A CASE STUDY IN PHYSICAL-LAYER STEGANOGRAPHY APPLIED TO MULTICARRIER TRANSMISSIONS

by

Daniel William Chew

A dissertation submitted to Johns Hopkins University in conformity with the requirements for the degree of Doctor of Engineering

Baltimore, Maryland

July, 2022

Abstract

Covert communications can be a force for good, such as providing a means of message authentication to prevent malicious actors from spoofing networks. This dissertation explores the design of a covert signal to be hidden inside the bandwidth of an Orthogonal Frequency Division Multiplexing (OFDM) signal. In order to make detection by unintended observers as difficult as possible, the covert signal operates as interference inside the OFDM signal and is set to a high Signal to Interference Ratio (SIR). Given the high SIR, the OFDM signal must be cancelled in order to recover the covert signal. The detectability of the covert signal is tested using multiple detectors with and without cancellation. Among the detectors used is a Convolutional Neural Network (CNN) designed for image classification that has been repurposed through transfer learning to detect signal activity in noise and interference. The CNN detector demonstrates resilience in the presence of narrowband interference. The cancellation algorithm is enhanced with an estimate of OFDM windowing as applied at the transmitter, which is an often-overlooked parameter in cancellation applications. The enhanced cancellation-algorithm improves the cancellation of OFDM signals by 5.3 dB in an over-the-air test. The enhanced cancellation-algorithm also improves the Packet Error Rate of OFDM signals and improves the recovery of the covert signal. The improved recovery has direct application to Power-Domain Non-orthogonal Multiple Access and Rate-Splitting Multiple Access, which both rely on successive interference cancellation. Lastly, to frustrate any efforts to analyze the covert waveform, the covert signal

is augmented with an adversarial waveform designed to exploit weaknesses in CNNs used for modulation classification. The classification system suffers from uncertainty in the bandwidth estimate of the covert signal. The system will likely err on the side of making the bandwidth wider than necessary. It is demonstrated that a wider bandwidth makes the attack more successful, as opposed to other estimation errors which prior literature has shown to weaken the effectiveness of these attacks.

Thesis Readers

A. Brinton Cooper (Primary Advisor)
Associate Research Professor
Department of Electrical and Computer Engineering
Johns Hopkins Whiting School of Engineering

Chris Baumgart
Program Manager
Force Projection Sector
Johns Hopkins University Applied Physics Laboratory

Trac-Duy Tran
Professor
Department of Electrical and Computer Engineering
Johns Hopkins Whiting School of Engineering

Acknowledgments

I would like to acknowledge the people who helped make this work a reality.

I am grateful to my advisors Brint Cooper and Chris Baumgart for their support and guidance through the Whiting School's Doctor of Engineering program. We worked together through the COVID-19 pandemic, and through many other rough patches on this journey.

I would like to acknowledge those who worked with me on the research in this dissertation, Samuel Berhanu, Christine Nguyen, and Daniel Barcklow. We will always remember the *Whiskey Rodeo*. That is an inside joke.

I would like to thank Stephan Frisbie and Doug Frome who both reviewed early manuscripts of my publications. I appreciate their attention to detail and their willingness to help. I would also like to thank Milan Yagodich for reading through a draft this dissertation.

Last but certainly not least, I want to thank my family and friends who supported me through this process. Without them, this work could not have been possible.

Dedication

I dedicate this work to my wife Lleona and to my children Marin, Everett, and Theodore.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Covert Communications as a Force for Good

The term *covert communications* may elicit assumptions of nefarious purposes. It is true that covert communications can be used for exfiltration and other exploits; but these covert communications can also be a **force for good**. Two example applications for this type of signal are as a means of message authentication to prevent malicious actors from spoofing networks [43], and to provide metadata and emergency alerts [32]. As an example, consider the consequences of an eavesdropper intercepting a signal from a medical device [42]. Just the act of intercepting the signal may allow the eavesdropper to know that the patient has such a device and then extrapolate private information. Intercepting the signal may also allow the eavesdropper to track the patient. Encryption will not mitigate either of those risks. In this case, hiding such signals helps protect the patient's privacy and security. Low probability of detection (LPD) signals are useful to reduce the probability of a wireless signal being exploited, as they remove the ability of a hacker to intercept a signal in the first place. Promoting strong privacy and security is the goal of this research. In order to accomplish that, the research described herein focuses on advancing the state-of-the-art in covert communications. Physical-layer steganography represents the cutting edge of covert communications waveforms. Therefore this research

shall be a thorough case study into the development of a covert signal by which physical-layer steganography is applied to multicarrier transmissions.

1.2 Low Probability of Detection Signals

The concept of LPD communications can be explained as a wireless link between "Alice" and "Bob," where the two characters want to exclude "Eve the Eavesdropper" from detecting their communications signal. This scenario is illustrated in Fig. 1.1. When Alice transmits to Bob, her transmitter emits a signal. The goal is to make the waveform in such a way as to make Eve's detection of the signal improbable. However, it is also important that Alice's signal be recoverable by Bob. One way of achieving this is for Alice and Bob to share a secret such as a spreading code. Bob knows the spreading code, and can de-spread Alice's signal. Eve cannot de-spread Alice's signal and therefore must employ a wider bandwidth than does Bob. Alice and Bob hope this makes Eve employ a receiver with a wider bandwidth than Bob's receiver, making Eve's eavesdropping less probable.

Among the earliest publications explicitly to address such covert LPD waveforms are [36], [39] and [34]. In reference [36], the authors analyze the capability of a radiometer[1] to detect a spread spectrum signal, and it is established that noise uncertainty reduces the effectiveness of this detection method, especially against spread spectrum signals. LPD waveforms are defined in [39] as effecting a low detection probability by way of having a Signal-to-Noise Ratio (SNR) at an intercepting receiver that is lower than the intercepting receiver's Minimum Detectable Signal (MDS).

The research in [34] compares the probability of detection of the LPD waveforms to that of more conventional non-LPD waveforms, identifying common detection

[1] Also known as an energy detector

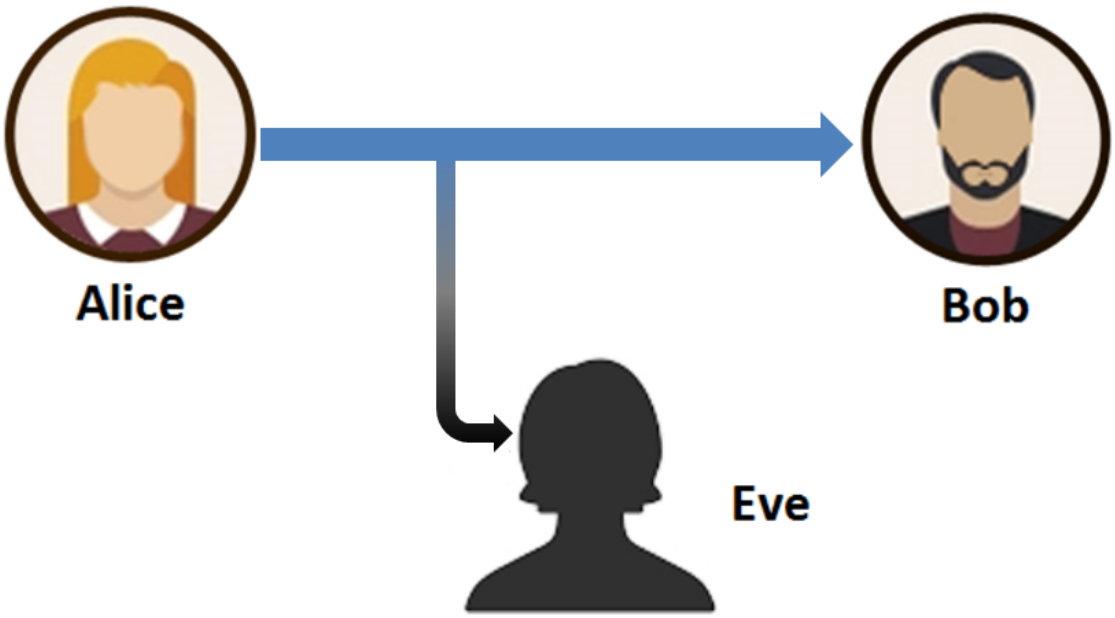

Figure 1.1: Alice, Bob, and Eve

techniques such as energy detectors, carrier regeneration, clock extraction, and exploits against unintended emissions from the transmitter. The latter, unintended emissions from the transmitter, is specific to flaws in the design of the transmitter. Carrier regeneration and clock extraction are both the results of mixing the signal with itself. For example, squaring a Binary Phase Shift Keying (BPSK) signal will result in a spike in the spectrum of the resulting product located at twice the carrier frequency.

1.3 Physical-Layer Steganography

The advent of ubiquitous wireless systems provides new opportunities in which to camouflage covert communications. The covert signal and the incumbent transmit concurrently and in the same bandwidth. The covert signal is thus conveyed as perturbations in the incumbent waveform. The covert signal described in this work can be called *physical-layer steganography* because it is hiding one waveform

inside another waveform. The concept is more generally referred to as "network steganography," as defined in [25], and the idea can be applied to any layer of the protocol stack of an incumbent signal.

Examples of steganography applied to the higher layers involve manipulating subsets of the payload bits. Another example of steganography applied to the Medium Access Layer involves modifying the timing of shared medium access in such a way as to convey a pulse-width modulated message [23]. The trouble with these methods is that they are of low data rate. Steganography applied to the physical layer offers a much higher data rate [13].

Applying steganography at the physical layer presents an advantage over the traditional LPD signal relying on noise at Eve's receiver. Being inside the incumbent offers significant concealment [13]. The features the traditional signal detectors rely upon, such as energy, are now extracted from the incumbent signal much more so than the covert signal. That the incumbent signal is the primary source of energy within the bandwidth will frustrate energy detection techniques. The feature-extraction techniques such as the carrier regeneration method in [34] will also expose the features of the incumbent signal more so than the covert signal.

1.4 Creating a Covert Signal using Physical-Layer Steganography

This dissertation explores covert communications where the covert signal is hidden inside the bandwidth of another transmission, the latter signal being termed the *incumbent*[2]. It is intended that the incumbent be an Orthogonal Frequency Division

[2]In terms common to steganography, the incumbent signal would be called the *cover* signal and the signal hidden inside of it would be the *covert* signal. The problem with that terminology is that *cover* and *covert* are only one letter apart. This work utilizes terminology common to the Cognitive Radio community where the term *incumbent* implies higher priority and/or ownership of the spectrum being used.

Multiplexed (OFDM) signal. OFDM is a multicarrier modulation scheme found in many signals used today, including but not limited to 4G cellular, 5G cellular, and IEEE 802.11 Wireless Local Area Networks (WLANs). OFDM signals were chosen as the target incumbent signal because they are ubiquitous and therefore offer ample opportunity for camouflage.

The intention is for this covert signal to provide a "sizeable throughput." Therefore, it is insufficient to simply lower the data rate of the covert link in order to recover the signal and reduce the impact on the OFDM signal. Novel means of signal recovery were explored. The throughput achieved was compared to the state of the art in the literature. This covert communications technique must also be feasible in a realistic setting; therefore experiments with over-the-air (OTA) data were performed.

1.5 Vision and Approach

There were several goals for the covert signal developed in this work:

- It is the responsibility of the covert signal to mask itself in the incumbent, no cooperation from the incumbent can be expected,
- The interference from the covert signal must not inflict a noticeable reduction in the throughput of the incumbent signal or else the covert signal may be exposed,
- The covert receiver does not have a copy of the incumbent signal in advance,
- The covert link must provide a sizeable throughput, and
- The covert transmitter must embed resistance in the covert signal to classification and reverse engineering in order to frustrate any exploitation efforts in the event of signal detection.

This list of goals guided the development of the covert waveform. For example, to reduce the detectability of the covert signal, and to reduce the impact of the covert signal on the incumbent, the covert signal operates at a high Signal-to-Interference Ratio (SIR). At a high SIR, the recovery of the covert signal required the development of novel cancellation methods in order to maintain a high throughput. Given that the OFDM incumbent does not cooperate with the covert signal, the covert receiver must model and estimate the OFDM incumbent sufficiently for cancellation purposes. The following sections detail the approach taken in developing the covert waveform and quantitatively measuring its performance.

1.5.1 Covert Communications through Imperfect Cancellation

Signal detection methods in academic literature were briefly discussed in section 1.2. The eavesdropper may employ more sophisticated methods, such as extracting cyclostationary features from the signal [37]. As a means of competing with more sophisticated detection techniques, the covert signal can be embedded within another signal as an alternative to hiding in AWGN [25]. Such a method for embedding a covert signal into an OFDM signal was developed in this work. The covert signal was designed as a Direct-Sequence Spread Spectrum (DSSS) signal and injected into the OFDM signal while complying with the spectral mask specified by the standard. The covert signal operates as interference inside the OFDM signal. The Packet Error Rate (PER) of the OFDM signal was measured as a function of the Signal to Interference Ratio (SIR) in order to quantify the impact of the covert signal on the incumbent. The Bit Error Rate (BER) of the covert signal was measured as a function of SIR. In order to make detection by unintended observers as difficult as possible, a high SIR was used. Given the high SIR, it was determined that de-spreading alone was insufficient to recover the covert signal. To overcome this, the covert signal was extracted by cancellation of the OFDM signal in the covert

receiver. The detectability of the covert signal was tested using an energy detector and a cyclostationary detector. The cyclostationary detector was based on the Fast Fourier Transform (FFT) Accumulation Method which provides an estimate of the spectral correlation function [33] [3]. That estimate is used to produce a cycle-frequency domain profile from which features are extracted for signal detection [22]. The energy detector and cyclostationary detector were both tested with and without the benefit of their own OFDM cancellation capability and it was found that the detectors needed to use cancellation. The recovery of the covert signal was tested with OTA recordings of 802.11 packets and demonstrated the effectiveness of the technique using that real-world OTA data. This work has been accepted to IH&MMSec 2022 [7].

