_i}{d_{\text{ab}}^{r/2}} q(\tau_m - \tau_i) + \sum_{k=1}^{K} \sum_{i=1}^{M_n} \frac{v_{k,i}^{(0)}}{d_{\text{jb}}^{r/2}} q(\tau_m - \tau_{k,i}^{(0)}) + N^{(b)}(\tau_m) * q(-\tau_m) \,.$$

Thus, the expected power of the interference and noise in one sample is:

$$P_I = E\left[\left(\sum_{i=1,i\neq m}^{M_n} \frac{f_i}{d_{\text{ab}}^{r/2}} q(\tau_m - \tau_i) + \sum_{k=1}^{K}\sum_{i=1}^{M_n} \frac{v_{k,i}^{(0)}}{d_{\text{jb}}^{r/2}} q(\tau_m - \tau_{k,i}^{(0)}) + N^{(b)}(\tau_m) * q(-\tau_m)\right)^2\right]$$
$$= \frac{\sigma_a^2}{d_{\text{ab}}^r} \sum_{i=1,i\neq m}^{M_n} q^2(\tau_m - \tau_i) + \frac{\sigma_j^2}{d_{\text{jb}}^r} \sum_{k=1}^{K}\sum_{i=1}^{M_n} q^2(\tau_m - \tau_{k,i}^{(0)}) + E\left[\left(N^{(b)}(\tau_m) * q(-\tau_m)\right)^2\right] \tag{3.13}$$

where (3.13) is obtained by the fact that $\{f_i\}_{i=1}^{M_n}$ and $\{v_{k,i}^{(0)}\}_{i=1,k=1}^{M_n,K}$ are sequences of i.i.d. zero-mean Gaussian random variables with variance σ_a^2 and σ_j^2, respectively. Let $P_I^{(a)} = \frac{\sigma_a^2}{d_{\text{ab}}^r}\sum_{i=1,i\neq m}^{M_n} q^2(\tau_m - \tau_i)$ be the power of Alice's ISI, $P_I^{(j)} = \frac{\sigma_j^2}{d_{\text{jb}}^r}\sum_{k=1}^{K}\sum_{i=1}^{M_n} q^2(\tau_m - \tau_{k,i}^{(0)})$ be the jammer's interference power, and let $P_I^{(N)} = E\left[\left(N^{(b)}(\tau_m) * q(-\tau_m)\right)^2\right]$ be the noise power. We will evaluate each term, seperately starting with $P_I^{(a)}$.

Let $c_d > 0$ be a small constant, and $M_{d,i}$ be the number of Alice's pulses whose offset τ_i is within a distance of $d_i \in \left((i-1)\frac{c_d}{W}, i\frac{c_d}{W}\right], i = 1, 2, \ldots, \frac{TW}{c_d}$ from τ_m; that is $|\tau_i - \tau_m| \in \left((i-1)\frac{c_d}{W}, i\frac{c_d}{W}\right]$. Then, since the pulses are uniformly located over $[0, T]$, for any $i = 1, 2, \ldots, \frac{TW}{c_d}$,

$$E[M_{d,i}] = \frac{2c_d M_n}{WT} \leq 2\alpha c_d\,. \tag{3.14}$$

Obviously, on each interval $\left((i-1)\frac{c_d}{W}, i\frac{c_d}{W}\right], i = 1, 2, \ldots, \frac{TW}{c_d}$, a pulse located the closest to τ_k maximizes the interference power. Upper bounding $P_I^{(a)}$ and taking the expectation over M_{d_i} yields:

$$E_{M_{d,i}}[P_I^{(a)}] \leq E_{M_{d,i}}\left[\frac{\sigma_a^2}{d_{\text{ab}}^{r/2}} \cdot \sum_{i=1}^{\frac{TW}{c_d}} M_{d,i} \cdot q^2\left((i-1)\cdot\frac{c_d}{W}\right)\right]$$

$$\leq \frac{\sigma_a^2}{d_{\text{ab}}^{r/2}} \cdot \sum_{i=1}^{\frac{TW}{c_d}} E[M_{d,i}] \cdot \frac{C_0}{\left(W^3 \cdot \left(\frac{(i-1)c_d}{W}\right)^3\right)^2} \tag{3.15}$$

$$\leq \frac{2\alpha c_d^{-5}\sigma_a^2 C_0}{d_{\text{ab}}^2} \cdot \sum_{i=1}^{\frac{TW}{c_d}} \frac{1}{i^6} \tag{3.16}$$

where C_0 is a constant, and (3.15) is obtained by the fact that $q(\tau)$ is a raised cosine pulse, which has a tail that decays in the order of $\frac{1}{(W\tau)^3}$ with τ being the timing offset. Since $\sum_{i=0}^{\frac{TW}{c_d}} \frac{1}{i^6}$ converges to a constant as $WT \to \infty$, and $\sigma_a^2 \leq P_A$, (3.16) is upper bounded by a constant.

Similarly, the jammer's interference can be upper bounded by:

$$E_{M_{d,i}}[P_I^{(j)}] \leq \frac{2K\alpha c_d^{-5}\sigma_j^2 C_0}{d_{\text{jb}}^2} \cdot \lim_{WT\to\infty} \sum_{i=0}^{\frac{TW}{c_d}} \frac{1}{i^6}\,. \tag{3.17}$$

Thus, since $\sigma_j^2 \leq P_J$, the power of the jammer's interference is also bounded by a constant. Since $N^{(b)}$ has power spectral density $N_0^{(b)}/2$, and $q(t)$ is a raised cosine pulse with asymptotic bandwidth W and a roll-off factor β, the noise power $P_I^{(N)} = N_0^{(b)}W/(1+\beta)$, which is independent of T.

By [5, Eq.(9)] and Jensen's inequality:

$$E_{M_{d_i}}[P_e^{(b)}] \leq 2^{M_n R - \frac{M_n}{2}\log_2\left(1+\frac{\sigma_a^2}{2\left(P_I^{(N)}+E_{M_{d_i}}[P_I^{(a)}]+E_{M_{d_i}}[P_I^{(j)}]\right)}\right)} \tag{3.18}$$

where R is the rate of Alice's transmission. Since Alice uses a transmission power that is independent of T, her transmission rate is also independent of T. Hence, Bob's error probability is upper bounded by a constant. Therefore, under this construction, Alice achieves both covert and reliable communications with a positive rate.

3.5 Extension to the Case Without Frame Synchronization

In discrete-time systems, frame synchronization between Alice and the jammer is assumed in [9]. It is plausible that Alice could listen to the jammer's signal before transmitting to determine the slot boundaries and then choose an upcoming slot in which to transmit. But since waiting for the upcoming slot incurs delay, and having Alice and Bob agree on the time of transmission makes it challenging for both of them to perform frame synchronization with negligible error, it is desirable to be able to transmit without worrying about such frame synchronization.

In continuous-time systems, the construction provided in the previous section allows Alice to achieve positive rate covert communication when there is perfect frame synchronization between her and the jammer. But the covert rate in the case without frame synchronization also needs to be established. In this section, we extend the construction to relax the requirement that the jammer and Alice both agree (or know) the codeword slot timing. Like before, Alice sends her pulses within time slot $[0, T]$ if she decides to transmit. The jammer transmits in all slots. Without loss of generality, we assume that $d^r_{\text{xy}} = 1$ for all transmitter and receiver pairs (x, y), with the exception that d^r_{aw} is not necessarily equal to one. We will show that positive rate covert communications can be achieved in continuous-time systems with no frame synchronization:

Theorem 3.2. Given the system model of Section 3.2 without frame synchronization between Alice and the jammer, there exists a construction for Alice and the jammer to achieve $\mathcal{O}(WT)$-bit covert and reliable transmission on a continuous-time channel employing asymptotic bandwidth W Hz for T seconds as $T \to \infty$.

3.5.1 Construction

3.5.1.1 Alice

The codebook construction is identical to Section 3.4. Over the time period $[0, T]$, Alice sends M_n pulses (if she decides to transmit). To avoid Willie's detection by simply looking for a pulse starting at time zero, Alice sends pulses during a period of length $T/2$ starting at time T_a drawn uniformly at random from $[0, T/2]$. The starting time of each pulse is uniformly distributed over a $[T_a, T_a + T/2]$. As before, she picks a power level uniformly at random from the power range $(P_a, P_a + \Delta_{P_a})$ to send her pulses. Thus, her waveform is given by:

$$x_a(t) = \sum_{i=1}^{M_n} f_i p(t - \tau_i)$$

where $f_{i\,i=1}^{M_n}$ is the sequence of i.i.d. code symbols, $p(t)$ is a square-root raised cosine pulse with roll-off factor β, $\{\tau_i\}_{i=1}^{M_n}$ is a sequence of i.i.d. pulse delays that are uniformly distributed on $[T_a, T_a + T/2]$.

3.5.1.2 Jammer

As shown in Fig. 3.7, the jammer sends multiple length-$\frac{T}{2}$ pulse trains, each consisting of M_n pulses, having a different power level, and starting at a random point in time. We term these random points the "starting points". The location of each pulse in a pulse train is uniformly distributed over a length-$\frac{T}{2}$ time interval starting from the starting point of that pulse train. Let λ_T denote the density of the starting points per time T. For each length-T period, the number of starting points, which we denote as M_T, follows a Poisson distribution with density λ_T, i.e., $M_T \sim Pois\,(\lambda_T)$. For each starting point, the jammer picks a power level uniformly at random from the power range $(P_j, P_j + \Delta_{P_j})$ to transmit one pulse train. The jammer's waveform due to pulse trains started within slot $[0, T]$ is given by:

$$x_j(t) = \sum_{k=1}^{M_T} \sum_{i=1}^{M_n} v_{k,i} p(t - \tau_{k,i}), \quad 0 \le t \le T$$

where $\{v_{k,i}\}_{i=1}^{M_n}$ is a sequence of i.i.d. zero-mean complex Gaussian symbols with variance being the power level associated with the kth starting point, and $\{\tau_{k,i}\}_{i=1}^{M_n}$ is a sequence of i.i.d. pulse delays that are uniformly distributed in a length-$\frac{T}{2}$ interval starting at the kth starting point.

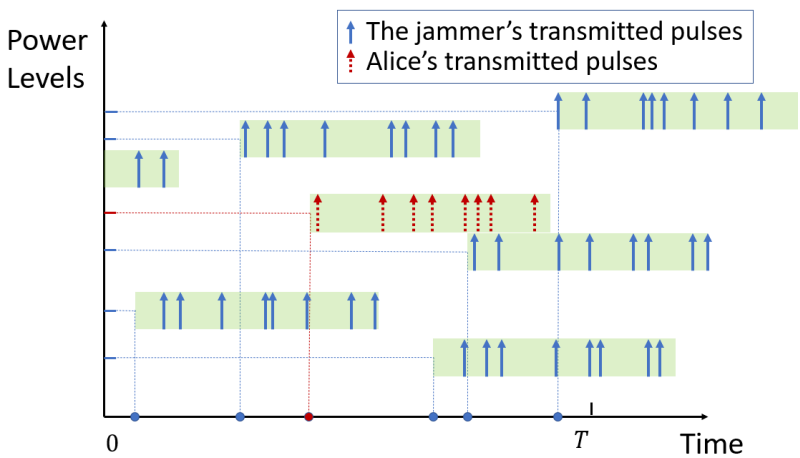


Figure 3.7: Pulse trains sent from Alice and the jammer over $[0, T]$. An impulse means a pulse sent by Alice or the jammer.

3.5.2 Analysis

For achievability, we derive an upper bound on the performance of Willie's optimal detector by assuming a genie provides Willie extra knowledge on the exact power range $\left[\frac{P_a}{d_{\text{aw}}^2}, \frac{P_a + \Delta_{P_a}}{d_{\text{aw}}^r}\right]$ received from Alice, the exact time range $[0, T]$ Alice will employ if she decides to transmit, the values of all power levels employed by the jammer and Alice (if she decides to transmit), the locations of the pulses, the starting points of all the pulse trains corresponding to each power level employed, and the distribution of the time instances. Willie does not know which power level is employed by whom, or for each time instance who starts to send pulses.

Willie makes his decision on Alice's transmission based on his observation over the time interval $[0, T]$. Note that any pulse train started outside $[0, T]$ does not paid Willie's detection, and thus the genie-aided Willie can ignore it. When Alice does not transmit, Willie's received signal due to pulse trains started within $[0, T]$ is given by:

$$z(t) \mid H_0 = \sum_{k=1}^{M_T} \sum_{i=1}^{M_n} v_{k,i} p(t - \tau_{k,i} - \tau_j) + N^{(w)}(t), \quad 0 \leq t \leq T\,.$$

When Alice transmits, Willie's received signal is given by:

$$z(t) \mid H_1 = \sum_{i=1}^{M_n} \frac{f_i}{d_{\text{aw}}^{r/2}} p(t - \tau_i - \tau_a) + \sum_{k=1}^{M_T} \sum_{i=1}^{M_n} v_{k,i} p(t - \tau_{k,i} - \tau_j) + N^{(w)}(t), \quad 0 \leq t \leq T\,.$$

3.5.3 Optimal Hypothesis Test

In this section, we show that the number of time instances over $[0, T]$ that correspond to pulse trains with power levels within the detection region is a sufficient statistic for the genie-aided Willie to decide between H_0 and H_1.

We denote M as the number of starting points in $[0, T]$, and M_1 and M_2 as the number of those points that corresponds to pulse trains with power levels within the detection region (including Alice's pulse train) and outside the detection region, respectively. The proof for M_1 being a sufficient statistic is analogous to the proof for $K_0^{(1)}$ being a sufficient statistic in Section 3.4.3. We provide it in Appendix 3.7.3.

Since M_1 is a sufficient statistic for the genie-aided Willie to detect the presence of Alice, hypotheses H_0 and H_1 can be stated as:

- H_0: M_1 follows $Pois\left(\frac{\Delta_{P_a}}{\Delta_{P_j} d_{\text{aw}}^r} \cdot \lambda_T\right)$;
- H_1: M_1 follows $Pois\left(\frac{\Delta_{P_a}}{\Delta_{P_j} d_{\text{aw}}^r} \cdot \lambda_T\right) + 1$.

3.5.4 Covert Limit

Similar to Section IV, we can derive a lower bound to λ_T:

$$\lambda_T \geq \frac{\Delta_{P_j} d_{\text{aw}}^2}{2\pi \Delta_{P_a} \epsilon^2} \tag{3.19}$$

such that covertness is achieved. As before, the maximum interference at Bob can be upper bounded by a constant, and thus reliability is achieved under the same construction. Also, Alice can send $M_n = \alpha \lfloor WT \rfloor$ pulses with a constant power (which does not decrease with WT), and thus $\mathcal{O}(WT)$ bits can be transmitted covertly and reliably from Alice to Bob.

3.6 Conclusion

In this chapter, we have studied covert communications in continuous-time systems, where transmitter Alice wants to reliably communicate with intended receiver Bob in the presence of a jammer without being detected by warden Willie. We first introduced an interference cancellation detector for Willie that outperforms the standard power detector, hence demonstrating that the continuous-time system will require different approaches than those proposed for discrete-time systems. We then established constructions that allow Alice to achieve covert communications in different cases: when there is perfect frame synchronization between Alice and the jammer, and when there is no frame synchronization. We proved that $\mathcal{O}(WT)$ covert information bits can be reliably transmitted from Alice to Bob on a channel with asymptotic bandwidth W in T seconds for both cases. In this work, an infinite number of key bits shared between Alice and Bob is assumed. A direction for future work is to consider the use of a finite number of key bits and the values of the scaling constants in the performance characterization.

3.7 Appendix

3.7.1 Discussion of the Bandwidth of the Constructions

Here we provide a discussion on the bandwidth of our construction. For the construction in Section 3.4, each of a constant number M_n of pulses with pulse shape $p(t)$ is multiplied by its corresponding symbol and then placed with delay randomly drawn from the interval $[0, T]$. This results in a waveform:

$$x_a(t) = \sum_{i=1}^{M_n} f_i p(t - \tau_i), \quad 0 \leq t \leq T$$

where $\{f_i\}_{i=1}^{M_n}$ is the i.i.d. sequence of zero-mean symbol values, and $\{\tau_i\}_{i=1}^{M_n}$ is the i.i.d. sequence of pulse delays. Since the delays are drawn uniformly over only the interval $[0, T]$, the process $x_a(t)$ is not wide-sense stationary and thus its bandwidth is not strictly defined. Hence, consider rather the following random process, which is an extension of the construction to the infinite interval:

$$\tilde{x}_a(t) = \sum_{k=-\infty}^{\infty} \sum_{i=1}^{M_n} f_i^{(k)} p(t - \tau_i^{(k)} - kT)$$

where $f_i^{(0)} = a_i$ and $\tau_i^{(0)} = \tau_i$, $i = 1, 2, \ldots, M_n$, and the values for the intervals outside of $[0, T]$ are chosen independently but according to the same construction as within $[0, T]$. The random process $\tilde{x}_a(t)$ is wide-sense stationary, and, through standard digital communication system analysis arguments, has power spectral density $S_{\tilde{x}_a}(f) = |P(f)|^2$, where $P(f)$ is the Fourier transform of $p(t)$. Hence, the bandwidth of $\tilde{x}_a(t)$ is the same as that of $P(f)$. Suppose $P(f)$ has a bandwidth of $W - \Delta_W$, where Δ_W is a very small constant, then $\tilde{x}_a(t)$ has a bandwidth of $W - \Delta_W$. Observing that $x_a(t)$ is a windowed version of $\tilde{x}_a(t)$:

$$x_a(t) = \text{rect}\left(\frac{t - T/2}{T}\right) \cdot \tilde{x}_a(t) \tag{3.20}$$

Taking the Fourier transform:

$$X_a(f) = e^{-j\pi ft} T\text{sinc}(fT) * \tilde{X}_a(f)\,. \tag{3.21}$$

In the limit of large T, we have:

$$\begin{aligned}
&\lim_{T\to\infty} \frac{1}{T}\int_W^\infty E\left[|X_a(f)|^2\right] df \\
&= \lim_{T\to\infty} \frac{1}{T}\int_W^\infty E\left[\left|e^{-j\pi ft}\int_{-W+\Delta_W}^{W-\Delta_W} T\text{sinc}(T(f-v))\cdot \tilde{X}_a(v)dv\right|^2\right] df \\
&\le \lim_{T\to\infty} T\int_W^\infty E\left[\left(\int_{-W+\Delta_W}^{W-\Delta_W} \frac{1}{T(f-W+\Delta_W)}\cdot |\tilde{X}_a(v)|dv\right)^2\right] df \\
&= \lim_{T\to\infty} T\int_W^\infty \frac{1}{T^2(f-W+\Delta_W)^2} E\left[\left(\int_{-W+\Delta_W}^{W-\Delta_W} |\tilde{X}_a(v)|dv\right)^2\right] df \\
&= \lim_{T\to\infty} \frac{1}{T}\int_W^\infty \frac{1}{(f-W+\Delta_W)^2} E\left[\int_{-W+\Delta_W}^{W-\Delta_W}\int_{-W+\Delta_W}^{W-\Delta_W} |\tilde{X}_a(v)|\cdot|\tilde{X}_a(u)|dvdu\right] df \\
&\le \lim_{T\to\infty} \frac{1}{T}\int_W^\infty \frac{1}{(f-W+\Delta_W)^2} E\left[\int_{-W+\Delta_W}^{W-\Delta_W}\int_{-W+\Delta_W}^{W-\Delta_W} \frac{1}{2}\left(|\tilde{X}_a(v)|^2 + |\tilde{X}_a(u)|^2\right) dvdu\right] df \\
&\le \lim_{T\to\infty} \frac{1}{T}\int_W^\infty \frac{1}{(f-W+\Delta_W)^2} E\left[\int_{-W+\Delta_W}^{W-\Delta_W} |\tilde{X}_a(v)|^2 dv\right] df \\
&\le \lim_{T\to\infty} \frac{1}{T}\int_W^\infty \frac{1}{(f-W+\Delta_W)^2} E\left[\int_{-W+\Delta_W}^{W-\Delta_W} |P(v)|^2 dv\right] df \qquad (3.22) \\
&= \lim_{T\to\infty} \frac{1}{T}\int_W^\infty \frac{1}{(f-W+\Delta_W)^2}\,(2W-2\Delta_W)\,df \\
&\le \lim_{T\to\infty} \frac{2W-2\Delta_W}{T\Delta_W} \qquad (3.23) \\
&= 0
\end{aligned}$$

where (3.22) is obtained by noting that $\int_{-W+\Delta_W}^{W-\Delta_W} |\tilde{X}_a(v)|^2 dv$ is the power of $\tilde{x}_a(t)$ and $\tilde{x}_a(t)$ has power spectral density $S_{\tilde{x}_a}(f) = |P(f)|^2$, and (3.23) is obtained by noting $p(t)$ is a unit-energy pulse.

3.7.2 Proof of Sufficient Statistic at Genie-Aided Willie Using the LRT

Recall that $\mathbf{L}$ denotes the locations of the pulses over $[0, T]$, $\mathbf{V}$ denotes the values of all power levels, and $\mathbf{S}$ denotes the original complex symbols sent by Alice (if she transmits) and the jammer. For $i = 1, 2, \ldots, K_0^{(1)}$ and all m, the variables $L_{i,m}, V_i$ and $S_{i,m}$ are associated with pulses sent with power levels within the detection region, and for $i = K_0^{(1)} + 1, K_0^{(1)} + 2, \ldots, K_0$ and all m, the variables $L_{i,m}, V_i$ and $S_{i,m}$ are associated with pulses sent with power levels outside the detection region. The LRT is given by:

$$
\begin{aligned}
\Lambda(x, l, v, k) =& \frac{P_{z(t),\mathbf{L},\mathbf{V},K_0|H_1}(x, l, v, k)}{P_{z(t),\mathbf{L},\mathbf{V},K_0|H_0}(x, l, v, k)} \\
=& \frac{P_{\mathbf{S},\mathbf{L},\mathbf{V},K_0|H_1}(s, l, v, k)}{P_{\mathbf{S},\mathbf{L},\mathbf{V},K_0|H_0}(s, l, v, k)} \qquad (3.24) \\
=& \frac{P_{\mathbf{S},\mathbf{V},\mathbf{L}|H_1,K_0}(s, l, v) \cdot P_{K_0|H_1}(k)}{P_{\mathbf{S},\mathbf{V},\mathbf{L}|H_0,K_0}(s, l, v) \cdot P_{K_0|H_0}(k)} \\
=& \frac{\prod_{i=1}^{K_0^{(1)}} \prod_{m=1}^{M_n} P_{S_{i,m},V_i|H_1,L_{i,m}}(s) \cdot P_{L_{i,m}|H_1}(l) \cdot P_{K_0^{(1)}|H_1}(k)}{\prod_{i=1}^{K_0^{(1)}} \prod_{m=1}^{M_n} P_{S_{i,m},V_i|H_0,L_{i,m}}(s) \cdot P_{L_{i,m}|H_0}(l) \cdot P_{K_0^{(1)}|H_0}(k)} \\
& \cdot \frac{\prod_{i=K_0^{(1)}+1}^{K_0} \prod_{m=1}^{M_n} P_{S_{i,m},V_i|H_1,L_{i,m}}(s) \cdot P_{L_{i,m}|H_1}(l) \cdot P_{K_0^{(2)}|H_1}(k)}{\prod_{i=K_0^{(1)}+1}^{K_0} \prod_{m=1}^{M_n} P_{S_{i,m},V_i|H_0,L_{i,m}}(s) \cdot P_{L_{i,m}|H_0}(l) \cdot P_{K_0^{(2)}|H_0}(k)} \\
=& \frac{\prod_{i=1}^{K_0^{(1)}} \prod_{m=1}^{M_n} P_{S_{i,m}|H_1,V_i}(s) \cdot P_{V_i|H_1}(v) \cdot P_{L_{i,m}|H_1}(l) \cdot P_{K_0^{(1)}|H_1}(k)}{\prod_{i=1}^{K_0^{(1)}} \prod_{m=1}^{M_n} P_{S_{i,m}|H_0,V_i}(s) \cdot P_{V_i|H_0}(v) \cdot P_{L_{i,m}|H_0}(l) \cdot P_{K_0^{(1)}|H_0}(k)} \\
& \cdot \frac{\prod_{i=K_0^{(1)}+1}^{K_0} \prod_{m=1}^{M_n} P_{S_{i,m}|H_1,V_i}(s) \cdot P_{V_i|H_1}(v) \cdot P_{L_{i,m}|H_1}(l) \cdot P_{K_0^{(2)}|H_1}(k)}{\prod_{i=K_0^{(1)}+1}^{K_0} \prod_{m=1}^{M_n} P_{S_{i,m}|H_0,V_i}(s) \cdot P_{V_i|H_0}(v) \cdot P_{L_{i,m}|H_0}(l) \cdot P_{K_0^{(2)}|H_0}(k)} \\
=& \frac{P_{K_0^{(1)}|H_1}(k)}{P_{K_0^{(1)}|H_0}(k)} \qquad (3.25)
\end{aligned}
$$

where (3.24) is obtained by the fact that $z(t)$ can be constructed from $p(t)$ given the locations $\mathbf{L}$ and the original complex symbols $\mathbf{S}$ sent, and (3.25) by recognizing: 1) $P_{S_{i,m}|H_1,V_i}(s)$ and $P_{S_{i,m}|H_0,V_i}(s)$ are both zero-mean Gaussian distributions with variance V_i for i and m; 2) $P_{V_i|H_1}(v)$ and $P_{V_i|H_0}(v)$ are both uniform distributions

over the detection region; 3) $P_{L_{i,m}|H_1}(l)$ and $P_{L_{i,m}|H_0}(l)$ are both uniform distributions over a length-T interval; and 4) $P_{K_0^{(2)}|H_1}$ and $P_{K_0^{(2)}|H_0}$ are both Poisson distributions with parameter $\lambda_j \left(1 - \frac{\Delta_{P_a}}{\Delta_{P_j} d_{\text{aw}}^r}\right)$.

3.7.3 Proof of M_1 Being a Sufficient Statistic for the Genie-Aided Willie

Recall that D denotes Alice's decision on transmission. Let L' denote the locations of all of the pulses over $[0, T]$, $\mathbf{S}'$ denote the original complex symbols sent, and $\mathbf{V}'$ denote the values of all power levels (within and outside the detection region) employed during $[0, T]$. The random variables D, M_1, $\mathbf{V}'$, $\mathbf{L}'$ and $\mathbf{S}'$ form a Markov chain shown in Fig. 3.8, which illustrates the transition from Alice's state D to Willie's received signal $z(t)$ in $[0, T]$. The transitions of the Markov chain are:

- $D \longrightarrow M_1$: M_1 and $M_1 - 1$ are characterized by a Poisson process with mean $\frac{\Delta_{P_a}}{\Delta_{P_j} d_{\text{aw}}^r} \cdot \lambda_T$ when Alice does not transmit ($D = H_0$) and when she does transmit ($D = H_1$), respectively.

- $M_1 \longrightarrow \mathbf{V}', \mathbf{L}$': Let $\{V_k' : k = 1, 2, \ldots, M_1\}$, be the values of power levels within the detection region, and $\{V_k' : k = M_1 + 1, M_1 + 2, \ldots, M\}$, be the values of power levels outside the detection region. Given M_1, the conditional distribution of $V_k', k = 1, 2, \ldots, M_1$, is uniform within the detection region. Note that M_2 is independent of D since the pulses sent with power levels outside the detection region can only come from the jammer, no matter if Alice transmits or not. Given M_2 (Poisson with mean $\left(1 - \frac{\Delta_{P_a}}{\Delta_{P_j} d_{\text{aw}}^r}\right) \lambda_T$), the distribution of $V_k', k = M_1 + 1, M_1 + 2, \ldots, M$, is uniform outside the detection region. Let $\{L_{k,m}' : k = 1, \ldots, M_1, m = 1, \ldots, M_n\}$ denote the locations (in $[0, T]$) of pulses sent with power within the detection region, and $\{L_{k,m}' : k = M_1+1, \ldots, M, m = 1, \ldots, M_n\}$ denote the locations of pulses sent with power outside the detection region. Given M_1 and the time instances each pulse train starts, the distribution of $L_{k,m}'$ for $k = 1, 2, \ldots, M_1$ and all m is uniform over the length-$\frac{T}{2}$ interval

starting from the time instances corresponding to each power level. Given M_2, the distribution of L_k for $k = M_1 + 1, M_1 + 2, \ldots, M$ and all m is also uniform over the length-$\frac{T}{2}$ interval starting from the time instances corresponding to each power level, which is independent from D.

- $\mathbf{V}', \mathbf{L}' \longrightarrow \mathbf{S}', \mathbf{L}'$: The conditional distribution of $S'_{k,m}$, $k = 1, 2, \ldots, M, m = 1, 2, \ldots, M_n$, given V'_k, is a zero-mean complex Gaussian random variable with variance $E[|S'_{k,m}|^2] = V'_k$.

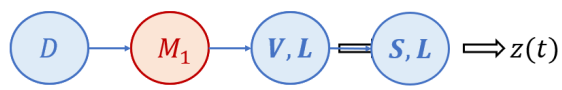


Figure 3.8: Markov chain illustrating the transition from Alice's decision D on transmission, to Willie's observed signal $z(t)$.

Given the pulse locations and the original complex symbols sent over $[0, T]$, the signal $z(t)$ can be constructed from $p(t)$ and the AWGN of Willie's channel. From the Markov chain shown in Fig. 3.8, we see that $z(t)$ conditioned on M_1 is independent of D. Therefore, M_1 is a sufficient statistic for Willie to decide between hypotheses H_0 and H_1.

CHAPTER 4

COVERT COMMUNICATIONS UNDER THE COVER OF A RADAR

4.1 Introduction

Covert communication allows a transmitter (Alice) to reliably send messages to a legitimate receiver (Bob) without being detected by an attentive warden (Willie). This is crucial for many applications where the existence of a transmission reveals sensitive information. Previous work studied the limits of reliable covert communications. Bash *et al.* first studied such limits over discrete-time AWGN channels in [5], establishing a square-root law (SRL): Alice can reliably and covertly transmit at most $\mathcal{O}(\sqrt{n})$ bits to Bob in n channel uses of a discrete-time AWGN channel. This SRL was then established in successive work over binary symmetric channels (BSCs) by Che *et al.* in [13], over discrete memoryless channels (DMCs) by Wang *et al.*, [16] and Bloch, [19], and over multiple-access channels [47] by Arumugam *et al.*. These works provide a thorough study of covert communications in common discrete-time channel models when Willie has an accurate statistical characterization of Alice's channel to him.

In covert communications, Willie attempts to determine whether he is only observing the background environment or a signal from Alice in that environment. Hence, uncertainty about the environment helps Alice to hide her transmission. Sobers *at al.*, [9] achieve positive rate by introducing a model with an interference source: an uninformed jammer that randomly generates interference, hence providing the required uncertainty at Willie. It is shown in [9] that Alice can covertly transmit $\mathcal{O}(n)$ bits in n channel uses over both discrete-time AWGN and block fading channels.