1.5.2 Spectrum Sensing in Interference and Noise Using Deep Learning

In this experiment, spectrum sensing was transformed into image recognition by taking the spectrogram [3] of the received data and processing that as an image. A well-known Convolutional Neural Network (CNN) for image classification, AlexNet [24], was repurposed as a signal detector through the process of "fine-tuning" [18]. By way of fine-tuning, AlexNet was retrained using a small training set of a few hundred samples. AlexNet was chosen for this task due to its demonstrated resilience to Gaussian noise [19]. The performance of the new detector was compared to the performance of more traditional energy detection. The new detector surpassed the performance of the energy detector in the presence of co-channel interference and did not require a separate noise estimate. This work was published in CISS 2020 [5]. The resilience of the new detector to co-channel interference made it a candidate detector for secondary signaling such as the covert signal developed

[3] A spectrogram is a two dimensional frequency vs time image of a signal.

in this dissertation. Another experiment was performed in which the new detector was trained in cancellation residue. The new detector was then tested against the covert signal. The results were compared to those from the energy detector and cyclostationary detector. Like the other detectors, the new detector was ineffective without its own cancellation capability. With a cancellation capability, the new detector performed well by comparison to the other two. This novel method can be used in any application in which signals must be detected in the presence of interference.

1.5.3 OFDM Window Estimation for Interference Cancellation

Windowing at the transmitter is a popular means to control the bandwidth of an OFDM signal because it is cheaper than filtering [15]. OFDM windowing at the transmitter is typically implemented by extending the cyclic prefix into the previous OFDM symbol and creating a cyclic suffix that is extended into the next OFDM symbol. These cyclic extensions are then tapered by multiplying their samples with the window coefficients. The effect is a tapered transition from one OFDM symbol into another. The problem with this windowing is that the extension into adjacent symbols is self-interference. This work demonstrated that the windowing applied at the transmitter has a significant impact on the ability to cancel that OFDM signal. In this work, windowing was added to the signal model used to cancel the OFDM signal. The window was estimated from received samples, and that window estimate was used to cancel the OFDM signal without prior knowledge of the windowing function. The suppression of the OFDM signal was measured with and without the window estimation improvement using both synthetic and OTA data. It was found that the window estimate offered a 5.3 dB improvement with the OTA data. The cancelling was also selectively applied only to the self-interference and improved the PER of the OTA OFDM signal. The improved cancelling algorithm

was then used to improve the recovery of the covert signal. This work is of active patent interest and a provisional patent has been filed [4]. This work is under review for publication.

1.5.4 Adversarial Attacks on Deep-Learning RF Classification in Spectrum Monitoring with Imperfect Bandwidth Estimation

In the event that the covert signal is detected, the likely next step of the eavesdropper would be to attempt to classify the modulation of the covert signal. CNNs are often used as automated modulation classifiers [30]. The covert signal was augmented with an adversarial waveform designed to exploit weaknesses in CNNs used for automated modulation classification in order to hamper any efforts to analyze the covert signal. An adversarial waveform is a small additive perturbation at the transmitter, and is generated similarly to adversarial examples used against image classifiers [38]. An adversarial waveform was developed against one CNN and then deployed against another CNN thus transferring the attack. This transfer of the attack was done without knowledge of the second CNN. This technique can be used to develop an attack against modulation classification CNNs and then deploy that attack against 3rd party spectrum monitors. The adversarial waveform was created by constraining the signal-to-interference ratio at the transmitter, which has the dual benefits of making the adversarial waveform easy to deploy and mitigating impairment to the communications link. The adversarial waveform does introduce interference in the covert link; however, the matched filter in the covert receiver filters out the majority of the power of the adversarial waveform, thus the addition of the adversarial waveform has only a small impact on the covert signal BER. Testing demonstrated that the vulnerability of a CNN classifier to this type of attack was in part a function of bandwidth uncertainty, where the classifier does not have an exact estimate of the bandwidth of the covert signal. The classification system

will likely err on the side of making the bandwidth wider than necessary, because to do otherwise would filter out needed information. It was demonstrated that a wider bandwidth makes the attack more successful, as opposed to other estimation errors which prior literature has shown to weaken the effectiveness of these attacks. A small over-estimation of signal bandwidth provides a significant increase to the effectiveness of the attack. This work has been accepted to WCNC 2022 [6].

1.5.5 Applications Outside of Covert Communications

This covert signal employs techniques found in other applications such as Paired Carrier Multiple Access (PCMA) for satellite communications [11], Message Authentication [43], RF watermarking in television transmission [32], Power-Domain Non-Orthogonal Multiple Access (PD-NOMA) [40] and Rate Splitting Multiple Access (RSMA) [8]. In each of these applications there signals are transmitted within the same *spectrum resource*, and a means of signal cancellation is required. A spectrum resource is defined by a period of time, a band of frequencies (a frequency channel), and often a geographic location. The sharing of spectrum resources is an active area of research. The techniques developed in this dissertation can be extended to those other areas.

1.6 Outline of this Work

Each chapter in this dissertation documents research conducted to explore an aspect or problem related to covert communications in an interference channel.

- Chapter 2, "Covert Communications through Imperfect Cancellation", details the design and operation of a covert signal deployed inside an OFDM incumbent. An algorithm for extracting the covert signal is developed. The chapter also provides measurements of the performance of the covert signal

as a means of communications and the performance of the OFDM incumbent in the presence of the covert signal. The sum of the work in this chapter was published in [7].

- Chapter 3, "Image Classification for Signal Detection", describes a novel signal-detection method in which an image-recognition deep-learning network is retrained and re-purposed for signal detection. The sum of the work in this chapter was published in [5].

- Chapter 4, "OFDM Window Estimation for Interference Cancellation", details the benefits of estimating the windowing function used at the transmitter by the OFDM incumbent. This parameter has a significant impact on the total cancellation of the incumbent. In addition to the impact on cancellation, it was discovered that the cancellation algorithm can be used to remove the self-interference caused by this OFDM windowing parameter and improve the Packet Error Rate (PER) of the OFDM incumbent itself.

- Chapter 5. "Exploiting Vulnerabilities Deep-Learning RF Classification using an Interference Signal", details the use of a secondary signal to frustrate the classification efforts of a spectrum monitor. The sum of the work in this chapter was published in [6].

- Chapter 6 provides summarizes the preceding topics and discusses future work.

The remainder of this introductory chapter provides background material and describes concepts important to the research documented in this dissertation

1.7 Cognitive Radio Terminology

Cognitive radio terminology, as it relates to the topic of communications in an interference channel, describes the operation of a secondary user attempting to gain access to the spectrum. In order to understand this concept, the idea of a primary and secondary user must first be defined. *Primary users* are defined by IEEE 1900.1 [20] as "Users with higher priority or legacy rights on the usage of a particular spectrum frequency band." This means that a primary user has the right to transmit on a given channel. The users in the channel who do not meet this criteria are *secondary users*. It is the responsibility of the secondary user to keep its interference on the primary user within set bounds. How those bounds are defined depends on numerous factors including the method of secondary access.

Primary users are called incumbent users in some literature, such as TV Whitespace applications where a broadcast station, or another user designated by the spectrum regulatory authority, has primary rights to a frequency channel. Broadcast stations and their exclusive use of a bandwidth are not the only examples of primary users.

The covert signal described in this work would be a secondary user by this definition. The covert signal described in this work would be classified as making use of "spectrum underlay" by the IEEE 1900.1 standard. The standard defines spectrum underlay to include any case where a secondary user concurrently transmits on the same spectrum resource as a primary user, and the underlaying secondary user keeps the resulting interference within established tolerable bounds. The concept of underlaying is illustrated in Fig. 1.2 from [14]. The abbreviation "IT" stands for *Interference Temperature*. This is to say that the secondary transmission must be kept below a specific power to avoid interfering with the primary signal beyond that which the wireless link of the primary user can tolerate. This is contrasted with

"spectrum overlay", also defined in the IEEE 1900.1 standard, in which a secondary user makes opportunistic use of spectrum whitespace. Whitespace is an unused spectrum resource in both frequency and time. Given a specific frequency channel, whitespace would be a period of time during which a primary user is not using that frequency channel. The secondary user senses the spectrum for these white spaces and then transmits when and where the primary user is not present. This way, the secondary user attempts to avoid inflicting any interference on the primary user.

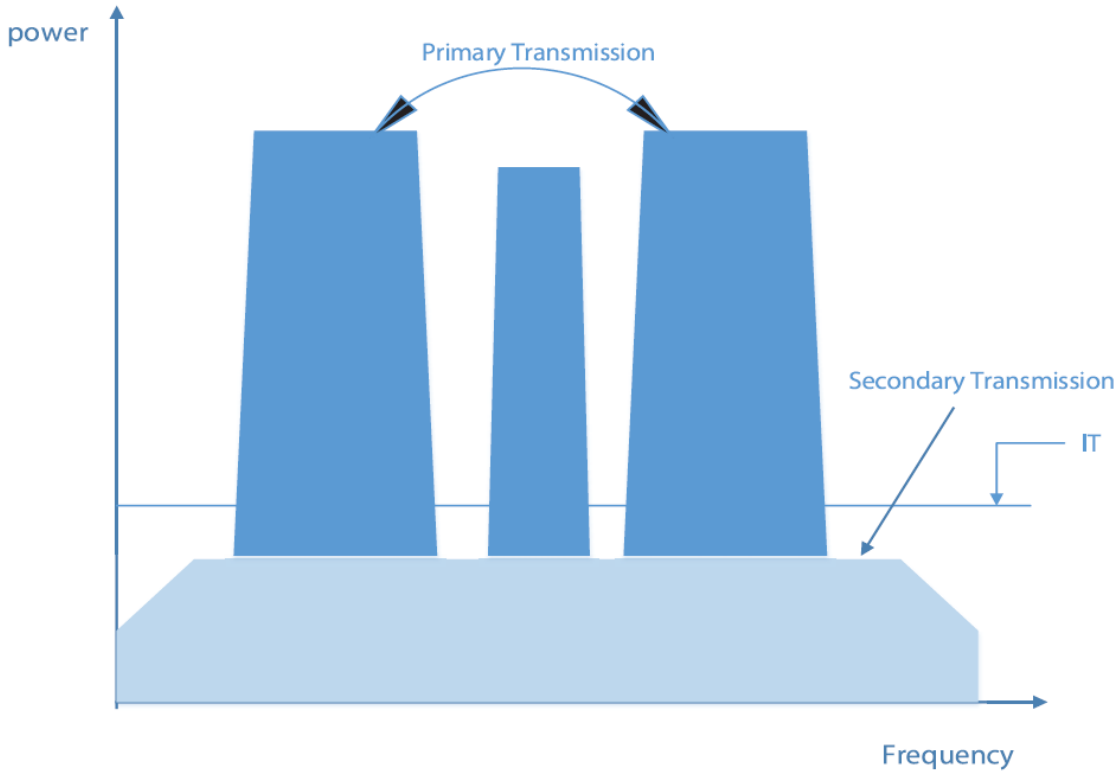

Figure 1.2: Underlaying [14] ©2017 IEEE

The terminology defined in IEEE 1900.1 is not universal in literature. A separate but related taxonomy is found in [17] that defines three categories of spectrum re-use: overlay, underlay and interweaving. In that taxonomy, underlay specifically refers to a secondary user occupying the same spectrum resource as a primary user, and therefore interferes with that primary user to some tolerable amount. In [17],

the term “overlay” refers to a secondary user (called the cognitive user) relaying the primary user’s signal within the secondary user’s signal, thus overlaying the two signals. The term “interweaved” is used in [17] to describe what the IEEE 1900.1 standard calls “overlaying”.

1.8 Interference Cancellation

Multi-User Interference Cancellation (MUIC) is a well-known concept that enables wireless users to address interference caused by multiple users accessing a single spectrum resource concurrently. MUIC is the process of modeling and reproducing a signal from one particular user for the purposes of removing that signal from the summed ensemble of all received signals. Successive Interference Cancellation (SIC) [28] is a category of MUIC implementations in which cancellation is applied sequentially over multiple users in the ensemble. Such cancellation is a topic explored for a variety of applications, among those applications is PD-NOMA. PD-NOMA is a multiple access scheme that accepts interference as part of the channel. Unlike other multiple access schemes that attempt to separate different emitters in frequency, time, and/or location, PD-NOMA requires the receiver to be able to remove unwanted signals.

A detailed description and argument in favor of PD-NOMA can be found in [40]. A summary is provided here: In a conventional system, the two users must take exclusive turns. The achievable overall throughput for the two users has a linear relationship with the number of turns allocated to each user. Hence, for every time-frequency resource that user 1 yields to user 2, user 2 gains a specific amount of throughput and user 1 loses a specific amount of throughput. Unlike the conventional scheme, in PD-NOMA signals for the two users are transmitted on the single spectrum resource concurrently. The signal for user 2 is transmitted

at a higher power than user 1. The interference caused by the lower-power signal for user 1 therefore factors into the achievable throughput for user 2. User 1 employs signal cancellation to remove the signal for user 2, and then access their own intended signal once the signal for user 2 has been removed. Both users can therefore enjoy a higher simultaneous data rate than they would have if they took exclusive turns accessing the spectrum resource.