In many scenarios, the jammer is assumed to be an intentional jammer that help Alice to achieve covert communications on purpose [33]. However, this requires an active non-covert Alice-Bob teammate, which may be difficult to provide in some situations. For example, in military communications where a transmitter attempts to covertly communicate with a receiver in enemy territory, it may be difficult to provide such a jammer. In addition, the jammer in [9] is itself not covert since the warden knows that the jammer is present and potentially trying to hide something. Hence, it is often useful to exploit an interference source that already exists in the environment, such as radars [34,35] and other existing communication sources [36,37], so that Alice and Bob can operate in such environment and hide under the cover of the interference. This chapter is the first work that analyzes the fundamental covert rate of a system that uses an unintentional interference source.

The work in [34] and [35] by Blunt *et al.* introduces an intra-pulse radar-embedded communication system where the transmitter attempts to covertly send transmission symbols to the radar. In order to achieve covertness, the transmission symbols are embedded with the incident radar pulses, and are hidden behind the backscatter induced by the radar reflections. The design of intra-pulse covert symbols based on the incident radar waveform is studied in [34] such that the covert symbols are sufficiently different from the ambient radar scattering to ensure acceptable bit error rate (BER) but at the same time sufficiently similar to the scattering to avoid detection by any adversary.

In this chapter, we exploit the idea of embedding covert symbols into radar pulses in a standard Alice-Bob-Willie covert communication scenario. When Alice decides to transmit, each incident radar pulse is remodulated into one communication waveform. Alice embeds her transmission waveforms into the scattering of the radar. We establish the theoretical limit on the covert rate of the transmission between Alice and Bob. In particular, we show that $\mathcal{O}(n)$ bits can be transmitted covertly regardless of

Willie's receiver in n samples of a radar pulse. We analyze Bob's decoding capability and provide an upper bound to his probability of error, which also shows that $\mathcal{O}(n)$ covert bits in n samples of a radar pulse can be transmitted reliably to Bob. A difference between this work and [9], where the interference comes from a jammer, is that here the interference from the radar scattering is correlated, and the variance of the average power of the local scattering is random. This leads to significant differences in the proofs of the covert limit and the error probability of Bob's decoding.

The rest of this chapter is organized as follows: after introducing the system model and metric in Section 4.2, we discuss the design of the covert communication waveform in Section 4.3. In Section 4.4, we prove the covertness of the system and analyze the covert rate. We then derive the error probability at Bob to determine reliability in Section 4.5. Finally, Section 4.6 draws the conclusions.

4.2 System Model and Metrics

4.2.1 System Model

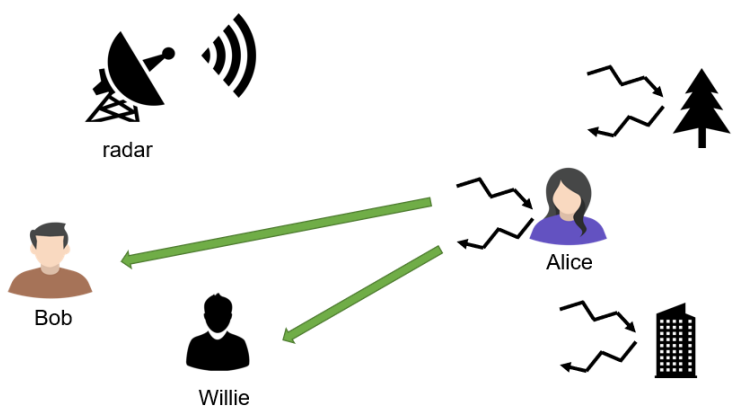


Figure 4.1: System model: Alice attempts to transmit covertly to Bob in a communication system with an illuminating radar.

Consider a radar-embedded covert communication system shown in Fig. 4.1, where the radar illuminates and introduces scattering due to the reflectors in the environ-

ment. The transmitter Alice attempts to covertly transmit messages to the receiver Bob in this environment. In order to do so, she embeds her transmission into the scattering of the radar to avoid being detected by the warden Willie.

Since each incident radar pulse is remodulated into one communication waveform, we are interested in Alice's ability to covertly transmit in a pulse with length equal to that of the codeword length n and Willie's ability to detect such a transmission. We assume that the environment (i.e, clutter) is dynamic such that it changes over different pulse repetition intervals (PRI) but is static within each PRI. Let $\mathbf{s} = [s_1, s_2, \ldots, s_n]^T$ be the sequence of samples of the illuminating radar waveform with average power σ_s^2, where $(\cdot)^T$ denotes the transpose operation. Since the ambient scattering can be assumed to be a linear time-invariant process, the received signal due to the clutter can be expressed as a convolution of the radar signal and the aggregation of the local scattering at a given range delay with respect to the receiver. Hence, the ambient radar scattering in the surrounding area can be modeled as $\mathbf{Sx}$ where $\mathbf{x}$ is a vector of random complex scattering coefficients, and:

$$\mathbf{S} = \begin{bmatrix} s_n & s_{n-1} & \ldots & s_1 & 0 & \ldots & 0 \\ 0 & s_n & & s_2 & s_1 & & 0 \\ \vdots & & \ddots & \vdots & \vdots & \ddots & \\ 0 & 0 & & s_n & s_{n-1} & \ldots & s_1 \end{bmatrix}$$

is an n-by-$(2n-1)$ matrix of delay shifts of $\mathbf{s}$ that characterizes the convolution of the radar signal with the local scattering, i.e., $\mathbf{Sx} = (\mathbf{s} * \mathbf{x})(n)$. Due to the scattering being a collection of the reflections from a large number of scatterers, we assume that $\mathbf{x}$ is a zero-mean independent and identically distributed (i.i.d.) complex Gaussian random vector with variance σ_x^2. In addition, since the environment (i.e., clutter) changes over different PRIs, we assume σ_x^2 is random and has a probability density function (pdf) of $f_{\sigma_x^2}$.

Without loss of generality, we assume that the distance between any transmitter and receiver is one. In the case of AWGN channels between all transmitters and receivers, the received vector at Willie when Alice does not transmit is given by:

$$\mathbf{z} = \mathbf{S}\mathbf{x} + \mathbf{u} \tag{4.1}$$

where $\mathbf{u}$ is the length-n noise vector observed at Willie, which is an i.i.d. zero-mean complex Gaussian vector with variance σ_u^2, i.e., $u_i \sim \mathcal{CN}(0, \sigma_u^2)$, $i = 1, 2, \ldots, n$. When Alice transmits, the received vector at Willie is given by:

$$\mathbf{z} = \mathbf{S}\mathbf{x} + \sigma_c \mathbf{c}_k + \mathbf{u} \tag{4.2}$$

where $\mathbf{c}_k$ is the length-n codeword Alice transmits, and σ_c is the square root of Alice's transmission power.

Bob observes the channel output $\mathbf{y}$, which is analogous to $\mathbf{z}$ but with the substitution of the noise $\mathbf{u}'$ for $\mathbf{u}$, where $\mathbf{u}'$ is an i.i.d. zero-mean complex Gaussian vector with variance $\sigma_{u'}^2$, and the substitution of the radar scattering $\mathbf{x}'$ for $\mathbf{x}$, where $\mathbf{x}'$ is a zero-mean random vector with variance $\sigma_{x'}^2$, and $\sigma_{x'}^2$ is a random variable with a pdf of $f_{\sigma_x^2}$.

4.2.2 Metric

Based on his observation $z(t)$, Willie will determine whether Alice transmitted or not. We define the null hypothesis (H_0) to be that Alice did not transmit, and the alternative hypothesis (H_1) to be that Alice transmitted a message. We assume Alice transmits a message with probability p. Willie tries to minimize his probability of error $\mathbb{P}_e^{(w)} = (1-p) \cdot \mathbb{P}_{FA} + p \cdot \mathbb{P}_{MD}$, where $\mathbb{P}_{FA}$ and $\mathbb{P}_{MD}$ are the probabilities of false alarm and missed detection at Willie, respectively. We assume that $\mathbb{P}(H_0)$ and $\mathbb{P}(H_1)$ are known to Willie. Since $\mathbb{P}_e^{(w)} \geq \min(\mathbb{P}(H_0), \mathbb{P}(H_1))(\mathbb{P}_{FA} + \mathbb{P}_{MD})$ [5], we say that

Alice achieves covert communication if, for a given $\epsilon > 0$, $\mathbb{P}_{FA} + \mathbb{P}_{MD} \geq 1 - \epsilon$ [5]. Reliability is measured by the probability $\mathbb{P}_e^{(b)}$ of Bob's decoding error averaged over all codebooks. The transmission is reliable if, for a given $\epsilon' > 0$, $\mathbb{P}_e^{(b)} < \epsilon'$.

We assume that Willie has full knowledge of the statistical model: The radar signal $\mathbf{s}$, his background noise power, i.e., σ_u^2, and the distribution $f_{\sigma_x^2}$ of the average power of the random local scattering of the radar. However, Willie does not know the exact coefficients $\mathbf{x}$.

4.3 Communication Waveform against Detection

4.3.1 Waveform Design

For a rate of R, we employ random coding arguments and independently generate $K = 2^{nR}$ codewords $\{\mathbf{c}_k, k = 1, 2, \ldots, 2^{nR}\}$ as described in detail below. If Alice decides to transmit, she selects the codeword corresponding to her message, scales it with σ_c to embed it into one radar pulse, and sends it over the AWGN channel. Unlike in [9] where Alice's codeword is independent of the jammer's signal, the codeword $\mathbf{c}_k$ in our scenario should possess some correlation with the radar scattering to avoid detection by simply being projected away from the radar signal. A basis within which to generate Alice's communication codeword can be obtained by the eigen-decomposition of the power-normalized correlation matrix of $\mathbf{Sx}$:

$$\frac{1}{\sigma_x^2} E\left[(\mathbf{Sx})(\mathbf{Sx})^H\right] = \mathbf{S}\mathbf{S}^H = \mathbf{V}\boldsymbol{\Lambda}\mathbf{V}^{\mathbf{H}} \tag{4.3}$$

where $\mathbf{V} = [\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n]$ is the set of n eigenvectors, Λ is a diagonal matrix that contains the associated eigenvalues $\{\lambda_i\}_{i=1}^n$ in descending order, i.e., $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$, and $(\cdot)^H$ denotes the Hermitian operation. Note that $\mathbf{V}\boldsymbol{\Lambda}\mathbf{V}^H$ does not depend on σ_x^2 in the left-hand side of (4.3) since it is cancelled out by σ_x^2 that comes out of the expectation.

The work in [34] introduces three design strategies for the communication waveform in an intra-pulse radar-embedded system: 1) use a subset of the nondominant eigenvectors (i.e., eigenvectors associated with small eigenvalues) as communication waveforms; 2) use a weighted combination of multiple nondominant eigenvectors; 3) project onto the non-dominant subspace. The general idea behind all three designs is to let the covert symbols be sufficiently different from the radar scattering to ensure acceptable BER at the receiver but at the same time similar enough to avoid detection by the adversary.

In our work, we will use both the dominant and non-dominant subspace of the radar scattering. Let $\mathbf{b}_k$, $k = 1, 2, \ldots, K$, be a normal random vector. We define Alice's codeword as:

$$\mathbf{c}_k = \mathbf{V}(\mathbf{\Lambda} + \mathbf{I})^{\frac{1}{2}}\mathbf{b}_k \tag{4.4}$$

where $\mathbf{I}$ is the identity matrix. Note that $\mathbf{b}_k$ is the secret key shared only between Alice and Bob but not Willie.

We want the dimension of the dominant subspace of the radar scattering to be of the same order as n. Thus, we analyze the matrix $\mathbf{\Lambda}$ in (4.3) containing the eigenvalues to show that there is a constant fraction of non-negligible eigenvalues as n becomes large.

Lemma 4.1. Given the continuous-time radar signal $s(t)$ and local scattering $x(t)$, if the Fourier transform of the windowed $(s*x)(t)$ is above a constant level for a constant fraction of time, then a constant fraction of the eigenvalues of $\mathbf{SS}^H$ as $n \to \infty$ will also be above a constant level.

Proof. The covariance matrix $\mathbf{SS}^H$ is an n-by-n Toeplitz and Hermitian matrix. Let T_n denote this n-by-n Toeplitz matrix. By [67], the sequence $\{T_n\}$ is asymptotically equivalent to a sequence of n-by-n circulant matrices $\{C_n\}$ derived from the Toeplitz

matrices as $n \to \infty$, and the asymptotic equivalence of the circulant and Toeplitz matrices implies the individual asymptotic convergence of the eigenvalues for T_n to those of C_n. By [68, Eq. (5)], $C_n \triangleq F_n^H \Delta_n F_n$, where Δ_n is a diagonal matrix containing the eigenvalues of C_n, and F_n is the n-by-n (discrete Fourier transform) DFT matrix. Thus, $\Delta_n \sim \Lambda_n$, i.e., the matrix containing the eigenvalues of T_n is asymptotically equivalent to Δ_n; hence, is also equivalent to the set of diagonal elements of the covariance matrix $F_n T_n F_n^H$ of the DFT $F_n \mathbf{Sx}$ of $\mathbf{Sx}$ [69]. We aim to study the fraction of non-negligible eigenvalues of the covariance matrix $\frac{1}{\sigma_x^2} E\left[(\mathbf{Sx})(\mathbf{Sx})^H\right]$ (i.e., the fraction of non-negligible diagonal entries of $\frac{1}{\sigma_x^2} E\left[(F_n\mathbf{Sx})(F_n\mathbf{Sx})^H\right]$) as n goes large. The sequence $\{\lambda_i\}_{i=1}^n$ of the eigenvalues is

$$\{\lambda_i\}_{i=1}^n = \frac{1}{\sigma_x^2} E\left[\left|F_n\mathbf{Sx}\right|^2\right] .$$

Since increasing n will only result in a higher resolution of the DFT, then as long as the Fourier transform of the windowed $(s * x)(t)$ is above a constant non-negligible level for a constant fraction of time, there will be a constant fraction of non-negligible eigenvalues of $\mathbf{SS}^H$ as $n \to \infty$. $\square$

In Fig. 4.2, we provide an example of the eigenvalues of the covariance matrix $\frac{1}{\sigma_x^2} E\left[(\mathbf{Sx})(\mathbf{Sx})^H\right]$ when the radar signal is a constant frequency modulated continuous-wave radar waveform. We observe that as n increases, the shape of the curves of the eigenvalues stays the same, and there is a constant fraction of non-negligible eigenvalues as n becomes large.

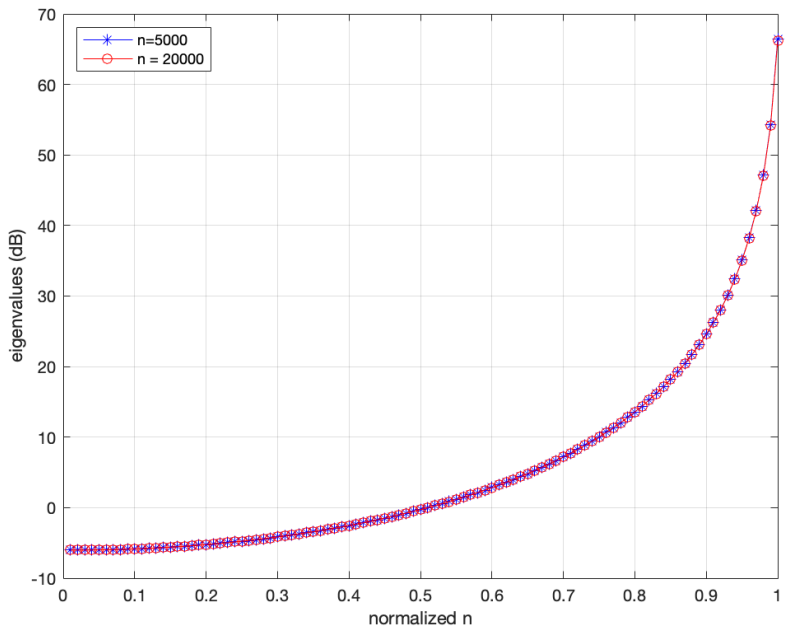


Figure 4.2: Comparison of the eigenvalues (in dB scale) of $\mathbf{S}\mathbf{S}^H$ when $n = 5000$ and $n = 20000$.