The *capacity* of a channel is the maximum rate at which information can be reliably transmitted over that communication channel, as defined in [21] and detailed in [2]. There are many different types of channels. One type of channel of interest to this research is the bandlimited Additive White Gaussian Noise (AWGN) channel the capacity of which is expressed in (1.1), and a detailed analysis can be found in [16]. The capacity in (1.1) is measured in bits/second, *gamma* is the SNR, and B is the bandwidth of the channel.

$$C(\gamma) = Blog_2(1+\gamma) \tag{1.1}$$

Now consider a situation where a single broadcaster wants to send independent messages to two different users. The two users are thus sharing one spectrum resource. That channel is a Gaussian Broadcast channel as defined and analyzed in [29]. This scenario is also discussed in [40] which advocates for PD-NOMA. A summary is provided here: The Gaussian Broadcast channel has a *capacity region* that is defined as the closure of the set of achievable rates for all users. For a multiple access channel with only two users, that set of rates can be called a *rate pair*, (R_1, R_2) where R_1 is the achievable data rate for user 1 and R_2 is the achievable data rate for user 2.

Two different cases of the Gaussian Broadcast channel are considered. The first case is where the users are given exclusive access to the spectrum resource. That

is to say, in traditional systems, a transmitter must stop using spectrum resources to transmit the signal for user 1 if it is going to transmit a signal for user 2. This case is called "Orthogonal Multiple Access" in [40]. Let the variable α represent the share of the spectrum resource available to user 1, and $1-\alpha$ represent the share available to user 2. The parameter α ranges from 0 to 1, where at 0 only user 2 gets the spectrum resource, and at 1 only user 1 gets the spectrum resource. γ_i represents the SNR for a given user i and is defined in (1.4) where h_i represents the channel loss for a given user i, σ_i^2 is the noise variance at the receiver, and P is the transmit power of the broadcaster. The possible data rates over this channel with exclusive access to the spectrum resource are expressed in (1.2) and (1.3). Fig. 1.3 illustrates the capacity region of such a system where the bandwidth is 20 MHz, user 1 enjoys an SNR of 25 dB but user 2 has an SNR of -6 dB.

$$R_1 = \alpha C(\gamma_1) \tag{1.2}$$

$$R_2 = (1-\alpha) C(\gamma_2) \tag{1.3}$$

$$\gamma_i = P|h_i|^2/\sigma_i^2 \tag{1.4}$$

For a PD-NOMA system, the achievable rates are expressed in (1.5) and (1.6). The variables h_1 and h_2 represent the coefficients of the propagation channel for each user. The variables σ_1 and σ_2 represents the noise at the receiver of each user. The broadcaster always transmits at power P. In the case of PD-NOMA, α represents the share of the transmit power of the base station dedicated to each of the two user signals. In this scenario, the bandwidth is 20 MHz, user 1 enjoys an SNR of 25 dB but user 2 has an SNR of -6 dB. Because user 1 has the higher SNR, user 1 will apply SIC. User 2 does not perform SIC in this scenario. Therefore, the signal for user 1

counts as noise in the calculation of the data rate for user 2. Note that this scenario assumes that user 2 will have the ability to cancel the signal from user 1 completely. Fig. 1.3 illustrates how the achievable rates in a PD-NOMA system are higher than that of the more traditional exclusive access model.

$$R_1 = C(\alpha P|h_1|^2/\sigma_1^2) \tag{1.5}$$

$$R_2 = C(\frac{(1-\alpha)P|h_2|^2}{\sigma_2^2 + \alpha P|h_2|^2}) \tag{1.6}$$

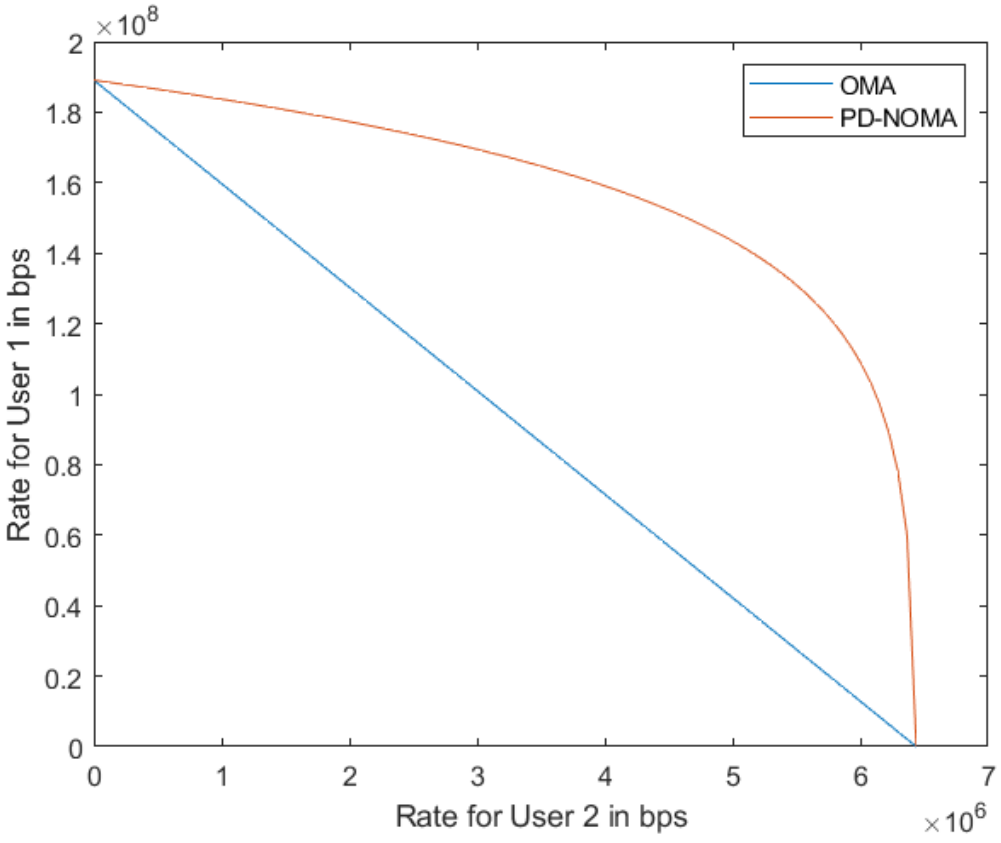

Figure 1.3: Capacity Region for PD-NOMA, Overlaid with the Capacity Region for OMA

MUIC schemes have been proposed for numerous wireless systems, e.g., MUIC and PD-NOMA have been proposed for 5G New Radio [10] [12] [1]. MUIC has

also been proposed in RSMA [8] which is a multiple access scheme that has been projected to outperform PD-NOMA [27]. Given that the waveform being used in these cases is OFDM, the question arises as to how well one can cancel another user's signal when using OFDM waveforms.

1.9 Complications in Interference Cancellation

The data rate equations in section 1.8 made several assumptions. That analysis assumed that the interference caused by user 1 on user 2 is the same as additional AWGN. In a single-carrier communication scheme, it is unlikely that the distribution of the single-carrier signal is Gaussian or that the samples are independent. In a multicarrier scheme like OFDM, one may be able to approximate the distribution of the signal as Gaussian by way of the central limit theorem. The passband of the signal is broad in the spectrum relative to the sample rate and very flat, thus simulating the spectrum of AWGN, even though the samples of the OFDM signal are not strictly independent.

The analysis in section 1.8 also assumes the perfect cancellation of user 2 from the bandwidth of user 1. The use of MUIC requires that the receiver be able to cancel a portion of the received signal. That assumes that the receiver has a sufficient model of the signal and can estimate the parameters of that model with sufficient accuracy. All the models begin with demodulating the signal and then remodulating. The remodulated signal is then augmented with estimated impairments for a better match at cancellation. This process is described in [28] but it does not define the parameters to be estimated. The signal model employed for cancellation in [9] estimates the channel coefficients. The signal model in [31] employs estimates for channel coefficients and "inter-channel interference" (ICI) meaning wireless impairments have caused the subcarriers of the OFDM symbol to no longer be

orthogonal. The signal model in [26] estimates amplitude and phase. These signal models for cancellation may be insufficient. There are signal model parameters that are often overlooked in MUIC and PD-NOMA literature.

As explained in [35], hardware imperfections are among the overlooked parameters in MUIC and PD-NOMA literature. It is common for wireless standards to limit the transmitter implementation in terms of transmitter error vector magnitude and frequency accuracy, and this allows significant room for individual vendors to employ hardware with various impairments including but not limited to I/Q imbalance and power amplifier nonlinearity. The degradation in the Bit Error Rate Curve of a radio node, U2, required to cancel the signal for another node, U1 is shown in Fig.1.4 from [35] and illustrates the effect on bit error rate of two different radio hardware impairments. Those two impairments are a non-linear power amplifier at the transmitter and I/Q imbalance at the receiver. The BER curves for two radio nodes, U1 and U2 are shown. The radio node U2 must cancel the signal for radio node U1 in this example. The SNR from AWGN for both nodes, not counting Multi-User Interference, is 25 dB. The independent axis is the power splitting parameter, "a", which is the same as α in Fig. 1.3.

In the OTA experiments performed in this work, it was found that the inclusion of OFDM windowing at the transmitter into the signal model provided a 5.3 dB improvement to the cancellation of the incumbent OFDM signal. The benefits of windowing at the transmitter are explored in detail in [15]. OFDM systems often employ windowing at the transmitter in order to meet the spectral mask imposed by a given wireless standard. The IEEE 802.11 standard suggests such a window but does not mandate it. The use, shape, and length of such a window are left to individual vendors. Fig. 1.5 illustrates windowing applied at the transmitter and the receiver. Windowing at the transmitter is applied in order to smooth the transitions between OFDM symbols and thus limit the bandwidth of the OFDM signal. The

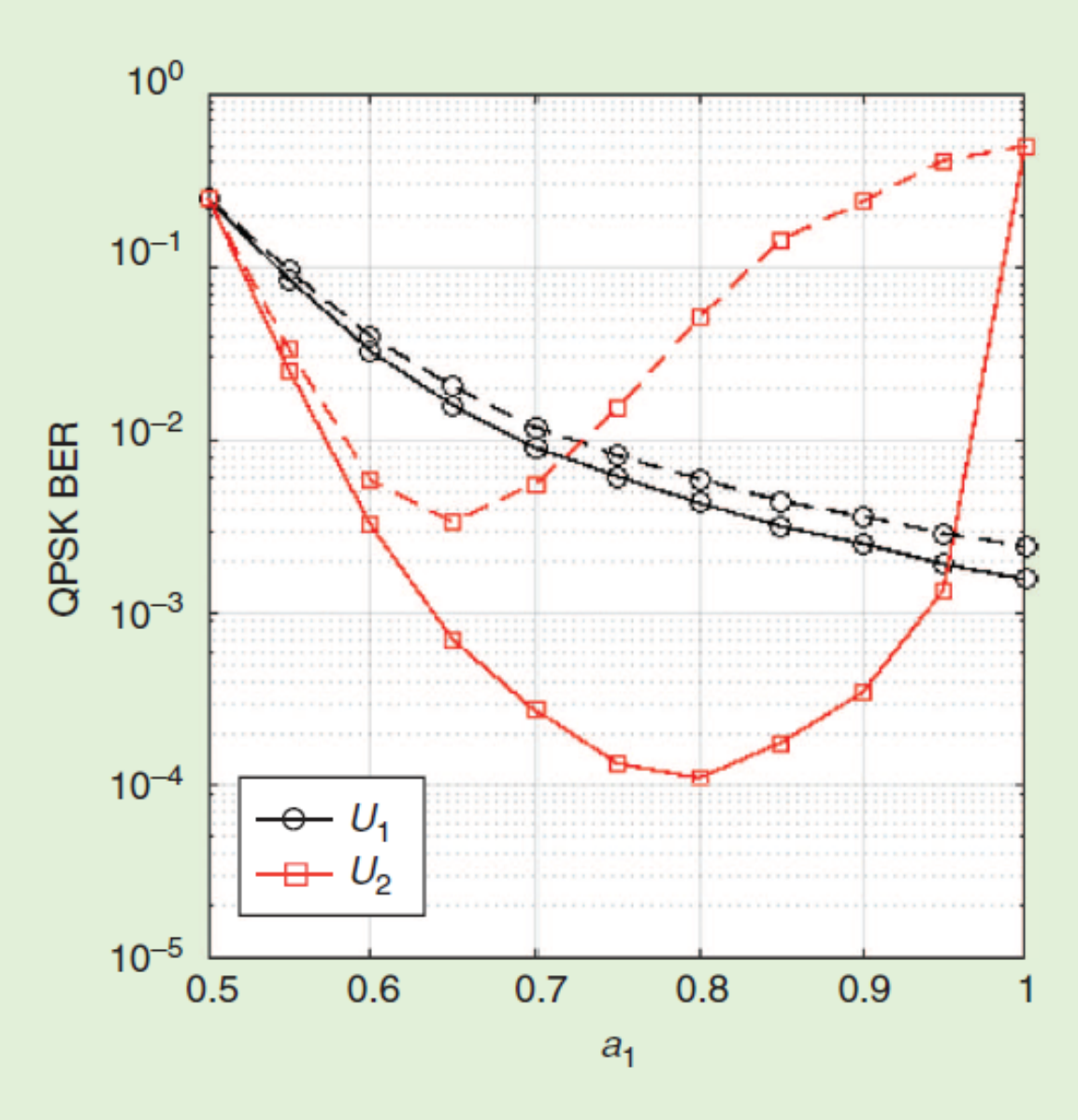

Figure 1.4: Degradation of the BER Curve in PD-NOMA with Radio Hardware Impairments [35] ©2019 IEEE

period T in Fig. 1.5 represents the total OFDM symbol period including the FFT period and the period for the required cyclic prefix. That OFDM symbol is extended by $0.5T_0$ into both adjacent symbols, and both adjacent symbol are extended by $0.5T_0$ into the current OFDM symbol. OFDM Windowing at the receiver is also illustrated in Fig. intro:fig:ofdmwindrxtx from [15]. OFDM windowing at the

receiver serves different purposes, has no impact on the transmitted signal, is not useful for cancellation, and is therefore not germane to this discussion.