4.3.2 Willie's Detection Capability

Willie's receiver observes the vector $\mathbf{z}$ given in (4.1) and (4.2) when H_0 and H_1 are true, respectively, and attempts to decide between the two hypotheses. Since Willie knows the radar signal and hence can learn the eigenstructure of $\mathbf{S}\mathbf{S}^H$, he can first perform $(\mathbf{\Lambda}+\mathbf{I})^{-\frac{1}{2}}\mathbf{V}^H\mathbf{z}$ to decorrelate $\mathbf{z}$ without loss of optimality, since the operation is invertible. Then, the two hypotheses become:

- H_0: $(\mathbf{\Lambda}+\mathbf{I})^{-\frac{1}{2}}\mathbf{V}^H\mathbf{z} = (\mathbf{\Lambda}+\mathbf{I})^{-\frac{1}{2}}\mathbf{V}^H(\mathbf{S}\mathbf{x}+\mathbf{u})$;
- H_1: $(\mathbf{\Lambda}+\mathbf{I})^{-\frac{1}{2}}\mathbf{V}^H\mathbf{z} = (\mathbf{\Lambda}+\mathbf{I})^{-\frac{1}{2}}\mathbf{V}^H(\mathbf{S}\mathbf{x}+\mathbf{u}) + \sigma_c\mathbf{b}_k$.

Let us denote $\mathbf{w} = (\mathbf{\Lambda}+\mathbf{I})^{-\frac{1}{2}}\mathbf{V}^H(\mathbf{S}\mathbf{x}+\mathbf{u})$. Since $\mathbf{x}$ is an i.i.d. zero-mean complex Gaussian random vector, the elements of $\mathbf{S}\mathbf{x}$ are jointly Gaussian. Then, $\mathbf{w}$ is a zero-mean complex Gaussian random vector with covariance matrix:

$$
\begin{aligned}
E[\mathbf{w}\mathbf{w}^H] =& (\mathbf{\Lambda}+\mathbf{I})^{-\frac{1}{2}}\mathbf{V}^H E[(\mathbf{S}\mathbf{x}+\mathbf{u})(\mathbf{S}\mathbf{x}+\mathbf{u})^H]\mathbf{V}(\mathbf{\Lambda}+\mathbf{I})^{-\frac{1}{2}} \\
=& \begin{bmatrix} \frac{\sigma_x^2\lambda_1+\sigma_u^2}{\lambda_1+1} & 0 & \dots & 0 \\ 0 & \frac{\sigma_x^2\lambda_2+\sigma_u^2}{\lambda_2+1} & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & \frac{\sigma_x^2\lambda_n+\sigma_u^2}{\lambda_n+1} \end{bmatrix}.
\end{aligned} \tag{4.5}
$$

A straightforward test for Willie is a power detector on $(\mathbf{\Lambda}+\mathbf{I})^{-\frac{1}{2}}\mathbf{V}^H\mathbf{z}$. However, this threshold test is not necessarily optimal for Willie. Typically, constructing the optimal detector for Willie is the most challenging task. Thus, many works just assume that the optimal detector is a threshold test on the power without establishing it. Sobers *et al.* in [9] shows that this is not true except for a limited set of distributions. In this work, we can show that $\mathbf{z}^H\mathbf{V}(\mathbf{\Lambda}+\mathbf{I})^{-1}\mathbf{V}^H\mathbf{z}$ is a sufficient statistic for a genie-aided Willie, which provides an upper bound for Willie's detection capability and guarantees achievability against an optimal Willie. The threshold test on $\mathbf{z}^H\mathbf{V}(\mathbf{\Lambda}+\mathbf{I})^{-1}\mathbf{V}^H\mathbf{z}$ is optimal when σ_x^2 is uniformly distributed according to [9].

While it is hard to prove the optimality of the power detector on $(\mathbf{\Lambda}+\mathbf{I})^{-\frac{1}{2}}\mathbf{V}^H\mathbf{z}$ when $\mathbf{w}$ is not an i.i.d. sequence, we can again upper bound Willie's performance by assuming Willie has a powerful receiver that reduces the interference power on certain eigendirections to result in an i.i.d. observing sequence, i.e., Willie eventually observes $\tilde{\mathbf{w}}$ and $\tilde{\mathbf{w}}+\sigma_c\mathbf{b}_k$ when H_0 and H_1 are true, respectively, where $\tilde{\mathbf{w}}$ is an i.i.d. zero-mean complex Gaussian random vector with variance $\min\limits_{0<i\leq n}\frac{\sigma_x^2\lambda_i+\sigma_u^2}{\lambda_i+1}$ which we denote as $\sigma_{\tilde{w}}^2$. Note that since we want our covert communication scheme to be effective against an optimal Willie, assuming a genie-aided Willie that is able to perform such an operation only helps Willie, and hence, as notes above, provides an upper bound on Willie's detection capability.

Let $\mathbf{z}' = \{z_i'\}_{i=1}^n$ denote the i.i.d vector observed by the genie-aided Willie after the above operation and θ denote the variance of the power of $\mathbf{z}'$. Then, $\theta = \sigma_{\tilde{w}}^2$ under

H_0 and $\theta = \sigma_{\tilde{w}}^2 + 1$ under H_1. Note that θ is random since σ_x^2 in $\sigma_{\tilde{w}}^2$ is random. The pdf of $\mathbf{z}'$ is given by:

$$\begin{aligned} f_{\mathbf{z}'}(\mathbf{z}') &= E_{\sigma_x^2}\left[\prod_{i=1}^{n} \frac{1}{\pi\theta} \cdot \exp\left(-\frac{|z_i'|^2}{\theta}\right)\right] \\ &= E_{\sigma_x^2}\left[\left(\frac{1}{\pi\theta}\right)^n \cdot \exp\left(-\frac{\sum_{i=1}^{n}|z_i'|^2}{\theta}\right)\right] \end{aligned} \tag{4.6}$$

Thus, by the Fisher-Neyman Factorization Theorem, the total power $\sum_{i=1}^{n}|z_i'|^2$ is a sufficient statistic for the test of the genie-aided Willie, based on which we will prove covertness of our communication system.

If σ_x^2 is uniformly distributed, then using a similar method to [9] (employing LRT and *stochastic ordering* [66]), it can be shown that the optimal test for the genie-aided Willie is a threshold test on $\sum_{i=1}^{n}|z_i'|^2$, i.e.,

$$\sum_{i=1}^{n}|z_i'|^2 \underset{H_0}{\overset{H_1}{\gtrless}} \tau \tag{4.7}$$

where τ is the threshold. However, an optimal threshold test on the power is not necessary for our proof in Section 4.4.

4.4 Covertness

In this section, we prove the covertness of the system with the genie-aided Willie when Alice sends the communication waveforms given in (4.4). Assuming we do not know the distribution of σ_x^2, we will only consider $\sum_{i=1}^{n}|z_i'|^2$ being a sufficient statistic for the genie-aided Willie, and prove the limit on the covert rate in this case.

Under H_0, we write $\sum_{i=1}^{n}|z_i'|^2 = \sum_{i=1}^{n}|\tilde{w}_i|^2$ with $\tilde{w}_i \sim \mathcal{CN}(0, \sigma_{\tilde{w}}^2)$ where $\sigma_{\tilde{w}}^2 = \min_{0<i\leq n} \frac{\sigma_x^2\lambda_i+\sigma_u^2}{\lambda_i+1}$; and under H_1, we write $\sum_{i=1}^{n}|z_i'|^2 = \sum_{i=1}^{n}|\tilde{w}_i|^2 + \sigma_c^2\mathbf{b}_k^H\mathbf{b}_k$. Define a region $\mathcal{R} = [0, L)$ on the real line such that L is a constant, and given any $\epsilon > 0$:

$$\mathbb{P}\left(\frac{1}{n}\sum_{i=1}^{n}|\tilde{w}_i|^2 \in \mathcal{R}\right) > 1 - \frac{\epsilon}{4}. \tag{4.8}$$

Define the region for $\frac{1}{n}\sum_{i=1}^{n}|z_i'|^2$ in which Willie decides there is no Alice's transmission as $\mathcal{R}_{H_0}$, i.e., if $\frac{1}{n}\sum_{i=1}^{n}|z_i'|^2 \in \mathcal{R}_{H_0}$, Willie will decide H_0 to be true. We also define the region in which Willie decides that Alice transmits as $\mathcal{R}_{H_1}$ (and $\mathcal{R}_{H_1} = \overline{\mathcal{R}_{H_0}}$). Conditioned on $\sigma_{\tilde{w}}^2 = t$, the false alarm probability of Willie's detection is then given by:

$$\begin{aligned}
\mathbb{P}_{FA}(t) &= \mathbb{P}\left(\frac{1}{n}\sum_{i=1}^{n}|z_i'|^2 \in \mathcal{R}_{H_1} \mid \sigma_{\tilde{w}}^2 = t, H_0\right) \\
&= \mathbb{P}\left(\frac{1}{n}\sum_{i=1}^{n}|\tilde{w}_i|^2 \in \mathcal{R}_{H_1} \mid \sigma_{\tilde{w}}^2 = t\right) \\
&\geq \mathbb{P}\left(\frac{1}{n}\sum_{i=1}^{n}|\tilde{w}_i|^2 \in \mathcal{R}_{H_1} \cup \overline{\mathcal{R}} \mid \sigma_{\tilde{w}}^2 = t\right) - \mathbb{P}\left(\frac{1}{n}\sum_{i=1}^{n}|\tilde{w}_i|^2 \in \overline{\mathcal{R}} \mid \sigma_{\tilde{w}}^2 = t\right) \\
&> \mathbb{P}\left(\frac{1}{n}\sum_{i=1}^{n}|\tilde{w}_i|^2 \in \mathcal{R}_{H_1} \cup \overline{\mathcal{R}} \mid \sigma_{\tilde{w}}^2 = t\right) - \frac{\epsilon}{4}
\end{aligned} \tag{4.9}$$

where $\overline{\mathcal{R}}$ denotes the complement of $\mathcal{R}$. Similarly, the conditional probability of missed detection is given by:

$$\begin{aligned}
\mathbb{P}_{MD}(t) &= \mathbb{P}\left(\frac{1}{n}\sum_{i=1}^{n}|z_i'|^2 \in \mathcal{R}_{H_0} \mid \sigma_{\tilde{w}}^2 = t, H_1\right) \\
&= \mathbb{P}\left(\frac{1}{n}\sum_{i=1}^{n}|\tilde{w}_i|^2 + \frac{1}{n}\sigma_c^2\mathbf{b}_k^H\mathbf{b}_k \in \mathcal{R}_{H_0} \mid \sigma_{\tilde{w}}^2 = t\right) \\
&\geq \mathbb{P}\left(\frac{1}{n}\sum_{i=1}^{n}|\tilde{w}_i|^2 + \frac{1}{n}\sigma_c^2\mathbf{b}_k^H\mathbf{b}_k \in \mathcal{R}_{H_0} \cup \overline{\mathcal{R}} \mid \sigma_{\tilde{w}}^2 = t\right) \\
&\quad - \mathbb{P}\left(\frac{1}{n}\sum_{i=1}^{n}|\tilde{w}_i|^2 + \frac{1}{n}\sigma_c^2\mathbf{b}_k^H\mathbf{b}_k \in \overline{\mathcal{R}} \mid \sigma_{\tilde{w}}^2 = t\right)
\end{aligned}$$

where $\frac{1}{n}\sigma_c^2\mathbf{b}_k^H\mathbf{b}_k$ can be dropped since it is a very small value due to the average power σ_c^2 is very small to allow covert communication for Alice. Thus,

$$\mathbb{P}_{MD}(\mathbf{t}) > \mathbb{P}\left(\frac{1}{n}\sum_{i=1}^{n}|\tilde{w}_i|^2 + \frac{1}{n}\sigma_c^2\mathbf{b}_k^H\mathbf{b}_k \in \mathcal{R}_{H_0}\cup\overline{\mathcal{R}} \mid \sigma_{\tilde{w}}^2 = t\right) - \frac{\epsilon}{4}. \tag{4.10}$$

Let χ_l^2 denote a chi-squared random variable with l degrees of freedom. Then, since $\tilde{w}_i \sim \mathcal{CN}(0,\sigma_{\tilde{w}}^2)$, we have $\frac{1}{n}\sum_{i=1}^{n}|\tilde{w}_i|^2 = \frac{1}{n}\sigma_{\tilde{w}}^2\chi_{2n}^2$ in distribution. By the weak law of large numbers, χ_{2n}^2/n converges in probability to one. Hence, for any $\delta > 0$, there exists an N_0 such that for any $n \geq N_0$,

$$\mathbb{P}\left(\frac{1}{n}\chi_{2n}^2 \in \left(1-\frac{\delta}{\sigma_{\tilde{w}}^2}, 1+\frac{\delta}{\sigma_{\tilde{w}}^2}\right)\right) > 1-\frac{\epsilon}{4}. \tag{4.11}$$

Hence, for any $n \geq N_0$,

$$\mathbb{P}\left(\frac{1}{n}\sum_{i=1}^{n}|\tilde{w}_i|^2 \in \left(\sigma_{\tilde{w}}^2-\delta, \sigma_{\tilde{w}}^2+\delta\right)\right) > 1-\frac{\epsilon}{4}. \tag{4.12}$$

Therefore, for any $n \geq N_0$, if $(t-\delta, t+\delta) \in \mathcal{R}_{H_1}\cup\mathcal{R}$, then from (4.9), we have $\mathbb{P}_{FA}(\mathbf{t}) > 1-\frac{\epsilon}{2}$. Likewise, following analogous arguments, there exists N_1 such that, for any $n \geq N_1$, if $(t-\delta, t+\delta) + \frac{1}{n}\sigma_c^2\mathbf{b}_k^H\mathbf{b}_k \in \mathcal{R}_{H_0}\cup\mathcal{R}$, then from (4.10), we have $\mathbb{P}_{MD}(\mathbf{t}) > 1-\frac{\epsilon}{2}$. Define $\mathcal{A}$ as the set of $\sigma_{\tilde{w}}^2$ such that:

$$(\sigma_{\tilde{w}}^2-\delta, \sigma_{\tilde{w}}^2+\delta) \in \left(\mathcal{R}_{H_1}\cup\overline{\mathcal{R}}\right) \cup \left(\mathcal{R}_{H_0}\cup\overline{\mathcal{R}} - \frac{1}{n}\sigma_c^2\mathbf{b}_k^H\mathbf{b}_k\right).$$

Then, for any $n \geq \max(N_0, N_1)$, $\mathbb{P}_{FA}(\mathbf{t}) + \mathbb{P}_{MD}(\mathbf{t}) > 1-\frac{\epsilon}{2}$ if $t \in \mathcal{A}$. As for the complement set $\overline{\mathcal{A}}$, it can be written as such that:

$$(\sigma_{\tilde{w}}^2-\delta, \sigma_{\tilde{w}}^2+\delta) \in (\mathcal{R}_{H_0}\cap\mathcal{R}) \cap \left(\mathcal{R}_{H_1}\cap\mathcal{R} - \frac{1}{n}\sigma_c^2\mathbf{b}_k^H\mathbf{b}_k\right).$$

Then, $\mathbb{P}_{FA}(\mathbf{t}) + \mathbb{P}_{MD}(\mathbf{t}) < \frac{\epsilon}{2}$ if and only if $t \in \overline{\mathcal{A}}$.