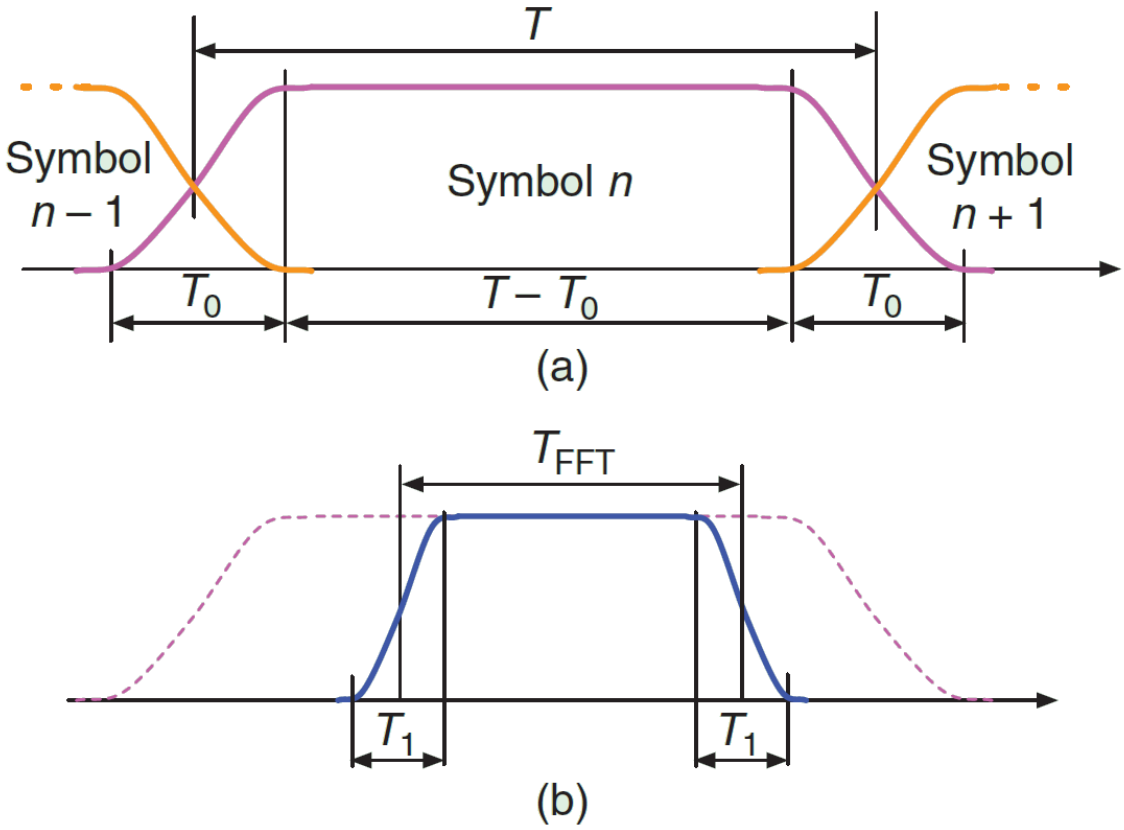

Figure 1.5: OFDM Windowing Transmit and Receive [15] ©2011 IEEE

OFDM windowing at the transmitter deliberately introduces inter-symbol interference between adjacent OFDM symbols. The effects of this self-interference on 802.11a/g OFDM signal were documented in [41]. Individual vendors are free to apply what windows they choose at the transmitter, but they must still meet the modulation accuracy requirements in the standard.

This work will show that the transmitter-windowing parameter in OFDM has a significant effect on MUIC performance. The estimation of the shape and length of the OFDM window is detailed in chapter 4.

1.10 References

[1] Ishan Budhiraja et al. "A Systematic Review on NOMA Variants for 5G and Beyond". In: *IEEE Access* 9 (2021), pp. 85573–85644. DOI: 10.1109/ACCESS.2021.3081601.

[2] "Channel Capacity". In: *Elements of Information Theory*. John Wiley & Sons, Ltd, 2005. Chap. 7, pp. 183–241. ISBN: 9780471748823. DOI: https://doi.org/10.1002/047174882X.ch7. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/047174882X.ch7. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/047174882X.ch7.

[3] Daniel Chew, Andrew L. Adams, and Jason Uher. "Secondary Spectrum Usage and Signal Detection". In: *Wireless Coexistence: Standards, Challenges, and Intelligent Solutions*. 2021, pp. 115–154. DOI: 10.1002/9781119584230.ch5.

[4] Daniel Chew, Chris Baumgart, and A. Brinton Cooper. *OFDM PER Improvement through Self-Interference Cancellation*. Tech. rep. JHU/APL File No. 6339-SPL. Baltimore, MD: Johns Hopkins University, 2021.

[5] Daniel Chew and A. Brinton Cooper. "Spectrum Sensing in Interference and Noise Using Deep Learning". In: *2020 54th Annual Conference on Information Sciences and Systems (CISS)*. 2020, pp. 1–6. DOI: 10.1109/CISS48834.2020.1570617443.

[6] Daniel Chew et al. "Adversarial Attacks on Deep-Learning RF Classification in Spectrum Monitoring with Imperfect Bandwidth Estimation". In: *Accepted to 2022 IEEE Wireless Communications and Networking Conference (WCNC)*. 2022. DOI: 10.1109/CISS48834.2020.1570617443.

[7] Daniel Chew et al. "Covert Communications through Imperfect Cancellation". In: *Accepted to Proceedings of the 2022 ACM Workshop on Information Hiding and Multimedia Security*. IH&MMSec '22. New York, NY, USA: Association for Computing Machinery, 2022.

[8] Bruno Clerckx et al. "Rate splitting for MIMO wireless networks: a promising PHY-layer strategy for LTE evolution". In: *IEEE Communications Magazine* 54.5 (2016), pp. 98–105. DOI: 10.1109/MCOM.2016.7470942.

[9] Bruno Clerckx et al. "Rate-Splitting Unifying SDMA, OMA, NOMA, and Multicasting in MISO Broadcast Channel: A Simple Two-User Rate Analysis". In: *IEEE Wireless Communications Letters* 9.3 (2020), pp. 349–353. DOI: 10.1109/LWC.2019.2954518.

[10] Linglong Dai et al. "Non-orthogonal multiple access for 5G: solutions, challenges, opportunities, and future research trends". In: *IEEE Communications Magazine* 53.9 (2015), pp. 74–81. DOI: 10.1109/MCOM.2015.7263349.

[11] Mark Dankberg. "Paired carrier multiple access (PCMA) for satellite communications". In: *17th AIAA International Communications Satellite Systems Conference and Exhibit*. 1998, p. 1398.

[12] Zhiguo Ding et al. "Application of Non-Orthogonal Multiple Access in LTE and 5G Networks". In: *IEEE Communications Magazine* 55.2 (2017), pp. 185–191. DOI: 10.1109/MCOM.2017.1500657CM.

[13] Aveek Dutta et al. "Secret Agent Radio: Covert Communication through Dirty Constellations". In: *Information Hiding*. Ed. by Matthias Kirchner and Dipak Ghosal. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 160–175. ISBN: 978-3-642-36373-3.

[14] Manal El Tanab and Walaa Hamouda. "Resource Allocation for Underlay Cognitive Radio Networks: A Survey". In: *IEEE Communications Surveys Tutorials* 19.2 (2017), pp. 1249–1276. DOI: 10.1109/COMST.2016.2631079.

[15] Behrouz Farhang-Boroujeny. "OFDM Versus Filter Bank Multicarrier". In: *IEEE Signal Processing Magazine* 28.3 (2011), pp. 92–112. DOI: 10.1109/MSP.2011.940267.

[16] "Gaussian Channel". In: *Elements of Information Theory*. John Wiley & Sons, Ltd, 2005. Chap. 9, pp. 261–299. ISBN: 9780471748823. DOI: https://doi.org/10.1002/047174882X.ch9. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/047174882X.ch9. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/047174882X.ch9.

[17] Andrea Goldsmith et al. "Breaking Spectrum Gridlock With Cognitive Radios: An Information Theoretic Perspective". In: *Proceedings of the IEEE* 97.5 (2009), pp. 894–914. DOI: 10.1109/JPROC.2009.2015717.

[18] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. MIT Press, 2016.

[19] K. Grm et al. "Strengths and weaknesses of deep learning models for face recognition against image degradations". In: *IET Biometrics* 7.1 (2018), pp. 81–89.

[20] "IEEE Standard for Definitions and Concepts for Dynamic Spectrum Access: Terminology Relating to Emerging Wireless Networks, System Functionality, and Spectrum Management". In: *IEEE Std 1900.1-2019 (Revision of IEEE Std 1900.1-2008)* (2019), pp. 1–78. DOI: 10.1109/IEEESTD.2019.8694195.

[21] "Introduction and Preview". In: *Elements of Information Theory*. John Wiley & Sons, Ltd, 2005. Chap. 1, pp. 1–12. ISBN: 9780471748823. DOI: https://doi.org/10.1002/047174882X.ch1. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/047174882X.ch1. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/047174882X.ch1.

[22] Kyouwoong Kim et al. “Cyclostationary Approaches to Signal Detection and Classification in Cognitive Radio”. In: *2007 2nd IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks*. 2007, pp. 212–215. DOI: 10.1109/DYSPAN.2007.35.

[23] Negar Kiyavash et al. “A Timing Channel Spyware for the CSMA/CA Protocol”. In: *IEEE Transactions on Information Forensics and Security* 8.3 (2013), pp. 477–487. DOI: 10.1109/TIFS.2013.2238930.

[24] A. Krizhevsky, I. Sutskever, and G. E. Hinton. “Imagenet classification with deep convolutional neural networks”. In: *Advances in Neural Information Processing Systems* (2012), pp. 1097–1105.

[25] Józef Lubacz, Wojciech Mazurczyk, and Krzysztof Szczypiorski. “Principles and overview of network steganography”. In: *IEEE Communications Magazine* 52.5 (2014), pp. 225–229. DOI: 10.1109/MCOM.2014.6815916.

[26] Talgat Manglayev, Refik Caglar Kizilirmak, and Yau Hee Kho. “Comparison Of Parallel And Successive Interference Cancellation For Non-Orthogonal Multiple Access”. In: *2018 International Conference on Computing and Network Communications (CoCoNet)*. 2018, pp. 74–77. DOI: 10.1109/CoCoNet.2018.8476815.

[27] Yijie Mao, Bruno Clerckx, and Victor O.K. Li. “Rate-splitting multiple access for downlink communication systems: bridging, generalizing, and outperforming SDMA and NOMA”. In: *J Wireless Com Network* 2018.133 (2018). DOI: 10.1186/s13638-018-1104-7.

[28] Nikolaos I. Miridakis and Dimitrios D. Vergados. “A Survey on the Successive Interference Cancellation Performance for Single-Antenna and Multiple-Antenna OFDM Systems”. In: *IEEE Communications Surveys Tutorials* 15.1 (2013), pp. 312–335. DOI: 10.1109/SURV.2012.030512.00103.

[29] “Network Information Theory”. In: *Elements of Information Theory*. John Wiley & Sons, Ltd, 2005. Chap. 15, pp. 509–611. ISBN: 9780471748823. DOI: https://doi.org/10.1002/047174882X.ch15. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/047174882X.ch15. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/047174882X.ch15.

[30] Timothy J. O’Shea, Johnathan Corgan, and T. Charles Clancy. “Convolutional Radio Modulation Recognition Networks”. In: *Engineering Applications of Neural Networks*. Ed. by Chrisina Jayne and Lazaros Iliadis. Cham: Springer International Publishing, 2016, pp. 213–226. ISBN: 978-3-319-44188-7.

[31] Erdal Panayirci, Habib Senol, and H. Vincent Poor. “Joint Channel Estimation, Equalization, and Data Detection for OFDM Systems in the Presence of Very High Mobility”. In: *IEEE Transactions on Signal Processing* 58.8 (2010), pp. 4225–4238. DOI: 10.1109/TSP.2010.2048317.

[32] Sung-Ik Park et al. "ATSC 3.0 Transmitter Identification Signals and Applications". In: *IEEE Transactions on Broadcasting* 63.1 (2017), pp. 240–249. DOI: 10.1109/TBC.2016.2630268.

[33] R.S. Roberts, W.A. Brown, and H.H. Loomis. "Computationally efficient algorithms for cyclic spectral analysis". In: *IEEE Signal Processing Magazine* 8.2 (1991), pp. 38–49. DOI: 10.1109/79.81008.

[34] R. Schoolcraft. "Low probability of detection communications-LPD waveform design and detection techniques". In: *MILCOM 91 - Conference record*. 1991, 832–840 vol.2. DOI: 10.1109/MILCOM.1991.258378.

[35] Bassant Selim et al. "Radio-Frequency Front-End Impairments: Performance Degradation in Nonorthogonal Multiple Access Communication Systems". In: *IEEE Vehicular Technology Magazine* 14.1 (2019), pp. 89–97. DOI: 10.1109/MVT.2018.2867646.

[36] A. Sonnenschein and P.M. Fishman. "Limitations on the detectability of spread-spectrum signals". In: *IEEE Military Communications Conference, 'Bridging the Gap. Interoperability, Survivability, Security'*. 1989, 364–369 vol.2. DOI: 10.1109/MILCOM.1989.103955.