We will show next that the expected probability that $\sigma_{\tilde{w}}^2 \in \overline{\mathcal{A}}$ can be upper bounded by $\frac{\epsilon}{2}$ when σ_c^2 is set to be a certain constant number. Define a region

$\mathcal{B}_0 \in \mathcal{R}_{H_0} \cap \mathcal{R}$ such that $\mathcal{B}_0$ contains all intervals in $\mathcal{R}_{H_0} \cap \mathcal{R}$ that have a length larger than or equal to 2δ. Since $\mathcal{R}_{H_0} \cap \mathcal{R}$ has a finite length, there are only a constant number (which we denote as m_0) of such intervals in $\mathcal{B}_0$. Similarly, define a region $\mathcal{B}_1 \in \mathcal{R}_{H_1} \cap \mathcal{R}$ such that $\mathcal{B}_1$ contains all intervals in $\mathcal{R}_{H_1} \cap \mathcal{R}$ that have a length larger than or equal to 2δ. Since $\mathcal{R}_{H_1} \cap \mathcal{R}$ has a finite length, there are only a constant number (which we denote as m_1) of such intervals in $\mathcal{B}_1$. Then:

$$\begin{aligned}
E_{\mathbf{b}_k}\left[P(\overline{\mathcal{A}})\right] &= E_{\mathbf{b}_k}\left[\mathbb{P}\left(\sigma^2_{\tilde{w}} - \delta, \sigma^2_{\tilde{w}} + \delta) \in \mathcal{B}_0 \cap \left(\mathcal{B}_1 - \frac{1}{n}\sigma_c^2 \mathbf{b}_k^H \mathbf{b}_k\right)\right)\right] \\
&\leq E_{\mathbf{b}_k}\left[\mathbb{P}\left(\sigma^2_{\tilde{w}} \in \mathcal{B}_0 \cap \left(\mathcal{B}_1 - \frac{1}{n}\sigma_c^2 \mathbf{b}_k^H \mathbf{b}_k\right)\right)\right] \\
&= E_{\mathbf{b}_k}\left[\int_{\mathcal{B}_0 \cap \left(\mathcal{B}_1 - \frac{1}{n}\sigma_c^2 \mathbf{b}_k^H \mathbf{b}_k\right)} f_{\sigma^2_{\tilde{w}}}(x) dx\right] \\
&\leq E_{\mathbf{b}_k}\left[m \cdot \frac{1}{n}\sigma_c^2 \mathbf{b}_k^H \mathbf{b}_k \cdot \max_x f_{\sigma^2_{\tilde{w}}}(x)\right] \\
&= m\sigma_c^2 \cdot \max_x f_{\sigma^2_{\tilde{w}}}(x) \qquad (4.13)
\end{aligned}$$

where $m = \min(m_0, m_1)$, and $f_{\sigma^2_{\tilde{w}}}(x)$ denotes the pdf of $\sigma^2_{\tilde{w}}$ (which can be obtained from $f_{\sigma^2_x}$). Hence, choosing $\sigma_c^2 = \frac{\epsilon}{2m \cdot \max_x f_{\sigma^2_{\tilde{w}}}(x)}$ yields $E_{\mathbf{b}_k}\left[\mathbb{P}(\overline{\mathcal{A}})\right] \leq \frac{\epsilon}{2}$, and we can upper bound $\mathbb{P}_{FA} + \mathbb{P}_{MD}$ as:

$$\begin{aligned}
\mathbb{P}_{FA} + \mathbb{P}_{MD} &= E_{\sigma^2_{\tilde{w}}, \mathbf{b}_k}\left[\mathbb{P}_{FA}(\sigma^2_{\tilde{w}}) + \mathbb{P}_{MD}(\sigma^2_{\tilde{w}})\right] \\
&\geq E_{\sigma^2_{\tilde{w}}, \mathbf{b}_k}\left[\mathbb{P}_{FA}(\sigma^2_{\tilde{w}}) + \mathbb{P}_{MD}(\sigma^2_{\tilde{w}}) \mid \mathcal{A}\right] E_{\mathbf{b}_k}\left[\mathbb{P}(\mathcal{A})\right] \\
&> 1 - \frac{\epsilon}{2}
\end{aligned}$$

Therefore, Alice can transmit with constant average power σ_c^2 and remain covert from Willie. This shows that Alice can covertly transmit $\mathcal{O}(n)$ bits to Bob in n samples of a radar pulse.

4.5 Reliability

Now we examine Bob's decoding error probability $\mathbb{P}_e^{(b)}$ averaged over all codebooks using similar methods as those in [5]. The key difference is that in our case, Bob observes colored noise due to the correlation of the interference from the clutter. So first, in the same way that Willie decorrelates $\mathbf{z}$, Bob performs $(\mathbf{\Lambda}+\mathbf{I})^{-\frac{1}{2}}\mathbf{V}^H\mathbf{y}$ to decorrelate $\mathbf{y}$. Then, when Alice transmits, Bob observes:

$$(\mathbf{\Lambda}+\mathbf{I})^{-\frac{1}{2}}\mathbf{V}^H\mathbf{y} = (\mathbf{\Lambda}+\mathbf{I})^{-\frac{1}{2}}\mathbf{V}^H(\mathbf{S}\mathbf{x}'+\mathbf{u}') + \sigma_c\mathbf{b}_k\,.$$

Let $\mathbf{w}' = (\mathbf{\Lambda}+\mathbf{I})^{-\frac{1}{2}}\mathbf{V}^H(\mathbf{S}\mathbf{x}'+\mathbf{u}')$. Then $\mathbf{w}'$ is a zero-mean complex Gaussian random vector with covariance matrix:

$$E[\mathbf{w}'\mathbf{w}'^H] = \begin{bmatrix} \frac{\sigma_x'^2\lambda_1+\sigma_u'^2}{\lambda_1+1} & 0 & \dots & 0 \\ 0 & \frac{\sigma_x'^2\lambda_2+\sigma_u'^2}{\lambda_2+1} & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & \frac{\sigma_x'^2\lambda_n+\sigma_u'^2}{\lambda_n+1} \end{bmatrix}. \tag{4.14}$$

Let Bob employ a maximum-likelihood (ML) decoder (i.e., minimum distance decoder) to process his observed vector $(\mathbf{\Lambda}+\mathbf{I})^{-\frac{1}{2}}\mathbf{V}^H\mathbf{y}$ when codeword $\mathbf{c}_k$ (i.e. key $\mathbf{b}_k$) was sent. Let $Err_i(\mathbf{b}_k)$ denote the event that the decoder suffers an error, i.e., when $(\mathbf{\Lambda}+\mathbf{I})^{-\frac{1}{2}}\mathbf{V}^H\mathbf{y}$ is closer to another key $\mathbf{b}_i$ for $i \neq k$. Then,

$$\begin{aligned} \mathbb{P}_e^{(b)} &= E_{\mathbf{b}_k}\left[\mathbb{P}\left(\cap_{i=1,i\neq k}^{2^{nR}} Err_i(\mathbf{b}_k)\right)\right] \\ &\leq \sum_{i=1,i\neq k}^{2^{nR}} E_{\mathbf{b}_k}\left[\mathbb{P}\left(Err_i(\mathbf{b}_k)\right)\right] \end{aligned} \tag{4.15}$$

where (4.15) is obtained using a union bound.

If $\mathbf{w}'$ is an i.i.d complex Gaussian random vector with variance $\sigma^2_{w'}$, then by [64]:

$$E_{\mathbf{b}_k}\left[\mathbb{P}\left(Err_i(\mathbf{b}_k)\right)\right] = E_{\mathbf{b}_k,\sigma^2_{w'}}\left[Q\left(\sqrt{\frac{\sigma_c^2||\mathbf{b}_k-\mathbf{b}_i||_2^2}{4\sigma^2_{w'}}}\right)\right]$$

where $Q(\cdot)$ denotes the Q-function, $||\cdot||^2$ denotes the $\mathcal{L}_2$ norm, and $||\mathbf{b}_k-\mathbf{b}_i||_2^2 = (\mathbf{b}_k-\mathbf{b}_i)^H(\mathbf{b}_k-\mathbf{b}_i)$ with $(\mathbf{b}_k-\mathbf{b}_i)\sim\mathcal{CN}(0,2)$. However, recall that in our case, $\mathbf{w}'$ is not identically distributed, but rather has covariance matrix given in (4.14). Since increasing the power of the interference can only degrade Bob's decoding capability, we can then upper bound $E_{\mathbf{b}_k}\left[\mathbb{P}\left(Err_i(\mathbf{b}_k)\right)\right]$ as:

$$\begin{aligned}E_{\mathbf{b}_k}\left[\mathbb{P}\left(Err_i(\mathbf{b}_k)\right)\right] &\leq E_{\mathbf{b}_k,\sigma^2_{x'}}\left[Q\left(\sqrt{\frac{\sigma_c^2||\mathbf{b}_k-\mathbf{b}_i||_2^2}{4\max\limits_{0<i\leq n}\frac{\sigma^2_{x'}\lambda_i+\sigma^2_{u'}}{\lambda_i+1}}}\right)\right]\\ &= E_{\mathbf{b}_k,\sigma^2_{x'}}\left[Q\left(\sqrt{\frac{\sigma_c^2||\mathbf{b}_k-\mathbf{b}_i||_2^2}{4(\sigma^2_{x'}\lambda_1+\sigma^2_{u'})}}\right)\right]\end{aligned}$$

Let $W=\frac{1}{2}(\mathbf{b}_k-\mathbf{b}_i)^H(\mathbf{b}_k-\mathbf{b}_i)$, then:

$$E_{\mathbf{b}_k}\left[\mathbb{P}\left(Err_i(\mathbf{b}_k)\right)\right] \leq E_{W,\sigma^2_{x'}}\left[Q\left(\sqrt{\frac{\sigma_c^2 W}{2(\sigma^2_{x'}\lambda_1+\sigma^2_{u'})}}\right)\right] \tag{4.16}$$

where $W\sim\chi^2_{2n}$. Since $Q(x)\leq\frac{1}{2}e^{-x^2/2}$ [65]:

$$
\begin{aligned}
E_{W,\sigma_{x'}^2}\left[Q\left(\sqrt{\frac{\sigma_c^2 W}{2(\sigma_{x'}^2\lambda_1+\sigma_{u'}^2)}}\right)\right] &\leq E_{W,\sigma_{x'}^2}\left[\exp\left(-\frac{\sigma_c^2 W}{4(\sigma_{x'}^2\lambda_1+\sigma_{u'}^2)}\right)\right] \\
&= E_{\sigma_{x'}^2}\left[\int_0^\infty \frac{e^{-\frac{\sigma_c^2 x}{4(\sigma_{x'}^2\lambda_1+\sigma_{u'}^2)}-\frac{x}{2}}2^{-n}x^{n-1}}{\Gamma(n)}dx\right] \\
&= E_{\sigma_{x'}^2}\left[2^{-n}\left(\frac{1}{2}+\frac{\sigma_c^2}{4(\sigma_{x'}^2\lambda_1+\sigma_{u'}^2)}\right)^{-n}\right] \\
&= E_{\sigma_{x'}^2}\left[2^{-n\log_2\left(1+\frac{\sigma_c^2}{2(\sigma_{x'}^2\lambda_1+\sigma_{u'}^2)}\right)}\right] \\
&\leq 2^{-n\log_2\left(1+\frac{\sigma_c^2}{2\left(E\left[\sigma_{x'}^2\right]\lambda_1+\sigma_{u'}^2\right)}\right)} \qquad (4.17)
\end{aligned}
$$

where $\Gamma(n) = \int_0^\infty x^{n-1}e^{-x}dx$ is the Gamma function, and (4.17) is obtained by Jensen's inequality. Thus, Bob's decoding error probability is upper bounded as:

$$
\mathbb{P}_e^{(b)} \leq 2^{nR-n\log_2\left(1+\frac{\sigma_c^2}{2\left(E\left[\sigma_{x'}^2\right]\lambda_1+\sigma_{u'}^2\right)}\right)}.
$$

If σ_c^2 is constant, then as long as $R < \log_2\left(1+\frac{\sigma_c^2}{2\left(E\left[\sigma_{x'}^2\right]\lambda_1+\sigma_{u'}^2\right)}\right)$, $\mathbb{P}_e^{(b)}$ approaches to zero exponentially as n becomes large, and Bob obtains $nR = n\log_2\left(1+\frac{\sigma_c^2}{2\left(E\left[\sigma_{x'}^2\right]\lambda_1+\sigma_{u'}^2\right)}\right)$ bits in n samples of a radar pulse.

4.6 Conclusion

In this chapter, we have studied covert communication that relies on the interference from a pulsed radar system. Covertness is achieved by embedding the transmission into the scattering of the radar. We have provided a design of the covert communication waveform. In the case of AWGN channels, we have analyzed the capability of both the optimal detection at Willie and the decoding at Bob. We have established that $\mathcal{O}(n)$ bits can be covertly and reliably transmitted from Alice to Bob in n samples of a radar pulse.

CHAPTER 5

FUNDAMENTAL LIMITS IN DETECTING WHETHER A SIGNAL HAS BEEN QUANTIZED

5.1 Introduction

In many scenarios, it is important to know whether a received signal was sent by a friend directly, or whether it was recorded by an adversary and then replayed. This is the common replay attack (or playback attack) in network security, where the attacker records messages from a transmitter to a receiver and replays the massages later to trick the receiver into unauthorized operations. This type of attack can be used effectively in many real-world applications like remote keyless-entry system for vehicles [73, 74] or text-dependent speaker verification [75].

Similar attacks also occur in radar jamming and deception to protect targets from being detected by enemy radar systems. Deceptive jamming uses techniques like range gate pull-off (RGPO) [76] to break a radar lock from the target. The basic idea is to generate a signal pulse very similar to the one that is reflected by the target, and then send it a fraction of time later so that the radar's range gate starts to follow the false pulse instead of the real reflection. Along with the appearance of digital radio frequency memory (DRFM) [77], a modern deception jammer can capture and retransmit the radiation signal of the target, producing a false signal that confuses the receiver radar and hides the target's real position or velocity.

Many approaches have been studied to prevent a replay attack in network security or deception jamming in radar systems: [78] presents a comparison of different feature extraction techniques and classifiers for replay attack detection; [79] provides

a mobile payment scheme based on radio frequency identification that can prevent replay attack; [80] theoretically studies the detection and classification of jamming signals by analyzing the adaptive coherent estimator and the generalized likelihood ratio test; and [81] proposes a DRFM deception jamming detection approach based on singular spectrum analysis. Although these methods are proposed to efficiently detect false signals that are recorded and replayed in the field of network security and radar jamming, the fundamental limits of such attacks with hardware imperfections has not been explored. In either the replay attack or deception jamming, imperfect hardware such as quantization or nonlinearities of RF components has a significant impact on the detection of signals. Here, motivated by our previous work in identifying transmitters based on subtle imperfections [82]– [84], we initiate a study in employing such imperfections in detecting a recording of the signal. Learning the fundamental thresholds for the characteristics of the hardware will provide us with both theoretical insight and application utility.