[37] C.M. Spooner and W.A. Gardner. "The cumulant theory of cyclostationary time-series. II. Development and applications". In: *IEEE Transactions on Signal Processing* 42.12 (1994), pp. 3409–3429. DOI: 10.1109/78.340776.

[38] Christian Szegedy et al. "Intriguing properties of neural networks". In: *International Conference on Learning Representations*. 2014. URL: http://arxiv.org/abs/1312.6199.

[39] L. Turner. "The evolution of featureless waveforms for LPI communications". In: *Proceedings of the IEEE 1991 National Aerospace and Electronics Conference NAECON 1991*. 1991, 1325–1331 vol.3. DOI: 10.1109/NAECON.1991.165935.

[40] Mojtaba Vaezi et al. "Non-Orthogonal Multiple Access: Common Myths and Critical Questions". In: *IEEE Wireless Communications* 26.5 (2019), pp. 174–180. DOI: 10.1109/MWC.2019.1800598.

[41] Bob Ward. "Non Rectangular Time Windowing Analysis for IEEE 802.11 OFDM System". In: *IEEE P802. 11 Working Group Contribution, IEEE 802.11-99/021* (1999).

[42] Shihao Yan et al. "Low Probability of Detection Communication: Opportunities and Challenges". In: *IEEE Wireless Communications* 26.5 (2019), pp. 19–25. DOI: 10.1109/MWC.001.1900057.

[43] Paul L. Yu, John S. Baras, and Brian M. Sadler. "Physical-Layer Authentication". In: *IEEE Transactions on Information Forensics and Security* 3.1 (2008), pp. 38–51. DOI: 10.1109/TIFS.2007.916273.

Chapter 2

Covert Communications through Imperfect Cancellation

2.1 Introduction

Low probability of detection (LPD) signals are useful to reduce the probability of detection of a signal by an unauthorized observer. Much of the literature on LPD communications involves hiding the signal in additive white Gaussian noise (AWGN). Spread spectrum is popular for covert communication when the unauthorized observer does not know the spreading code and must, therefore, operate with a wider bandwidth than the intended recipient does and admit greater noise power. Detection is more likely to be successful if the unauthorized observer attempts to extract cyclostationary features from the signal [17] as opposed to e.g. energy detection.

In order to compete with more sophisticated detection techniques, a signal can be embedded within another signal as an alternative to hiding in AWGN [13]. The signal used for camouflage is the *incumbent signal*, and the signal intended to be hidden shall be referred to as the *covert signal*. The covert signal can be injected into any layer of the protocol stack of an incumbent signal, for example, by modifying the timing of the MAC layer in order to convey a pulse-width modulated message

[12]. Hiding a covert signal in the physical layer offers a much higher data rate as compared to using other layers of the stack thus motivating our work [6]. This work presents a novel means of covert communications in which a covert signal will be transmitted inside an OFDM signal and employ imperfect cancellation to minimize detectability of the covert signal while maintaining a high covert data rate. The contributions of this work are: 1) The design of a covert DSSS signal to be hidden in a OFDM signal (Section 2.3); 2) Test of the impact of this covert signal on the incumbent signal (Section 2.4.2); 3) An OFDM cancellation scheme designed to maintain white noise for the covert receiver (Section 2.5); 4) Test of the detectability of this covert signal (Section 2.7); and 5) Application of this cancellation scheme against a popular OFDM waveform (802.11) using both synthetic (Section 2.6) and OTA data (Section 2.8).

In the experiments, measurements of the PER of the OFDM signal and the BER of the covert signal were taken, both as functions of the signal-to-interference power ratio (SIR). The results show that the despreading alone was insufficient to recover the covert signal at an SIR high enough to preserve the PER of the OFDM signal. To address this problem, a cancellation stage was developed to provide sufficient suppression of the OFDM signal to recover the covert signal. The cancellation scheme was designed to avoid shaping the noise in the covert receiver.

The structure of this chapter is as follows: Section 2.2 provides a literature survey into related work. Section 2.3 describes the covert spread signal. Section 2.4 considers the recovery of the covert signal using despreading alone. Section 2.5 establishes the OFDM signal model and presents our method for partially canceling the OFDM signal. Section 2.6 demonstrates that with this cancellation, we can retrieve the covert signal with significant improvements to the BER. Section 2.7 details an experiment where the covert signal is tested against two variants of two detectors empirically proving that the covert signal is well hidden inside the OFDM

packet. Section 2.8 applies the technique to Over-the-Air data and demonstrates that the cancellation technique is effective in a real world environment. Section 2.9 summarizes the results, compares this work to prior work in covert communications under OFDM incumbent[1] signals, and provides some direction for future research.

A portion of this chapter has been accepted to be published in [4].

2.2 Related Work

Both [6] [7] make use of *dirty constellations*, by which a small error is added to the constellation of a incumbent signal and the value of that error conveys the covert signal. Tracking the error between what was expected of the incumbent signal and what was received is related to the cancellation technique discussed in this work. One problem we see with the *dirty constellations* is that the method relies on the received signal being equalized and thus shapes the noise in the covert receiver. This was avoided in our implementation as explained in section 4.2.2.

Four covert channels in Wi-Fi are analyzed in [5]. Those four covert channels were 1) modifying the symbols in the Short Training Field (STF) to add phase modulation for the covert signal once per OFDM packet, 2) adjusting the carrier frequency offset, 3) adding additional signals into the unused subcarriers of the OFDM symbol (as introduced by [9]), and 4) replacing parts of the OFDM Cyclic Prefix as introduced by [8].

Method 1 in [5] hides one covert symbol per STF. As there is only one STF per OFDM packet, the covert signal in this work has a higher data rate than method 1. Method 3 violates the spectral mask of the standard. The claim is that a spectrum monitor won't notice because the covert signal is attempting to look like another

[1] in terms common to steganography, this OFDM signal would be called the *cover* signal and the signal hidden inside of it would be the *covert* signal. The problem with that terminology is that *cover* and *covert* are only one letter apart. This work will rely on terminology common to the Cognitive Radio community and name the OFDM signal the *incumbent* signal.

version of the standard that the authors concede is not in use at the target location. Method 4 replaces the cyclic prefix with a covert signal, and this has negative consequences in multipath channels.

The most promising method in [5] is method 2 where one covert symbol is embedded in every 1 OFDM symbol, that being 80 samples at 20 MSPS. This is a slightly faster symbol rate than the covert signal developed in this work which transmits one covert signal every 128 samples. Method 2 is an FSK covert signal that multiplies a carrier offset of each individual OFDM symbol. Because method 2 is an FSK waveform, and does not have its own power to noise ratio, the covert BER is a function of the frequency deviation and the SNR of the OFDM signal. The BER of the covert signal improves as the frequency deviation is increased. Unfortunately, increasing the frequency deviation of the covert signal increases the impairment on the OFDM signal, which increases the PER curve of the OFDM signal and makes the covert signal noticeable. The results for method 2 showed that it required a a frequency deviation of 5 kHz or lower to avoid detection, however those results were not well explained. A 5 kHz frequency deviation over a 4 μs period results in a 0.04π radian phase shift. It was unclear under what criteria the authors declared this phase error to have no perceptible effect on the error rate of the OFDM signal. Using a smaller frequency deviation, such as 1 kHz, will reduce the impact on the incumbent but require the covert receiver to be closer to the OFDM transmitter in order to enjoy the high SNR values that would be needed to demodulate the covert signal at that frequency deviation.

In [16], a covert signal was created by introducing an artificial channel effect at the transmitter resulting in transmitting one covert symbol per OFDM packet. The work in [15] adds distortion to a 802.11 signal and uses a neural network to classify that distortion resulting in a covert signal data rate of 93.75 kbps. The covert signal presented in this work exceeds these data rates.

2.3 Covert Signal Design

The covert signal used here is a direct sequence spread spectrum (DSSS) signal that distributes signal power over the OFDM signal. The covert signal is injected into the unused (null) OFDM bin and is constrained by the spectral mask imposed by the 802.11 standard [10] in order to prevent self-revelation. Spreading mitigates interference and multipath, and despreading will spread the OFDM signal, reducing the total interference to the covert signal. The spread carrier conveys BPSK information at 156.25 kbaud and uses a pseudo-random sequence derived from the polynomial $z^6 + z + 1$. The peak-to-sidelobe ratio is 20.56 dB. The autocorrelation function of the spreading code is illustrated in Fig. 2.1.

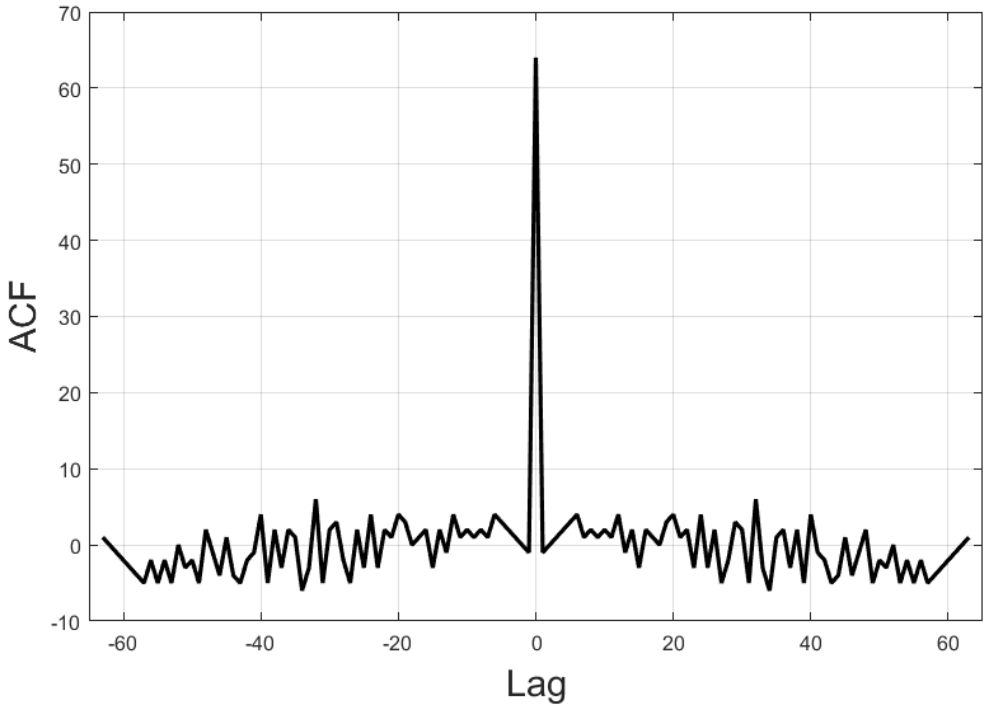

Figure 2.1: Spreading Code ACF

Each covert symbol requires 64 chips. At 10 Mcps, this requires 6.4 microseconds per covert symbol. When an 802.11 signal is used as the incumbent, the 6.4

microseconds period is twice the IFT frame of the OFDM symbol. The 802.11 signal applies a 0.8 microsecond cyclic prefix to each OFDM symbol making the total length 4 microseconds per OFDM symbol. These time values must be taken into account when injecting the covert signal into the incumbent signal. In order for the covert signal to remain within the time duration of an incumbent OFDM packet, the total transmit time of the covert signal must be less than that of the OFDM packet. Each instance of the covert signal can only have as many spread symbols as will fit inside the OFDM packet. The relationship in time is illustrated in Fig. 2.2.

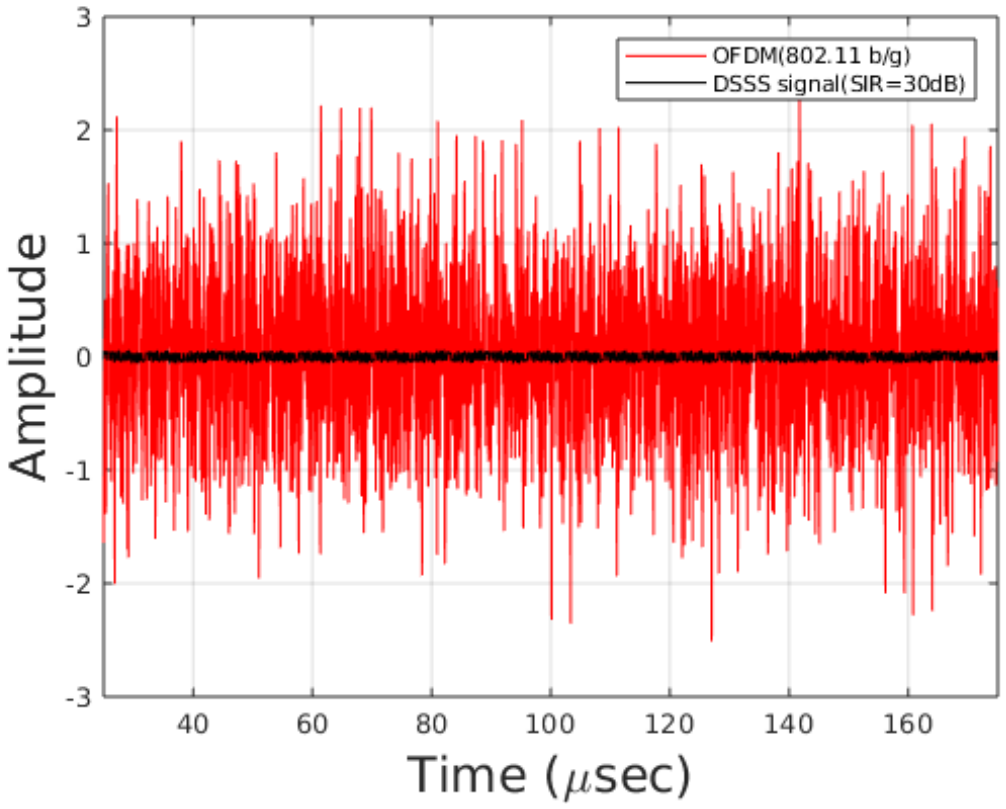

Figure 2.2: The Covert and OFDM Signal in Time

The spectrum of the OFDM signal and the covert signal are shown in Fig. 2.3. In this case the 802.11 signal is 23 dB above the noise power. The power of the covert DSSS signal is 35 dB below the OFDM signal. The covert signal adds no significant

power to the sidelobes of the OFDM signal or the center bin. The effect that this covert signal has on the OFDM signal will be measured in the following sections.

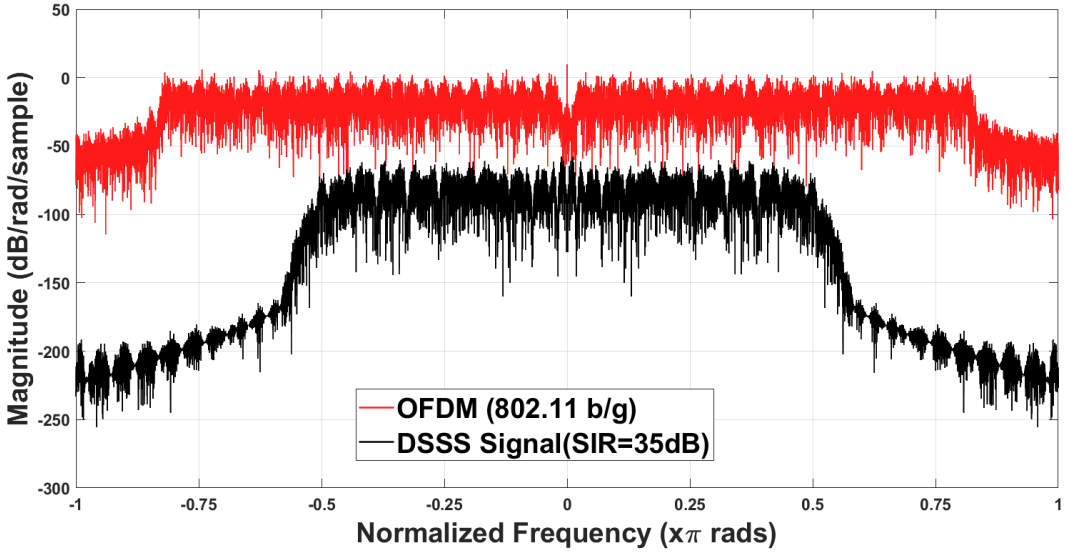

Figure 2.3: The Covert and OFDM Signal in Frequency

2.4 System Performance Without Cancellation

2.4.1 Performance Baseline

The first step in our experiment is to establish a baseline of the performance of the 802.11 receiver in terms of PER as a function of SNR without the covert signal present. This information will be used to determine what range of cascaded noise figures this receiver can handle. A noise figure is a characterization of a component that describes the degradation of the SNR from the input to the output of that component. A cascaded noise figure is an aggregate noise figure resulting from multiple cascaded components [14]. That range of cascaded noise figures will establish whether or not this can represent realistic receiver hardware.

Synthetic packets are generated to measure PER at different SNRs. Each packet has 1000 octets of random data. We iterate through SNR values from 0 to 25dB in

steps of 0.5 dB, passing the synthetic packet through an AWGN channel. For each SNR, we attempt to demodulate 5000 random packets with noise, recording the packet error we receive. This process is repeated for all modulation and coding schemes allowed by 802.11 b/g for a 20 MHz channel. For this experiment, we do not add synchronization or multipath impairments. Fig. 2.4 illustrates the results. Table 2.1 records the SNR value required to meet a 1% PER and shows the minimum input sensitivity requirements in dBm specified by the standard [10] at which the receiver must achieve a PER of 10% of less. The minimum sensitivity is -65 dBm for a data rate of 54 Mbps. The power spectral density (PSD) of thermal noise at 290° K is -174 dBm/Hz. Using a channel bandwidth of 20 MHz and a PSD of -174 dBm/Hz, the thermal noise power is -101 dBm. Therefore the receiver used in this work can handle a cascaded noise figure of up to 14.5 dB and still operate at a PER of 1% which exceeds the requirements of the standard.

2.4.2 OFDM PER as a Function of SIR

When an interferer tries to establish a covert channel, the predominant factor of outage for the OFDM signal can be attributed to that interferer. The work in [2] and [1] develops outage probabilities as a function of SIR. If the covert signal is too powerful, the PER of the OFDM incumbent will rise and will force the incumbent network to adapt or fail. Any monitor noticing this phenomenon may attribute the event to the presence of an interferer, thus negating covertness. It is therefore necessary to limit the power of the covert signal to prevent an increase in PER that would raise suspicions. At the receiver of the OFDM signal, PERs that deviate from the performance established in section 2.4.1 are good indicators for the presence of interferers. Such an outage, in the case of our covert channel, can be mitigated by reducing the power spectral density (PSD) of the covert signal further after the DSSS operation. The reduction has the unwanted effect of lowering the covert

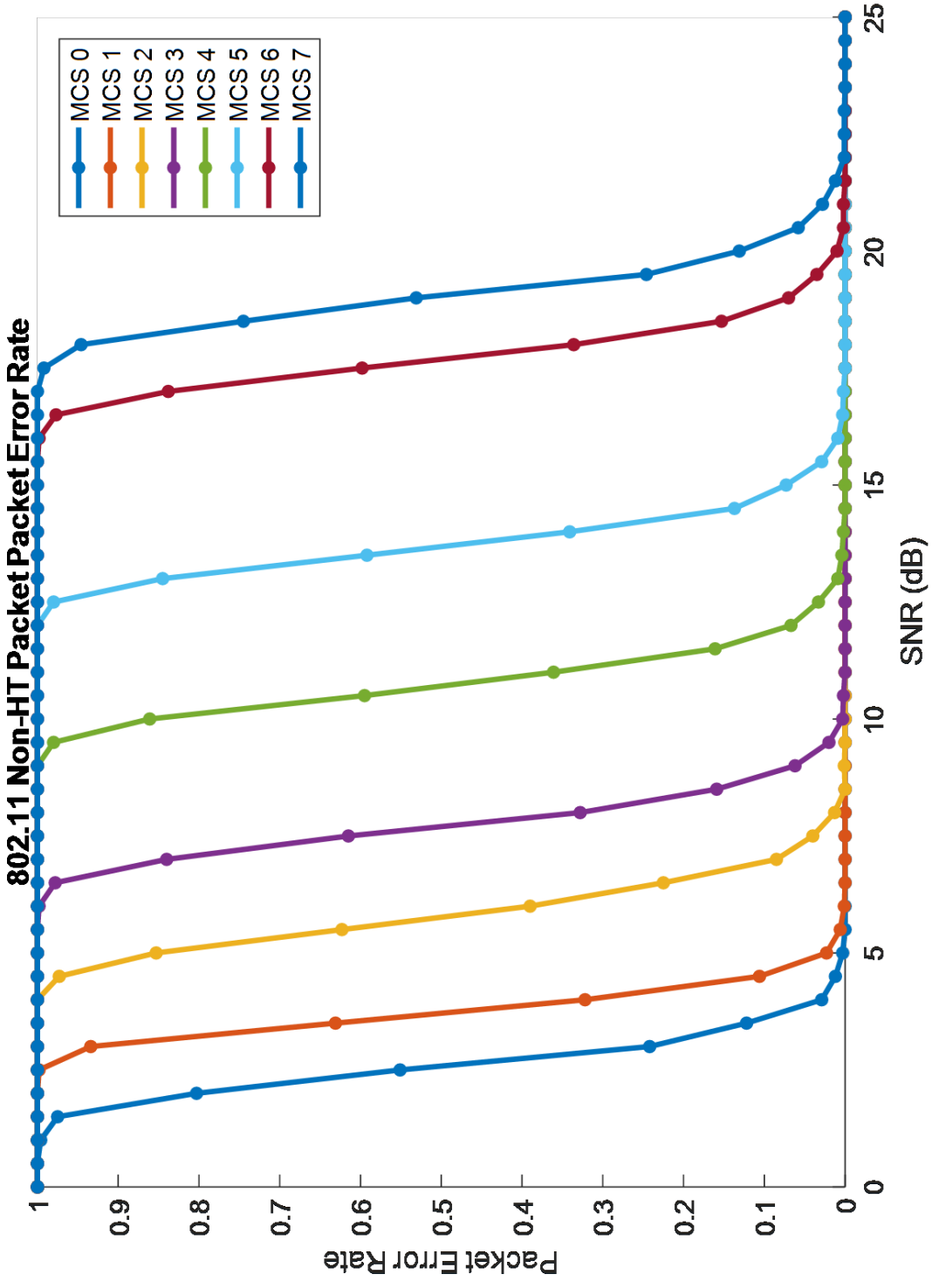

Figure 2.4: Baseline PER of 802.11 MODEM

Table 2.1: SNR Needed to Achieve 1% PER for each Data Rate

Modulation	**Coding Rate**	**Coded Bits Per Subcarrier**	**Coded Bits Per OFDM Symbol**	**Data Bits Per OFDM Symbol**	**Data Rates in Mbps**	**SNR (dB) Required for 1% PER**	**Minimum Sensitivity (dBm) Requirement**
BPSK	1/2	1	48	24	6	4.5	-82
BPSK	3/4	1	48	36	9	5.5	-81
QPSK	1/2	2	96	48	12	8	-79
QPSK	3/4	2	96	72	18	9.75	-77
16-QAM	1/2	4	192	96	24	13	-74
16-QAM	3/4	4	192	144	36	16	-70
64-QAM	2/3	6	288	192	48	20	-66
64-QAM	3/4	6	288	216	54	21.5	-65

channel's capacity as it reduces the energy available per symbol for the covert signal. Without a mechanism to separate the OFDM signal from the covert channel, the processing gain achieved from despreading does not provide a margin big enough to achieve the desired bit error performance. This is especially true if the competing objective to keep a low PER for the OFDM signal is desired. This leads to a search for a SIR operating point at which one finds the lowest BER for the covert signal and the performance of the OFDM signal is within a target PER. In order to evaluate this operating point, PER of the OFDM signal is measured over a range of SIR values. Noise power is kept constant in these measurements. Multipath and synchronization impairments were not added to the OFDM or covert signals. The OFDM signal was set to a data rate of 54 Mb/s (subcarrier modulation of 64 QAM with a coding rate of 3/4), which represents the highest data rate for 802.11 b/g. Our choice of the highest data rate for 802.11 b/g establishes the worst-case scenario in terms of PER performance for the 802.11 receiver. The results of this experiment are shown in Fig. 2.5. Four different SNRs were tested: no noise, 21 dB, 23 dB, and 25 dB. The SIR for each case was ranged from 0 dB to 45 dB. Note that the lowest PER possible for each curve relating to a finite SNR is reached asymptotically indicating an irreducible error rate determined by the presence of AWGN. These results show that the SIR must be kept above 30 dB in order for the covert signal to not be the dominant contributor to the PER in the OFDM signal.

2.4.3 Covert Signal BER as a Function of SIR

Fig. 2.6 depicts the bit error rate of the covert signal when the OFDM data rate is 54 Mb/s. The four curves represent different SNR values of the OFDM packet, including one with no AWGN. The driving factor is SIR rather than SNR. The difference in BER performance under different SNR scenarios is negligible. This is a reiteration of the fact that interference from the OFDM signal, and not ambient

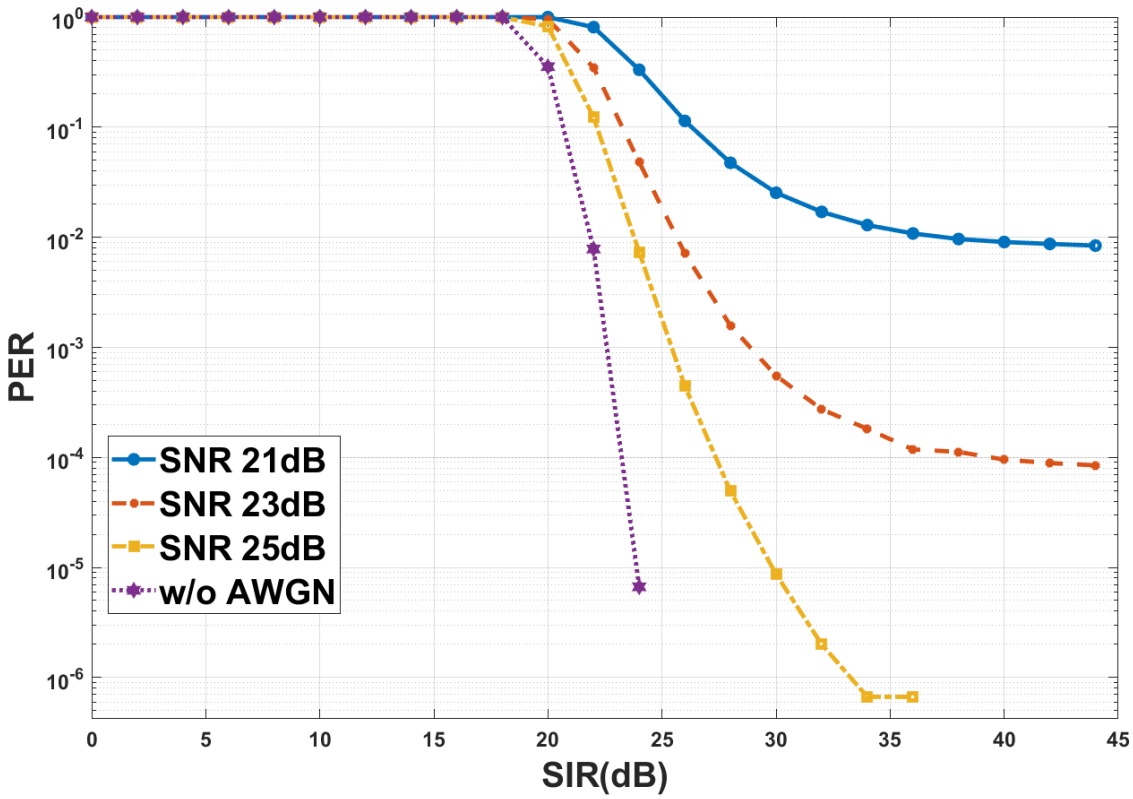

Figure 2.5: PER for 54 Mb/s 802.11 with a Covert Signal Present

noise, is the driving factor for the BER performance of the covert channel. Moreover, the SIR required to operate at 10^{-5} to 10^{-4} BER for the DSSS signal is around 12 dB. One can choose a lower SIR operating point for the covert signal but the limitation, as discussed earlier, is precipitated by reduced performance of the OFDM receiver.