We start by studying the theoretical limits in the detection of quantized signals. Analogous to [5] and [11] where a theoretical limit on the amount of information transmitted reliably without detection is presented, we are interested in finding the characteristics of the quantizer employed in a replay system that avoid its detection. Based on the mathematics of statistical hypothesis testing, we provide a limit on the quantization bits and the quantizer span such that the quantized signal essentially cannot be detected. We consider a discrete-time model where signals are discrete-time series. We will show that if a signal with length m is sent, and the quantization is uniform with b bits, then $2^b = \omega\left(\sqrt{m}\right)$ and a quantizer span of $\omega\left(\sqrt{\ln m}\right)$ are sufficient to avoid detection. Conversely, having $2^b = \mathcal{O}\left(\sqrt{m}\right)$ or a quantizer span of $o\left(\sqrt{\ln m}\right)$ results in detection by the observer with high probability as $m \to \infty$.

The rest of this chapter is organized as follows: after introducing the system model and metric in Section 5.2, we derive the optimal hypothesis test and the probability of

error in distinguishing the original signal and the quantized signal in Section 5.3. We prove the achievability and converse results in Section 5.4 and Section 5.5 respectively. Section 5.6 draws the conclusion.

5.2 System Model and Metrics

We employ a discrete-time model with real-valued elements. The framework of our system is shown in Fig. 5.1, where Alice sends a vector $\boldsymbol{X} = \{X_i\}_{i=1}^m$ of m real values and an observer receives a vector $\boldsymbol{Y} = \{Y_i\}_{i=1}^m$. If the signal $\boldsymbol{X}$ is sent directly, which we term as the "original signal", then $Y_i = X_i + N_i$ for $i = 1, 2, \ldots, m$, is an independent and identically distributed (i.i.d.) sequence of Gaussian random variables, where $N_i \sim \mathcal{N}(0, \sigma^2)$ is the noise on the channel between the observer and Alice. However, if the signal is recorded and replayed, then the signal $\boldsymbol{X}$ is first quantized before being sent through the channel; in this case, $Y_i = Q(X_i) + N_i$ for $i = 1, 2, \ldots, m$, where $Q(\cdot)$ is the quantization function taking the original signal as input and outputing the quantized signal.

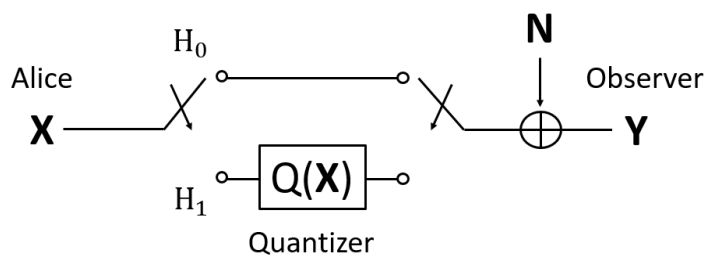


Figure 5.1: System framework: Alice sends a real-valued vector $\boldsymbol{X}$ and an observer attempts to classify his observed vector $\boldsymbol{Y}$ as either a vector $\boldsymbol{X} + \boldsymbol{N}$ of the original signal through the channel or a vector $Q(\boldsymbol{X}) + \boldsymbol{N}$ of the quantized signal through the channel.

The goal of our work is to study how the parameters of the quantizer affect the ability of the observer to distinguish the original signal and the quantized signal. In order to distinguish the two signals, the observer performs a statistical hypothesis test on his observed vector $\boldsymbol{Y}$. We define:

- H_0: $\boldsymbol{Y} = \boldsymbol{X} + \boldsymbol{N}$;
- H_1: $\boldsymbol{Y} = Q(\boldsymbol{X}) + \boldsymbol{N}$.

Our metric is the probability of error. We assume that the observer uses classical hypothesis testing with equal prior probabilities of each of the two hypotheses being true. The rejection of H_0 when it is true is known as a type I error (or false alarm), and we denote its probability as α [85]. The acceptance of H_0 when it is not true is known as a type II error (or missed detection), and we denote β to be its probability. Thus, the error probability for the observer to distinguish the two hypotheses can be written as $P_e = \frac{\alpha+\beta}{2}$.

We also assume that the observer knows $\boldsymbol{X}$ and the realization of the quantization function Q, and he also knows the variance σ^2 of the Gaussian noise on the channel between the observer and Alice. So, the observer is aware of the statistics of the two hypotheses H_0 and H_1.

5.3 Optimal Test and the Probability of Error

The observer's goal is to determine whether his observed vector is the original signal sent through the channel or the signal that is quantized and then sent through the channel. The observer will make his decision based on the optimal test between H_0 and H_1. The likelihood ratio test (LRT) is:

$$\Omega(\boldsymbol{Y} = \boldsymbol{y}) = \frac{f_{\boldsymbol{Y}|H_0}(\boldsymbol{y})}{f_{\boldsymbol{Y}|H_1}(\boldsymbol{y})} \underset{H_1}{\overset{H_0}{\gtrless}} 1$$

where $f_{\boldsymbol{Y}|H_0}(\boldsymbol{y})$ is the probability distribution of $\boldsymbol{Y}$ given H_0 is true and $f_{\boldsymbol{Y}|H_1}(\boldsymbol{y})$ is the probability distribution of $\boldsymbol{Y}$ given H_1 is true.

Given that $\boldsymbol{X} = \boldsymbol{x}$, since the observer knows $\boldsymbol{x}$ and $Q(\boldsymbol{x})$, we have $Y_i \sim \mathcal{N}(x_i, \sigma^2), i = 1, 2, \ldots, m$ when H_0 is true (i.e., $Y_i = x_i + N_i$), and $Y_i \sim \mathcal{N}(Q(x_i), \sigma^2), i = 1, 2, \ldots, m$ when H_1 is true (i.e., $Y_i = Q(x_i) + N_i$). Thus, the LRT can be written as:

$$\Omega(\boldsymbol{Y} = \boldsymbol{y}) = \frac{\prod_{i=1}^{m} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y_i - x_i)^2}{\sigma^2}}}{\prod_{i=1}^{m} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y_i - Q(x_i))^2}{\sigma^2}}} \underset{H_1}{\overset{H_0}{\gtrless}} 1$$

which could be further written as:

$$\sum_{i=1}^{m} \left[-(y_i - x_i)^2 + (y_i - Q(x_i))^2 \right] \underset{H_1}{\overset{H_0}{\gtrless}} 0\,,$$

i.e.,

$$\sum_{i=1}^{m} \left[2y_i(x_i - Q(x_i)) + (Q(x_i))^2 - x_i^2 \right] \underset{H_1}{\overset{H_0}{\gtrless}} 0\,.$$

Now we would like to derive the error probability for the above test. Let us denote $Z_i = 2Y_i(X_i - Q(X_i)) + (Q(X_i))^2 - X_i^2, i = 1, 2, \ldots, m$. Given H_0 is true and $\boldsymbol{X}$ is known to the observer, then $Y_i = X_i + N_i$ and $Y_i \sim \mathcal{N}(X_i, \sigma^2), i = 1, 2, \ldots, m$. In this case, we see that Z_i is a Gaussian random variable with mean $2X_i(X_i - Q(X_i)) + (Q(X_i))^2 - X_i^2$ and variance $4(X_i - Q(X_i))^2\sigma^2$. Thus, under the assumption of equal prior probabilities of H_0 and H_1, the probability of false alarm is given by:

$$\alpha = P\left(\sum_{i=1}^{m} Z_i < 0 \right) = \frac{1}{2} - \frac{1}{2}\mathrm{erf}\left(\frac{\sqrt{\sum_{i=1}^{m} U_i^2}}{2\sqrt{2}\sigma} \right)$$

where $U_i = X_i - Q(X_i)$, $i = 1, 2, \ldots, m$, is the quantization error which is uniformly distributed over $[-\frac{\Delta}{2}, \frac{\Delta}{2}]$. The quantization step size is denoted as Δ.

By the symmetry of the problem, the probability of missed detection has the same form ($\alpha = \beta$). Thus, the error probability for the observer to distinguish the two hypotheses is given by:

$$P_e(\boldsymbol{X}) = \frac{1}{2} - \frac{1}{2}\text{erf}\left(\frac{\sqrt{\sum_{i=1}^{m} U_i^2}}{2\sqrt{2}\sigma}\right). \tag{5.1}$$

In order to confuse the observer, we want the error probability P_e to be close to $\frac{1}{2}$ so that it is equivalent to a random guess for the observer in distinguishing the original and the quantized signal, meaning that we want $P_e \geq \frac{1}{2} - \epsilon$ for any $\epsilon > 0$ [5]. Conversely, we want $P_e \leq \epsilon$ for any $\epsilon > 0$ for the observer to be able to detect the quantized signal with high probability.

5.4 Achievability

In this section, we will state the achievability theorems under the assumption that the quantizer in our system is uniform. We seek the sufficient quantization step size Δ as a function of the vector length m. Intuitively, Δ should be small, and since we do not want Δ goes to infinity, we must have $\Delta = \mathcal{O}(1)$ for all circumstances. We can write $\Delta = \frac{C}{2^b}$ where b is the number of quantization bits and C is a constant.

Theorem 5.1 (Achievability under uniform quantization with no overflows)**.** Suppose that Alice sends a length-m discrete-time signal that never exceeds the range of the quantizer, and the quantizer is uniform with b bits of quantization. If $2^b = \omega\left(\sqrt{m}\right)$ (in particular, $b \geq \log_2 \frac{C\sqrt{m}}{4\sqrt{6}\sigma\text{erf}^{-1}(2\epsilon)}$, where C is a constant and σ is the standard deviation of the noise on the channel), then the observer can only distinguish the original signal and the quantized signal with error probability $P_e \geq \frac{1}{2} - \epsilon$ for any $\epsilon > 0$.

Proof. We will first take the expectation of $P_e(\boldsymbol{X})$ in (5.1) over $U_i \in [-\frac{\Delta}{2}, \frac{\Delta}{2}]$ for $i = 1, 2, \ldots, m$. Since $-\text{erf}(\cdot)$ is a convex function, then by Jensen's inequality we have:

$$P_e = E_{\boldsymbol{X}}[P_e(\boldsymbol{X})] \geq \frac{1}{2} - \frac{1}{2}\text{erf}\left(\frac{\sqrt{E[\sum_{i=1}^{m} U_i^2]}}{2\sqrt{2}\sigma}\right) \tag{5.2}$$

which is a tight lower bound since $\text{erf}(x)$ is approximately linear at small x. Substituting $E[\sum_{i=1}^{m} U_i^2] = \frac{m\Delta^2}{12}$ and $\Delta = \frac{C}{2^b}$ in (5.2) yields:

$$P_e \geq \frac{1}{2} - \frac{1}{2}\text{erf}\left(\frac{\frac{C}{2^b}\sqrt{\frac{m}{12}}}{2\sqrt{2}\sigma}\right).$$

If $b \geq \log_2 \frac{C\sqrt{m}}{4\sqrt{6}\sigma \text{erf}^{-1}(2\epsilon)}$ for any $\epsilon > 0$, then $\frac{1}{2}\text{erf}\left(\frac{\frac{C}{2^b}\sqrt{\frac{m}{12}}}{2\sqrt{2}\sigma}\right) \leq \epsilon$. This implies that $P_e \geq \frac{1}{2} - \epsilon$ for any $\epsilon > 0$. Thus, $2^b = \omega(\sqrt{m})$ is sufficient to prevent detection by the observer. □

In many scenarios, the transmitted signals can have a wide range of values; for example, a Gaussian signal would have values ranging from negative infinity to infinity. In this case, the quantizer would have overflows that can assist the observer in detecting the signal. So, we next assume that the original signal $X_i \sim \mathcal{N}(0, \sigma_0^2), i = 1, 2, \ldots, m$. We consider that the system employs quantization with a span of $[-l, l]$. If its input has value within this span, it outputs the quantized value using the quantization function Q. If the input value is outside of the span, it overflows and outputs either $-l$ or l. To keep the observer from detecting the overflows, l should be scaled with the length m of the transmitted signal. Intuitively, l must go to infinity as $m \to \infty$ or the quantization is readily detected. Thus, we consider $l = \omega(1)$ for all circumstances. We obtain the achievability result in this case as below.

Theorem 5.2 (Achievability under uniform quantization with overflows)**.** Suppose that Alice sends a discrete-time signal with length m and the quantizer is uniform with b bits of quantization and a span of $[-l, l]$. Then, if $2^b = \omega\left(\sqrt{m}\right)$ bits and $l = \omega\left(\sqrt{\ln m}\right)$, the observer can only distinguish the original signal and the quantized signal with error probability $P_e \geq \frac{1}{2} - \epsilon$ for any $\epsilon > 0$.

Proof. Since the quantizer has a span of $[-l, l]$, then if hypothesis H_1 is true, Y_i can be written as:

$$Y_i = \begin{cases} Q(X_i), & |X_i| < l \\ l, & X_i > l \\ -l, & X_i < -l \end{cases}, \quad i = 1, 2, \ldots, m$$

and the quantization error can be written as:

$$U_i = \begin{cases} X_i - Q(X_i), & |X_i| < l \\ X_i - l, & X_i > l \\ X_i + l, & X_i < -l \end{cases}, \quad i = 1, 2, \ldots, m\,. \tag{5.3}$$

Recall that $X_i \sim \mathcal{N}(0, \sigma_0^2)$ and $E[U_i^2] = \frac{\Delta^2}{12}$ when $U_i = X_i - Q(X_i)$ for $i = 1, 2, \ldots, m$. Then, we derive the expectation of U_i^2 for the general case in (5.3) as:

$$\begin{aligned} E[U_i^2] &= P\left(|X_i| < l\right) \cdot \frac{\Delta^2}{12} + P\left(X_i > l\right) \cdot \int_0^\infty \frac{1}{\sqrt{2\pi\sigma_0^2}} e^{-\frac{(x+l)^2}{2\sigma_0^2}} x^2 dx \\ &\quad + P\left(X_i < -l\right) \cdot \int_{-\infty}^0 \frac{1}{\sqrt{2\pi\sigma_0^2}} e^{-\frac{(x-l)^2}{2\sigma_0^2}} x^2 dx \\ &= \frac{\Delta^2}{12} \mathrm{erf}\left(\frac{l}{\sqrt{2}\sigma_0}\right) + \left(\frac{\sigma_0^2 l e^{-\frac{l^2}{2\sigma_0^2}}}{\sqrt{2\pi}} + \frac{l^2 + \sigma_0^2}{2}\left(1 - \mathrm{erf}\left(\frac{l}{\sqrt{2}\sigma_0}\right)\right)\right)\left(1 - \mathrm{erf}\left(\frac{l}{\sqrt{2}\sigma_0}\right)\right). \end{aligned} \tag{5.4}$$

Substituting (5.4) in (5.2) we have a lower bound for P_e in this case when the quantizer has overflows:

$$P_e \geq \frac{1}{2} - \frac{1}{2}\text{erf}\left(\frac{\sqrt{mE[U_i^2]}}{2\sqrt{2}\sigma}\right). \tag{5.5}$$

Note that l goes to infinity as $m \to \infty$. If $2^b = \omega\left(\sqrt{m}\right)$, then:

$$\frac{m\Delta^2}{12}\text{erf}\left(\frac{l}{\sqrt{2}\sigma_0}\right) \leq \frac{m\Delta^2}{12} \to 0\,,$$

i.e., the first term of $mE[U_i^2]$ goes to zero as $m \to \infty$.