Because the primary source of interference for the covert signal is the OFDM signal, some means of suppressing that interference must be introduced. The next section will introduce a method of partially canceling the OFDM signal.

2.4.4 Analysis of Incumbent PER and Covert BER as a Function of SIR

It was found that the SIR must be kept above 30 dB in order for the covert signal to not be the dominate cause of the PER for the incumbent OFDM signal. It was also

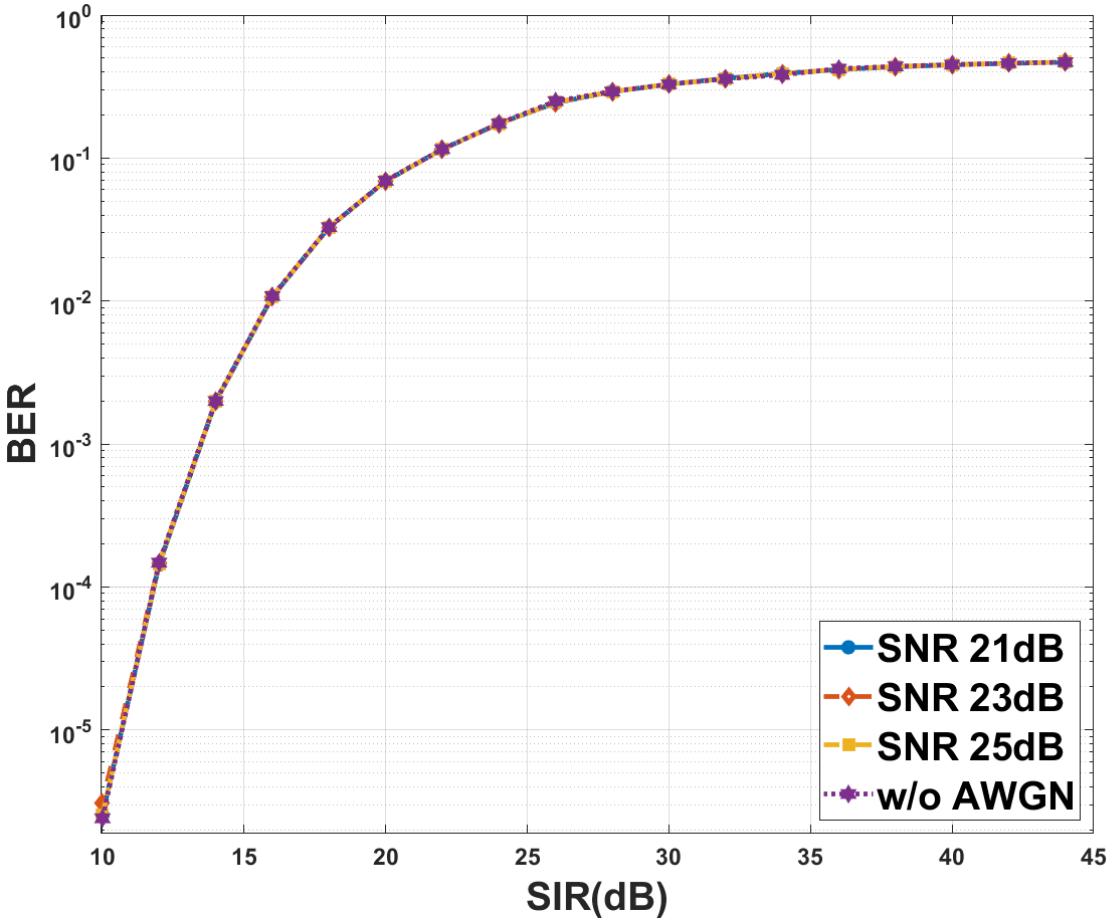

Figure 2.6: BER for Covert Signal inside 54 Mb/s 802.11

found that the SIR required to operate at 10^{-5} to 10^{-4} BER for the covert signal is around 12 dB. These two requirements have no overlap. Therefore, it is not possible to reliably communicate the covert signal in the OFDM incumbent without some additional processing applied.

2.5 OFDM Signal Cancellation

2.5.1 OFDM Signal Model

An OFDM packet is a concatenation of N OFDM symbols $\vec{o}_p$, $p = 0, 1, ...N-1$, where each individual symbol is created as the length-M Inverse Fourier transform (IFT) of a complex data vector $\vec{d}_p$ as shown in (2.1) where k represents discrete time

for that OFDM symbol. Each element of d_p will modulate a subcarrier in the IFT. The data frame d_p is composed of modulated complex-valued symbols, both for pilot subcarriers and data subcarriers, as well as null values for null subcarriers.

$$o_p[k] = \sum_{m=0}^{M-1} d_p[m] e^{j\frac{2\pi mk}{M}} \tag{2.1}$$

Most OFDM systems use a mandatory cyclic prefix which extends the OFDM symbols to be longer in time than the number of points in the IFT. This cyclic prefix helps mitigate the effects of multipath channels. As an example, an 802.11g OFDM symbol uses a 64-length IFT (M=64) at 20 MSPS for a frame period of 3.2 μs.That is then extended with a cyclic prefix by 0.8 μs for a total OFDM symbol period of 4 μs. This will be the OFDM symbol period used in the subsequent experiments in this chapter.

An OFDM packet is a concatenation of N OFDM symbols in time, where each individual symbol is denoted $o_p[k]$. The subscript p indicates the placement of an OFDM symbol in the packet and k indexes the individual samples in a given OFDM symbol. A model for an OFDM packet $\vec{s}$ is in (2.2). The initial OFDM symbols are often defined to serve as a preamble to aid synchronization and channel estimation at the receiver.

$$\vec{s} = \{\vec{o}_0 || \vec{o}_1 ... || \vec{o}_{N-1}\} \tag{2.2}$$

2.5.2 ODFM Signal Parameter Estimation

The demodulation process begins with receiving the preamble of the OFDM packet and performing frame synchronization. The 802.11 receiver firsts detects the packet and estimates the timing sample offset. The sample rate is not corrected.

The OFDM packet as seen by the receiver $\vec{r}$ after frame synchronization is shown

in (2.3) where matrix H represents the multipath channel, the diagonal matrix Λ_Θ represents the carrier frequency offset, and $\vec{n}$ noise at the receiver.

$$\vec{r} = \Lambda_\Theta H\vec{s} + \vec{n} \tag{2.3}$$

The carrier offset diagonal matrix Λ_Θ contains phase offsets for each sample of $\vec{s}$ as shown in (2.4). If the carrier frequency offset is constant then when integrated in time it will produce a phase ramp in Λ_Θ. The magnitude of each element is unity. Each element on the diagonal represents a phase offset value $\Theta[k]$ shown as a phase ramp in (2.5). The carrier frequency offset can be reversed by multiplying the diagonal matrix with its conjugate. The product results in an identity matrix, $\Lambda_\Theta^* \Lambda_\Theta = I$.

$$\Lambda_\Theta = \text{diag}(e^{j\vec{\Theta}}) \tag{2.4}$$

$$\Theta[k] = \omega k + \phi \tag{2.5}$$

Once sample timing of the 802.11 packet has been determined, estimates are needed for Λ_Θ and H, those being $\widehat{\Lambda_\Theta}$ and $\widehat{H}$. After those parameters are estimated, the conjugate and inverse of those estimates $\widehat{\Lambda_\Theta}^*$ and $\widehat{H}^{-1}$ will be found. Because the estimates $\widehat{\Lambda_\Theta}$ and $\widehat{H}$ are not perfect representations of Λ_Θ and H, there will be some error when the inverse of the estimate is applied.

The first impairment to be estimated is the center frequency impairment defined in (2.4). In order to correct the impairments, the estimates must be applied to the received samples. However, these impairment estimations are imperfect. The error resulting from applying the conjugate of the estimate of the carrier frequency offset is represented as the diagonal matrix $\Lambda_{\epsilon_\Theta}$ defined in (2.6) indicating that some

residual offset remains.

$$\widehat{\Lambda_\Theta}^* \Lambda_\Theta = \Lambda_{\epsilon_\Theta} \tag{2.6}$$

The frequency correction estimate $\widehat{\Lambda_\Theta}^*$ is applied to the received samples $\vec{r}$ before an estimate of the channel matrix can be calculated. Therefore, the error $\Lambda_{\epsilon_\Theta}$ propagates into the estimation of the channel impulse response. The estimate of the channel impulse response is initially calculated using the known sequence in the 802.11 preamble. This only provides estimates for the 52 non-zero subcarriers. The channel estimate is linearly interpolated over the null-subcarriers in phase and magnitude.

If the channel impulse response has an inverse, then the inverse channel matrix H^{-1} can be multiplied with H and the result will be the identity matrix, $H^{-1}H = I$. The inverse of the channel estimate $\widehat{H}$ is calculated and used to equalize the received signal. The channel estimate is imperfect and therefore the equalization will be imperfect. The combined error resulting from the imperfections of both the inverse-channel estimate and the carrier frequency offset estimate is represented as $\epsilon_{H\Theta}$ as shown in (2.7). This error trends to the identity matrix as the estimates come closer to the actual impairments.

$$\widehat{H}^{-1} \Lambda_{\epsilon_\Theta} H = \epsilon_{H\Theta} \tag{2.7}$$

2.5.3 Applying Cancellation

The cancellation process is illustrated in Fig. 2.7. The OFDM transmitter modulates and transmits an OFDM packet. That packet passes through a multipath channel. The covert receiver estimates carrier frequency offset, channel impulse response, and demodulates the OFDM packet. The demodulated bits can be encrypted and

the covert receiver has no need of decrypting these bits. The covert receiver will re-modulate the bits creating a local version of the received OFDM packet. The covert receiver will use this local copy to cancel the received OFDM signal. If the cancellation provides sufficient suppression of the OFDM signal, the covert signal can be recovered.

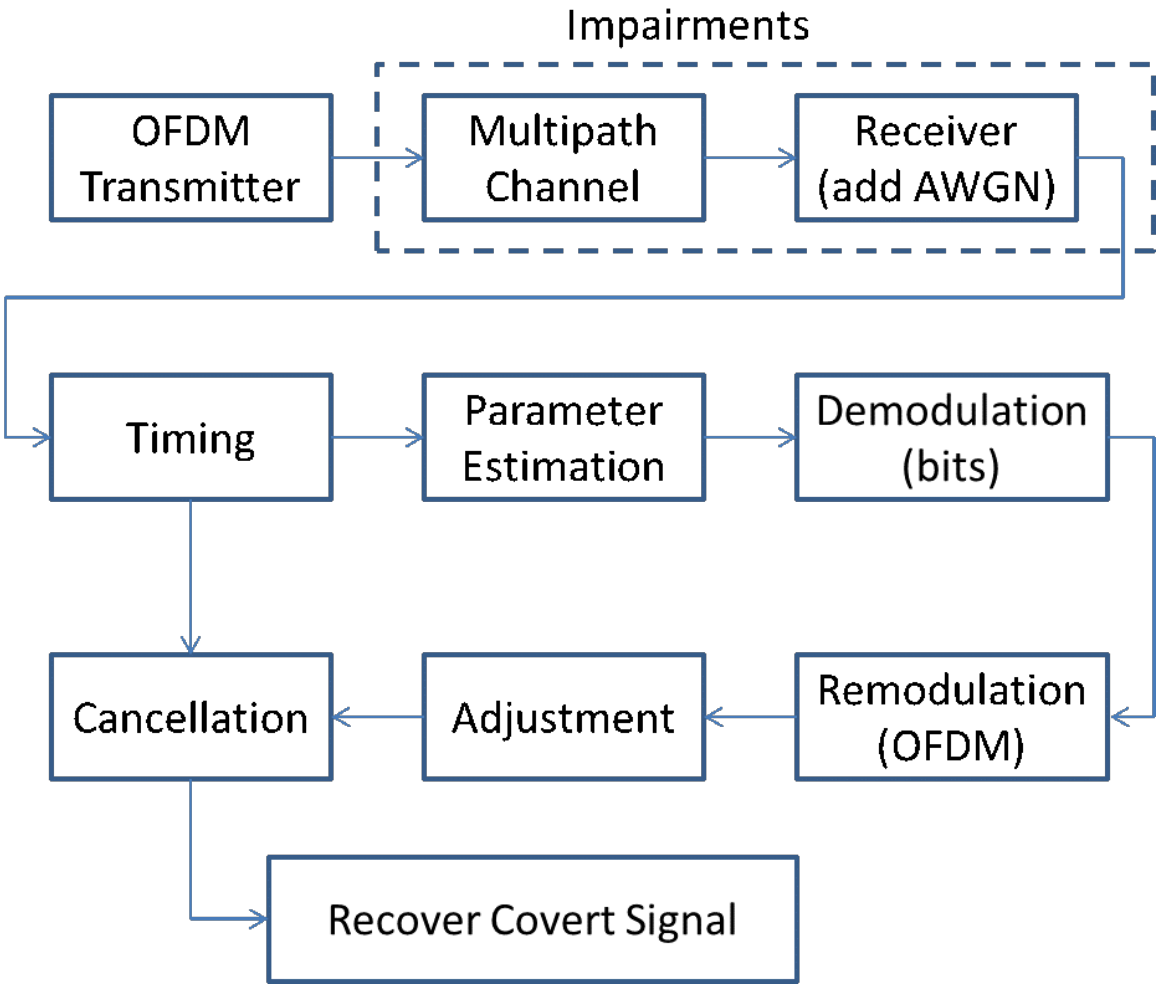

Figure 2.7: Covert Signal Recovery with Cancellation

The re-modulation produces an estimated OFDM waveform which will be represented as $\hat{s}$. The waveform $\vec{s}$ represents the OFDM signal at the transmitter with no impairments of any kind. After the estimated corrections have been applied to the received samples, the estimated OFDM signal can be subtracted from the

received samples resulting in a residue $\vec{u}$ as shown in (2.8).