On the other hand, if $l = \omega\left(\sqrt{\ln m}\right)$, we have:

$$m\left(\frac{\sigma_0^2 l e^{-\frac{l^2}{2\sigma_0^2}}}{\sqrt{2\pi}} + \frac{l^2+\sigma_0^2}{2}\cdot\frac{e^{-\frac{l^2}{2\sigma_0^2}}}{\sqrt{\pi}\frac{l}{\sqrt{2}\sigma_0}}\right)\frac{e^{-\frac{l^2}{2\sigma_0^2}}}{\sqrt{\pi}\frac{l}{\sqrt{2}\sigma_0}} \to 0 \tag{5.6}$$

since keeping only the dominant terms in (5.6) yields $\frac{m\sigma_0(\sigma_0^2+1)e^{-\frac{l^2}{\sigma_0^2}}}{\pi} \to 0$ when $l = \omega\left(\sqrt{\ln m}\right)$. If we take the Taylor series expansion of $\text{erf}\left(\frac{l}{\sqrt{2}\sigma_0}\right)$ at $\frac{l}{\sqrt{2}\sigma_0} = \infty$, then the second term of $mE[U_i^2]$ is upper bounded by (5.6), and hence goes to zero as $m \to \infty$.

Therefore, letting $2^b = \omega\left(\sqrt{m}\right)$ and $l = \omega\left(\sqrt{\ln m}\right)$, we get $mE[U_i^2] \to 0$ as $m \to \infty$, which implies that $\frac{1}{2}\text{erf}\left(\frac{\sqrt{mE[U_i^2]}}{2\sqrt{2}\sigma}\right) \leq \epsilon$ for any $\epsilon > 0$. By (5.5), we have $P_e \geq \frac{1}{2} - \epsilon$ for any $\epsilon > 0$.

□

5.5 Converse

In this section, we provide the converse results under the assumption that the quantizer is uniform with step size Δ.

Theorem 5.3 (Converse under uniform quantization with no overflows)**.** Suppose that Alice sends a length-m discrete-time signal that never exceeds the range of the quantizer, and the quantizer is uniform with b bits of quantization levels. If $2^b = \mathcal{O}\left(\sqrt{m}\right)$ (in particular, $b \leq \log_2 \frac{C\sqrt{m}}{8\sqrt{2}\sigma\sqrt{\ln\frac{1}{2\epsilon}}}$ for any $\epsilon > 0$, where C is a constant and σ is the standard deviation of the noise on the channel), then the observer can distinguish the original signal and the quantized signal with arbitrarily low probability of error.

Proof. For the achievability result, we give a tight lower bound (5.2) to the error probability P_e. Now, we need an analogous upper bound. Using the fact that $\mathrm{erf}(x) \geq 1 - e^{-x^2}$ [72], we upper bound $P_e(\boldsymbol{X})$ in (5.1) as:

$$P_e(\boldsymbol{X}) \leq \frac{1}{2} e^{-\frac{\sum_{i=1}^{m} U_i^2}{8\sigma^2}} = \frac{1}{2} \prod_{i=1}^{m} e^{-\frac{U_i^2}{8\sigma^2}}, \tag{5.7}$$

and taking the expectation yields:

$$\begin{aligned} P_e &= E_{\boldsymbol{X}}[P_e(\boldsymbol{X})] \\ &\leq \frac{1}{2}\left(\int_{-\frac{\Delta}{2}}^{\frac{\Delta}{2}} e^{-\frac{x^2}{8\sigma^2}} \frac{1}{\Delta} dx\right)^m \\ &= \frac{1}{2}\left(\frac{2\sqrt{2\pi}\sigma}{\Delta}\mathrm{erf}\left(\frac{\Delta}{4\sqrt{2}\sigma}\right)\right)^m . \end{aligned} \tag{5.8}$$

If $2^b = \mathcal{O}\left(\sqrt{m}\right)$, Δ is small, and thus we take the Taylor series expansion of $\mathrm{erf}\left(\frac{\Delta}{4\sqrt{2}\sigma}\right)$ at $\frac{\Delta}{4\sqrt{2}\sigma} = 0$:

$$
\begin{aligned}
P_e &\leq \frac{1}{2}\left(\frac{4\sqrt{2}\sigma}{\Delta}\left(\frac{\Delta}{4\sqrt{2}\sigma} - \frac{\Delta^3}{3(4\sqrt{2}\sigma)^2} + \frac{\Delta^5}{10(4\sqrt{2}\sigma)^5}\right)\right)^m \\
&= \frac{1}{2}\left(1 - \frac{\Delta^2}{3(4\sqrt{2}\sigma)^2} + \frac{\Delta^4}{10(4\sqrt{2}\sigma)^4}\right)^m .
\end{aligned}
$$

Note that at $\frac{\Delta}{4\sqrt{2}\sigma} = 0$:

$$
e^{-\frac{\Delta^2}{4(4\sqrt{2}\sigma)^2}} = 1 - \frac{\Delta^2}{4(4\sqrt{2}\sigma)^2} + \frac{\Delta^4}{32(4\sqrt{2}\sigma)^4} + \mathcal{O}(\Delta^5) .
$$

When Δ is small, the first and second terms are dominant. Thus, we have:

$$
P_e \leq \frac{1}{2} e^{-\frac{m\Delta^2}{4(4\sqrt{2}\sigma)^2}} + \mathcal{O}\left(\Delta^3\right) . \tag{5.9}
$$

For large m, we can ignore the error term $\mathcal{O}\left(\Delta^3\right)$. Then, if $b \leq \log_2 \frac{C\sqrt{m}}{8\sqrt{2}\sigma\sqrt{\ln\frac{1}{2\epsilon}}}$ for any $\epsilon > 0$, $P_e \leq \epsilon$ for any $\epsilon > 0$. Thus, $2^b = \mathcal{O}\left(\sqrt{m}\right)$ is sufficient for the observer to detect the signal. □

Now we consider the case that the system employs a quantizer with overflows; in particular, the quantizer has a span of $[-l, l]$, and we assume that the original signal $X_i \sim \mathcal{N}(0, \sigma_0^2), i = 1, 2, \ldots, m$. We provide the converse result in this case.

Theorem 5.4 (Converse under uniform quantization with overflows)**.** Suppose that Alice sends a discrete-time signal with length m and the quantizer is uniform with b bits of quantization levels. Then, if $b = \mathcal{O}\left(\sqrt{m}\right)$ or $l = o\left(\sqrt{\ln m}\right)$, the observer can distinguish the original signal and the quantized signal with arbitrarily low probability of error.

Proof. We derive the upper bound for the error probability P_e in the case that the quantizer has overflows. In this case, the quantization error U_i for $i = 1, 2, \ldots, m$ is given by (5.3). Following (5.7), we write:

$$
\begin{aligned}
P_e &= E_{\boldsymbol{X}}[P_e(\boldsymbol{X})] \\
&\leq \frac{1}{2}\Bigg[P\left(|X_i| < l\right) \cdot \int_{-\frac{\Delta}{2}}^{\frac{\Delta}{2}} e^{-\frac{x^2}{8\sigma^2}} \frac{1}{\Delta} dx + P\left(X_i > l\right) \cdot \int_0^{\infty} \frac{1}{\sqrt{2\pi\sigma_0^2}} e^{-\frac{(x+l)^2}{2\sigma_0^2}} e^{-\frac{x^2}{8\sigma^2}} dx \\
&\quad + P\left(X_i < -l\right) \cdot \int_{-\infty}^{0} \frac{1}{\sqrt{2\pi\sigma_0^2}} e^{-\frac{(x-l)^2}{2\sigma_0^2}} e^{-\frac{x^2}{8\sigma^2}} dx\Bigg]^m \\
&= \frac{1}{2}\Bigg[\underbrace{\operatorname{erf}\left(\frac{l}{\sqrt{2}\sigma_0}\right) \frac{2\sqrt{2\pi}\sigma}{\Delta} \operatorname{erf}\left(\frac{\Delta}{4\sqrt{2}\sigma}\right)}_{A} \\
&\quad + \underbrace{\frac{\left(1 - \operatorname{erf}\left(\frac{l}{\sqrt{2}\sigma_0}\right)\right) e^{-\frac{l^2}{8\sigma^2 + 2\sigma_0^2}} \left(1 - \operatorname{erf}\left(\frac{\sqrt{2}l}{\sqrt{\frac{\sigma_0^4}{\sigma^2} + 4\sigma_0^2}}\right)\right)}{\sqrt{\frac{\sigma_0^2}{\sigma^2} + 4}}}_{B}\Bigg]^m.
\end{aligned}
$$

If $2^b = \mathcal{O}\left(\sqrt{m}\right)$ and l is arbitrary, then by the discussion from (5.8) to (5.9), we have:

$$
\begin{aligned}
A^m &\leq \left(\frac{2\sqrt{2\pi}\sigma}{\Delta} \operatorname{erf}\left(\frac{\Delta}{4\sqrt{2}\sigma}\right)\right)^m \\
&\leq e^{-\frac{m\Delta^2}{4(4\sqrt{2}\sigma)^2}} + \mathcal{O}\left(\Delta^3\right)
\end{aligned}
\tag{5.10}
$$

which goes to zero as $m \to \infty$. On the other hand, recall that $l = \omega\left(1\right)$, then for any $a \geq 1$, $B^a \to 0$ as l becomes large. This can be seen by noting that if we ignore all of the constants in B and use the fact that $1 - \operatorname{erf}(x) < e^{-x^2}$ [72], we have $B < e^{-3l^2}$. Thus, if $2^b = \mathcal{O}\left(\sqrt{m}\right)$, $P_e \leq \epsilon$ for any $\epsilon > 0$.

If $l = o\left(\sqrt{\ln m}\right)$ and b is arbitrary, then:

$$A^m \leq \left(\operatorname{erf}\left(\frac{l}{\sqrt{2}\sigma_0}\right)\right)^m \to 0$$

as $m \to \infty$. Again, for any $a \geq 1$, we have $B^a \to 0$ as l becomes large. Thus, when $l = o\left(\sqrt{\ln m}\right)$, $P_e \leq \epsilon$ for any $\epsilon > 0$.

Therefore, if $2^b = \mathcal{O}\left(\sqrt{m}\right)$ or $l = o\left(\sqrt{\ln m}\right)$, the error probability P_e at the observer is arbitrarily small, which establishes the converse result. $\square$

5.6 Conclusion

In many applications such as the detection of a replay attack in network security or the detection of deception jamming in radar systems, it is important to know whether a received signal was sent directly, or was recorded and then replayed. Many approaches to this problem have been proposed in prior work; however, the fundamental limits of such detection with hardware imperfections have not been explored. Thus, we have studied this limit and analyzed the characteristics of the hardware, in particular the quantizer, that affect the detection. Specifically, if a signal with length m is sent and a uniform b-bit quantizer is employed, then $2^b = \omega\left(\sqrt{m}\right)$ and a quantizer span of $\omega\left(\sqrt{\ln m}\right)$ are sufficient to avoid detection; that is, the error probability at the observer is bounded as $P_e \geq \frac{1}{2} - \epsilon$ for any $\epsilon > 0$. Conversely, having $2^b = \mathcal{O}\left(\sqrt{m}\right)$ or a quantizer span of $o\left(\sqrt{\ln m}\right)$ results in detection by the observer with high probability as $m \to \infty$.

CHAPTER 6

CONCLUSION

This dissertation explored covert communications, where a transmitter Alice intends to communicate with a legitimate receiver Bob without being detected by a warden Willie. We focused on addressing critical aspects of moving towards implementation of covert communications.

Chapter 2 considered a power allocation problem for covert communications using a standard discrete-time model as a cornerstone to understand the underlying mechanisms of covert communications. In particular, with the information about the gain on the channel between Alice and Bob, Alice can adapt her transmit power to this gain to achieve a certain rate and meet the requirement on covertness such that Willie detects the presence of the transmission with low probability. We provided exact optimal power adaptation schemes in different scenarios that significantly outperform standard power adaptation schemes.

In Chapter 3, we studied covert communications in a true continuous-time model. To demonstrate that the power detector for Willie which is optimal in a discrete-time model may not be optimal in the continuous-time case, we provided and analysed an interference cancellation detector that outperforms the power detector. This shows that Willie can benefit from the continuous-time setting, which has a significant impact on the true covert throughput of the system. We then established novel constructions that allow Alice to achieve covert communications on the continuous-time model in two different cases: when there is perfect frame synchronization between Alice and the jammer, and when there is no such frame synchronization. We proved

that $\mathcal{O}(WT)$ covert information bits can be reliably transmitted from Alice to Bob in T seconds on a channel with asymptotic bandwidth W for both cases.

Since it is useful to exploit an interference source that already exists in the environment rather than a friendly jammer, Chapter 4 considers the case where Alice and Bob operating in an environment with an illuminating radar. We designed a communication waveform that embeds covert symbols in the radar signals and established a construction such that the messages are sent under the cover of the radar scattering. We showed that $\mathcal{O}(n)$ bits in n samples of the radar signal can be transmitted covertly and reliably from Alice to Bob.

Finally, in Chapter 5, we researched fundamental limits of the number of quantization bits and quantizer span of a quantizer that prevent or allow an observer to determine whether a signal has been recorded and then replayed. In particular, if a signal with length m is sent and a uniform b-bit quantizer is employed, then $2^b = \omega\left(\sqrt{m}\right)$ and a quantizer span of $\omega\left(\sqrt{\ln m}\right)$ are sufficient to avoid detection. Conversely, having $2^b = \mathcal{O}\left(\sqrt{m}\right)$ or a quantizer span of $o\left(\sqrt{\ln m}\right)$ results in detection by the observer with high probability as $m \to \infty$.

BIBLIOGRAPHY

[1] S. Chen, R. Wang, X. Wang and K. Zhang, "Side-Channel Leaks in Web Applications: A reality today, a challenge tomorrow," *2010 IEEE Symposium on Security and Privacy*, Berkeley/Oakland, CA, 2010, pp. 191-206.

[2] A. D. Wyner, "The wire-tap channel," in *The Bell System Technical Journal*, vol. 54, no. 8, pp. 1355-1387, Oct. 1975.

[3] I. Csiszar and J. Korner, "Broadcast channels with confidential messages," in *IEEE Transactions on Information Theory*, vol. 24, no. 3, pp. 339-348, May 1978.

[4] F. A. .P Petitcolas, R J. Anderson, Markus and G. Kuhn, "Information hiding – A survey," *Proceedings of the IEEE (special Issue)*, 1999, pp. 1062-1078.

[5] B. A. Bash, D. Goeckel and D. Towsley, "Limits of reliable communication with low probability of detection on AWGN channels," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 9, pp. 1921-1930, September 2013.

[6] B. A. Bash, D. Goeckel and D. Towsley, "Square root law for communication with low probability of detection on AWGN channels," *2012 IEEE International Symposium on Information Theory Proceedings*, Cambridge, MA, 2012, pp. 448-452.

[7] S. Lee and R. Baxley, "Achieving positive rate with undetectable communication over AWGN and Rayleigh channels," *2014 IEEE International Conference on Communications (ICC)*, Sydney, NSW, 2014, pp. 780-785.

[8] D. Goeckel, B. Bash, S. Guha and D. Towsley, "Covert communications when the warden does not know the background noise power," *IEEE Communications Letters*, vol. 20, no. 2, pp. 236-239, Feb. 2016.

[9] T. V. Sobers, B. A. Bash, S. Guha, D. Towsley and D. Goeckel, "Covert communication in the presence of an uninformed jammer," *IEEE Transactions on Wireless Communications*, vol. 16, no. 9, pp. 6193-6206, Sept. 2017.

[10] T. V. Sobers, PhD thesis, *Covert wireless communications in a dynamic environment*, University of Massachusetts Amherst, May 2017.

[11] T. V. Sobers, B. A. Bash, S. Guha, D. Towsley and D. Goeckel, "Covert communications on continuous-time channels in the presence of jamming," *51st Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, 2017, pp. 1697-1701.

[12] V. Korzhik, G. Morales-Luna and M. H. Lee, "On the existence of perfect stegosystems," in Proceedings of the 4th International Workshop on Digital Watermarking, Siena, Italy, Sep. 2005, pp. 30-38.

[13] P. H. Che, M. Bakshi and S. Jaggi, "Reliable deniable communication: hiding messages in noise," *2013 IEEE International Symposium on Information Theory*, Istanbul, 2013, pp. 2945-2949.

[14] P. H. Che, M. Bakshi, C. Chan and S. Jaggi, "Reliable deniable communication with channel uncertainty," *2014 IEEE Information Theory Workshop (ITW 2014)*, Hobart, TAS, 2014, pp. 30-34.

[15] L. Wang, G. W. Wornell and L. Zheng, "Limits of low-probability-of-detection communication over a discrete memoryless channel," *2015 IEEE International Symposium on Information Theory (ISIT)*, Hong Kong, 2015, pp. 2525-2529.

[16] L. Wang, G. W. Wornell and L. Zheng, "Fundamental Limits of Communication With Low Probability of Detection," in *IEEE Transactions on Information Theory*, vol. 62, no. 6, pp. 3493-3503, June 2016.

[17] L. Wang, "Optimal throughput for covert communication over a classical-quantum channel," *IEEE Information Theory Workshop (ITW)*, Cambridge, 2016, pp. 364-368.

[18] M. Bloch, "A channel resolvability perspective on stealth communications," 2015 IEEE International Symposium on Information Theory (ISIT), Hong Kong, 2015, pp. 2535-2539.

[19] M. R. Bloch, "Covert communication over noisy channels: a resolvability perspective," *IEEE Transactions on Information Theory*, vol. 62, no. 5, pp. 2334-2354, May 2016.

[20] M. Tahmasbi and M. R. Bloch, "Second-order asymptotics of covert communications over noisy channels," 2*IEEE International Symposium on Information Theory (ISIT)*, Barcelona, 2016, pp. 2224-2228.

[21] M. Tahmasbi and M. R. Bloch, "First- and second-order asymptotics in covert communication," *IEEE Transactions on Information Theory*, vol. 65, no. 4, pp. 2190-2212, April 2019.

[22] S. Lee, R. J. Baxley, J. B. McMahon and R. Scott Frazier, "Achieving positive rate with undetectable communication over MIMO Rayleigh channels," *2014 IEEE 8th Sensor Array and Multichannel Signal Processing Workshop (SAM)*, A Coruna, 2014, pp. 257-260.

[23] S. Lee, R. J. Baxley, M. A. Weitnauer, and B. Walkenhorst, "Achieving undetectable communication" *IEEE Journal of Selected Topics in Signal Processing*, vol. 9, no. 7, pp. 1195â€"1205, Oct. 2015.

[24] J. Hu, S. Yan, X. Zhou, F. Shu and J. Li, "Covert wireless communications with channel inversion power control in Rayleigh fading," *arXiv:1803.07812v2*, Aug 2018.

[25] S. Yan, B. He, X. Zhou, Y. Cong, A. L. Swindlehurst, "Delay-intolerant covert communications with either fixed or random transmit power", *IEEE Transactions on Information Forensic and Security*, vol. 14, no. 1, pp. 129-140, Jan. 2019.

[26] K. Shahzad, X. Zhou and S. Yan, "Covert communication in fading channels under channel uncertainty," *2017 IEEE 85th Vehicular Technology Conference (VTC Spring)*, Sydney, NSW, 2017, pp. 1-5.

[27] M. Forouzesh, P. Azmi, N. Mokari and D. Goeckel, "Robust power allocation in covert communication: imperfect CDI", *arXiv:1901.04914*, Jan. 2019.

[28] W. A. Gardner, *Cyclostationarity in Communications and Signal Processing*. New Jersey: IEEE Press, 1993.

[29] W. A. Gardner, "Signal interception: a unifying theoretical framework for feature detection," *IEEE Transactions on Communications*, vol. 36, pp. 897-906, Aug. 1988.

[30] W. Gardner, "Spectral correlation of modulated signals: Part I - analog modulation," *IEEE Transactions on Communications*, vol. 35, pp. 584-594, June 1987.

[31] W. Gardner, W. Brown, and C.-K. Chen, "Spectral correlation of modulated signals: Part II - digital modulation," *IEEE Transactions on Communications*, vol. 35, pp. 595-601, June 1987.

[32] K. Kim, I. A. Akbar, K. K. Bae, J. Um, C. M. Spooner and J. H. Reed, "Cyclostationary approaches to signal detection and classification in cognitive radio," 2007 2nd IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks, Dublin, 2007, pp. 212-215.

[33] W. Xiong, Y. Yao, X. Fu and S. Li, "Covert communication with cognitive jammer," *IEEE Wireless Communications Letters*, vol. 9, no. 10, pp. 1753-1757, 2020.

[34] S. D. Blunt, P. Yatham and J. Stiles, "Intrapulse radar-embedded communications," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 46, no. 3, pp. 1185-1200, July 2010.

[35] S. D. Blunt, J. G. Metcalf, C. R. Biggs and E. Perrins, "Performance characteristics and metrics for intra-pulse radar-embedded communication," *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 10, pp. 2057-2066, December 2011.

[36] Z. Hijaz and V. S. Frost, "Exploiting OFDM systems for covert communication," *IEEE Military Communications Conference*, pp. 2149-2155, 2010.

[37] G. Shabsigh and V. S. Frost, "Covert communications in wideband OFDMA primary networks," *IEEE Globecom Workshops*, 2015.

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

[39] H. ElSawy and E. Hossain, "On stochastic geometry modeling of cellular uplink transmission with truncated channel inversion power control," *IEEE Transactions on Wireless Communications*, vol. 13, no. 8, pp. 4454-4469, Aug. 2014.

[40] A. Limmanee, S. Dey, E. Nekouei, "Optimal power policies and throughput scaling analyses in fading cognitive broadcast channels with primary outage probability constraint", *EURASIP Journal on Wireless Communications and Networking*, vol. 2014, no. 35, pp. 1-18, 2014.

[41] T. Rakia, H. Yang, F. Gebali and M. Alouini, "Power adaptation based on truncated channel inversion for hybrid FSO/RF transmission with adaptive combining," *IEEE Photonics Journal*, vol. 7, no. 4, pp. 1-12, Aug. 2015, Art no. 7903012.

[42] C. Zenger, H. Vogt, J. Zimmer, A. Sezgin and C. Paar, "The passive eavesdropper affects my channel: secret-key rates under real-world conditions," *IEEE Globecom Workshops (GC Wkshps)*, Washington, DC, 2016, pp. 1-6.

[43] X. He, H. Dai, W. Shen, P. Ning and R. Dutta, "Toward proper guard zones for link signature," *Transactions on Wireless Communications*, vol. 15, no. 3, pp. 2104-2117, March 2016.

[44] R. Wilson, D. Tse and R. A. Scholtz, "Channel identification: secret sharing using reciprocity in ultrawideband channels," *IEEE Transactions on Information Forensics and Security*, vol. 2, no. 3, pp. 364-375, Sept. 2007.

[45] M. A. Khojastepour and B. Aazhang, "The capacity of average and peak power constrained fading channels with channel side information," *IEEE Wireless Communications and Networking Conference*, Atlanta, GA, 2004, pp. 77-82 Vol. 1.

[46] I. M. Gelfand and S. V. Fomin, *Calculus of variations*, Mineola, NY, USA: Dover Publications, 2000.

[47] K. S. K. Arumugam and M. R. Bloch, "Covert Communication Over a K -User Multiple-Access Channel," in *IEEE Transactions on Information Theory*, vol. 65, no. 11, pp. 7020-7044, Nov. 2019.

[48] J. Hu, S. Yan, X. Zhou, F. Shu and J. Wang, "Covert communications without channel state information at receiver in IoT systems," in *IEEE Internet of Things Journal (Early Access)*, May. 2020.

[49] Z. Liu, J. Liu, Y. Zeng and J. Ma, "Covert wireless communication in IoT network: from AWGN channel to THz band," *IEEE Internet of Things Journal*, vol. 7, issue. 4, pp. 3378-3388, Jan. 2020.

[50] Z. Liu, J. Liu, Y. Zeng and J. Ma, "Covert wireless communications in IoT systems: Hiding information in interference", *IEEE Wireless Communications*, vol. 25, no. 6, pp. 46-52, Dec. 2018.

[51] J. Hu, S. Yan, X. Zhou, F. Shu and J. Wang, "Covert communication in wireless relay networks," *IEEE Global Communications Conference*, Singapore, 2017, pp. 1-6.

[52] J. Hu, S. Yan, X. Zhou, F. Shu, J. Li, J. Wang, "Covert communication achieved by a greedy relay in wireless networks," *IEEE Trans. Wireless Commun.*, vol. 17, no. 7, pp. 4766-4779, May. 2018.

[53] J. Wang, W. Tang, Q. Zhu, X. Li, H. Rao and S. Li, "Covert communication with the help of relay and channel uncertainty," *IEEE Wireless Communications Letters*, vol. 8, no. 1, pp. 317-320, Feb. 2019.

[54] M. Forouzesh, P. Azmi, A. Kuhestani and P. L. Yeoh, "Covert communication and secure transmission over untrusted relaying networks in the presence of multiple Wardens," *IEEE Trans. Commun.*, vol. 68, no. 6, pp. 3737-3749, June 2020.

[55] K. Cho, S. Lee and V. Y. F. Tan, "Throughput scaling of covert communication over wireless Adhoc networks," *2019 IEEE International Symposium on Information Theory (ISIT)*, pp. 2164-2168, 2019.

[56] X. Zhou, S. Yan, J. Hu, J. Sun, J. Li and F. Shu, "Joint optimization of a UAV's trajectory and transmit power for covert communications," in *IEEE Transactions on Signal Processing*, vol. 67, no. 16, pp. 4276-4290, Aug. 2019.

[57] H. Wang, Y. Zhang, X. Zhang and Z. Li, "Secrecy and covert communications against UAV surveillance via multi-hop networks," *IEEE Transactions on Communications*, vol. 68, no. 1, pp. 389-401, Jan. 2020.

[58] Y. Jiang, L. Wang and H. Chen, "Covert communications in D2D underlaying cellular networks with antenna array assisted artificial noise transmission," *IEEE Trans. Veh. Techn.*, vol. 69, no. 3, pp. 2980-2992, March. 2020.

[59] Y. Jiang, L. Wang, H. Zhao and H. Chen, "Covert communications in D2D underlaying cellular networks with power domain NOMA," in *IEEE Systems Journal (Early Access)*, Feb. 2020.

[60] L. Wang, "On covert communication over infinite-bandwidth Gaussian channels," *IEEE 19th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, 2018, pp. 1-5.

[61] L. Wang, "On Gaussian covert communication in continuous time," *EURASIP Journal on Wireless Communications and Networking*, 2019, no. 1, pp. 283.

[62] L. Wang, "The continuous-time Poisson channel has infinite covert communication capacity," *2018 IEEE International Symposium on Information Theory (ISIT)*, Vail, CO, 2018, pp. 756-760.

[63] Q. E. Zhang, M. R. Bloch, M. Bakshi and S. Jaggi, "Undetectable radios: covert communication under spectral mask constraints," *2019 IEEE International Symposium on Information Theory (ISIT)*, Paris, France, 2019, pp. 992-996.

[64] U. Madhow, *Fundamentals of Digital Communication*, Cambridge, UK: Cambridge University Press, 2008.

[65] M. Chiani, D. Dardari, and M. K. Simon, "New exponential bounds and approximations for the computation of error probability in fading channels", *IEEE Transactions on Wireless Communications*, vol. 2, no. 4, pp. 840-845, 2003.

[66] M. Shaked and J. Shanthikumar, *Stochastic Orders and their Applications*, San Diego, CA, USA: Academic, 1994.

[67] Z. Zhu and M. B. Wakin, "On the asymptotic equivalence of circulant and Toeplitz matrices," in *IEEE Transactions on Information Theory*, vol. 63, no. 5, pp. 2975-2992, May 2017.

[68] J. Pearl, "On coding and filtering stationary signals by discrete Fourier transforms," *IEEE Trans. Inf. Theory*, vol. 19, no. 2, pp. 299-232, Mar. 1973.

[69] B. A. Bash, D. Goeckel and D. Towsley, "Asymptotic optimality of equal power allocation for linear estimation of WSS random processes," in *IEEE Wireless Communications Letters*, vol. 2, no. 3, pp. 247-250, June 2013.

[70] H. Arslan and K. Molnar, "Cochannel interference suppression with successive cancellation in narrow-band systems," *IEEE Communications Letters*, vol. 5, no. 2, pp. 37-39, Feb 2001.

[71] W. Qardaji and L. Ninghui, "Anonymizing network traces with temporal pseudonym consistency," *IEEE 32nd Int. Conf. on Distributed Computing Systems Workshops (ICDCSW)*, pp. 622-633, 2012.

[72] S. H. Chang, P. C. Cosman and L. B. Milstein, "Chernoff-type bounds for the gaussian error function," *IEEE trans. on commun.*, vol. 59, pp. 2939-2944, Nov. 2011.

[73] S. van de Beek and F. Leferink, "Vulnerability of remote keyless entry systems against pulsed electromagnetic interference and possible improvements," *IEEE Transactions on Electromagnetic Compatibility*, vol. 58, no. 4, pp. 1259-1265, Aug. 2016.

[74] A. Francillon, B. Danev and S. Capkun, "Resisting relay attacks on vehicular passive keyless entry and start systems," *2012 9th International Conference on Fuzzy Systems and Knowledge Discovery*, pp. 2232-2236.

[75] Z. Wu, S. Gao, E. S. Cling and H. Li, "A study on replay attack and antispoofing for text-dependent speaker verification," *Signal and Information Processing Association Annual Summit and Conference (APSIPA)*, pp. 1- 5, 2014.

[76] D. L. Adamy, *Electronic Warfare Moeling and Simulation*, Artech House, 2003.

[77] S. J. Roome, "Digital radio frequency memory," *Electronics & Communication Engineering Journal*, vol. 2, no. 4, pp. 147-153, Aug. 1990.

[78] C. Hanilçi, "Features and classifiers for replay spoofing attack detection," *2017 10th International Conference on Electrical and Electronics Engineering (ELECO)*, pp. 1187-1191, 2017.

[79] Y. Fu and Q. Fu, "Scheme and secure protocol of mobile payment based on RFID," *2009 3rd International Conference on Anti-counterfeiting, Security, and Identification in Communication*, pp. 631-634, 2009.

[80] M. Greco, F. Gini and A. Farina, "Radar Detection and Classification of Jamming Signals Belonging to a Cone Class," *IEEE Transactions on Signal Processing*, vol. 56, no. 5, pp. 1984-1993, May 2008.

[81] Y. Lu and S. Li, "CFAR detection of DRFM deception jamming based on singular spectrum analysis," *2017 IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC)*, pp. 1-6, 2017.

[82] A. C. Polak, S. Dolatshahi and D. L. Goeckel, "Identifying wireless users via transmitter imperfections," IEEE Journal on Selected Areas in Communications, vol. 29, no. 7, pp. 1469-1479, Aug. 2011.

[83] A. C. Polak and D. L. Goeckel, "Wireless device identification based on RF oscillator imperfections," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 12, pp. 2492-2501, Dec. 2015.

[84] A. C. Polak and D. L. Goeckel, "Identification of wireless devices of users who actively fake their RF fingerprints with artificial data distortion," *IEEE Transactions on Wireless Communications*, vol. 14, no. 11, pp. 5889-5899, Nov. 2015.

[85] E. Lehmann and J. Romano, *Testing Statistical Hypotheses*, 3rd ed. New York: Springer, 2005.