$$\vec{u} = \epsilon_{H\Theta}\vec{s} + \widehat{H}^{-1}\widehat{\Lambda_\Theta}^{*}\vec{n} - \vec{s} \tag{2.8}$$

Assuming the demodulation process saw no bit errors, and assuming $\widehat{H}$ and $\widehat{\Lambda_\Theta}$ are perfect estimates, then $\widehat{s} = \vec{s}$. In this case, the residue reduces to (2.9).

$$\vec{u} = (\epsilon_{H\Theta} - I)\vec{s} + \widehat{H}^{-1}\widehat{\Lambda_\Theta}^{*}\vec{n} \tag{2.9}$$

The signal-error-term of the residue is $\epsilon_{(H\Theta - I)\vec{s}}$. As the estimation functions improve, $\epsilon_{H\Theta} \to I$, driving the residue error to zero. The noise-term of the residue is $\widehat{H}^{-1}\widehat{\Lambda_\Theta}^{*}\vec{n}$. The diagonal matrix $\widehat{\Lambda_\Theta}^{*}$ will not change the autocorrelation of the noise. The inverse channel response $\widehat{H}^{-1}$ may shape the noise in spectrum meaning the autocorrelation is no longer an impulse and therefore samples of noise are no longer independent. In addition to shaping the noise, applying the equalizing inverse channel may affect the covert signal. Consider the case in (2.10) where a covert signal $\vec{c}$ is present and may have different impairments than the incumbent signal $\vec{s}$.

$$\vec{r} = \Lambda_{\Theta_s} H_s \vec{s} + \Lambda_{\Theta_c} H_c \vec{c} + \vec{n} \tag{2.10}$$

If $H_s \neq H_c$ then applying the equalization $\widehat{H}_s^{-1}$ on all of $\vec{r}$ would further impair the covert signal. Given the potential for problems, an alternate approach to cancellation is used. Instead of applying the inverse estimated channel to the received samples, the estimated channel is applied to the estimated signal as shown in (2.11).

$$\vec{u} = \Lambda_\Theta H\vec{s} + \vec{n} - \widehat{\Lambda_\Theta}\widehat{H}\widehat{s} \tag{2.11}$$

Using the same assumption about the transmitter and bit errors, the residue reduces to (2.12).

$$\vec{u} = (\Lambda_\Theta H - \widehat{\Lambda_\Theta}\widehat{H})\vec{s} + \vec{n} \tag{2.12}$$

The signal-error-term of the residue is $(\Lambda_\Theta H - \widehat{\Lambda_\Theta}\widehat{H})\vec{s}$. As the estimation functions improve, the estimated parameters approach the actual parameters, $\widehat{\Lambda_\Theta}\widehat{H} \rightarrow \Lambda_\Theta H$. As that happens, the signal-error term goes to zero. This would leave only the noise term in the residue. We do not expect that our signal parameters will be perfect. The cancellation technique is expected to be imperfect and we expect some remainder of the OFDM signal to be present in the residue. The goal is to estimate a sufficient number of signal parameters to create a copy of the OFDM with enough accuracy to provide significant suppression of the OFDM signal.

2.6 System Improvement With Cancellation

For this experiment, we recover the spread signal after cancellation in our simulation as described in section 2.5. Estimates for carrier frequency offset and the channel impulse response are created as part of the demodulation process. However, we found that the estimates of the channel impulse response and carrier frequency offset could be improved by comparing the remodulated signal to the original received signal. We adjust the power of the remodulated signal $\widehat{s}$ to match the received signal $\vec{r}$. We also compare the phase difference between $\widehat{s}$ and $\vec{r}$ and rotate $\widehat{s}$ with a phase (not frequency) offset to lower the mean difference in phase between the two signals.

After cancellation, the ratio of the power of the residue and the OFDM signal represents the suppression of the OFDM signal, and that is calculated to be -20 dB. Given that the noise is 23 dB below the OFDM signal, and the residue is 20 dB below the OFDM signal, we conclude that there remains some OFDM signal power in the residue. That some OFDM signal power remains is to be expected, as the

cancellation is imperfect. If the cancellation were perfect, then the residue would contain only noise.

The second part of this experiment injects the covert signal into the OFDM signal before any parameter estimation is performed. The power of the covert signal is varied relative to the power of the OFDM signal, varying the SIR. The results are plotted in Fig. 2.8. The covert signal BER dramatically improves when aided by the cancellation algorithm. Note that as SIR decreases below 24 dB, the BER of the covert signal begins to increase. This is due to the covert signal interfering with the OFDM signal so severely that it causes bit errors and thus degrades the demodulation process illustrated in Fig. 2.7.

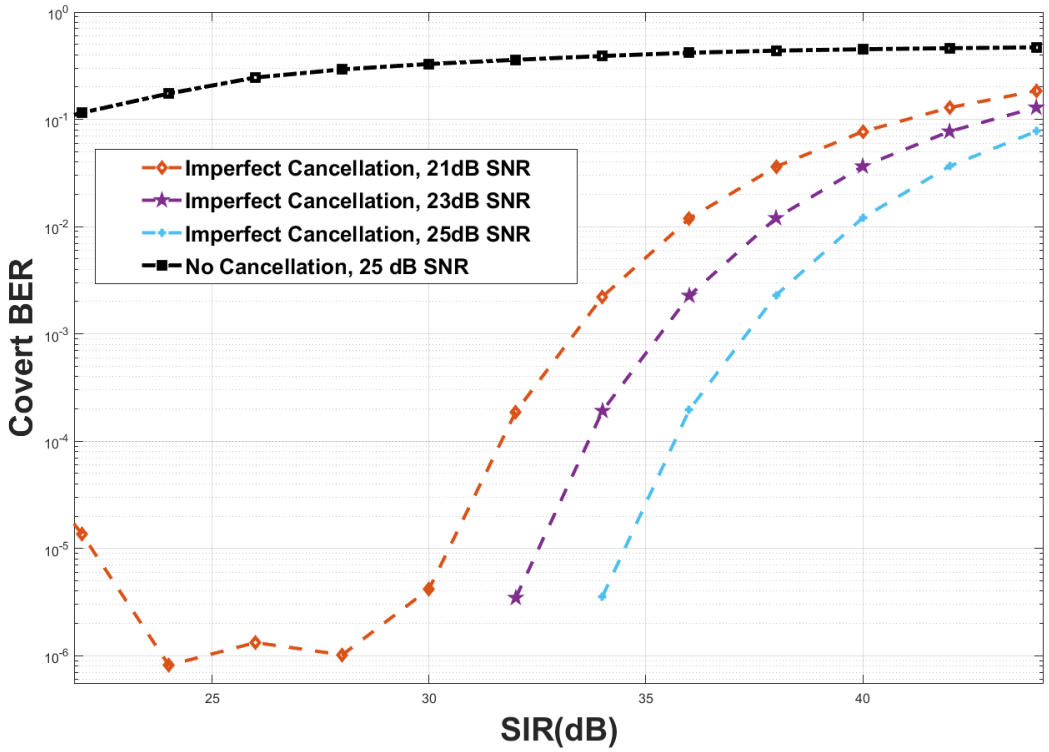

Figure 2.8: Covert BER Improvement

2.7 Detection of the Covert Signal

In this experiment, we test the detectability of the covert signal against a cyclostationary detector and an energy detector. The cyclostationary detector based on the FFT Accumulation Method which provides an estimate of the spectral correlation function [3]. That estimate is used to produce a cycle-frequency domain profile from which features are extracted for signal detection [11]. These tests are performed on synthetic data. We test two cases for each detector. The first test case executes the detection against the received signal as defined in (2.3) (no covert signal) and (2.10) (covert signal present). This test case represents the performance of the covert signal against detectors with no augmentation. The second test case gives the detectors the same cancellation advantage as the covert receiver in section 2.6, thus the detectors are run against the residue defined in (4.7). Sample packets from the incumbent OFDM signal are created with an SNR of 25 dB. The covert signal is injected beneath half of those packets at an SIR of 31 dB. The results of these experiments are shown in the receiver operating characteristic (ROC) plot in Fig. 2.9, illustrating the results of the energy detector (ED) with and without cancellation and the cyclostationary detector (CSA) with and without cancellation. In all cases without cancellation, the detectors are incapable of reliably detecting the presence of the covert signal.

The cyclostationary detector augmented with cancellation performs well, yielding a ROC curve with a considerable area under the curve. This means that a low probability of false alarm can be selected and still maintain a high probability of detection. However, this means that the cyclostationary detector must include an OFDM cancellation capability. The energy detector does not show much promise with or without the cancellation, although cancellation does yield some improvement to the ROC curve of the energy detector.

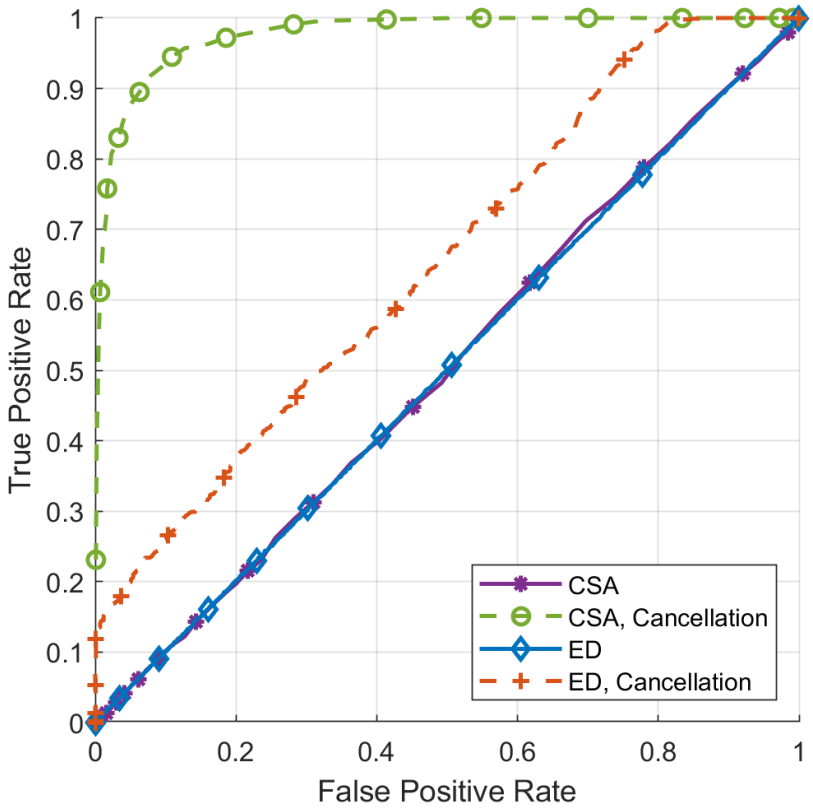

Figure 2.9: Receiver Operating Characteristic curves of the Four Detectors

2.8 OTA Experiment

For this experiment, complex baseband recordings were made at 40 MSPS of OTA 802.11g Wi-Fi packets with a data rate of 54 Mb/s. The Wi-Fi signals in these recordings exhibit real-world multipath channels and carrier frequency offsets. The cancellation process is the same as in sections 2.6 and 2.7, except for two additional steps. The first modification was to compare the phase difference between $\widehat{r}$ and $\vec{r}$, where $\widehat{r} = \widehat{\Lambda_\Theta}\widehat{H}\widehat{s}$, and rotate $\widehat{s}$ with a phase offset to lower the mean difference in phase between the two signals. The second additional step was to normalize the power between $\widehat{r}$ and $\vec{r}$.

We injected the covert signal into the OTA Wi-Fi recordings. The SNR estimates of the OTA packets varied from 29 to 31 dB. The covert signal is not perfectly aligned

with the Wi-Fi packet as the Wi-Fi packet has a carrier frequency offset. Additionally, the covert signal is not impaired with a multipath channel whereas the OTA Wi-Fi Packet does exhibit that impairment. This is to say that the covert signal shares neither a frequency reference nor a multipath channel with the incumbent signal. The resulting BER will therefore provide a best-case for the covert signal given the impairments of the OTA incumbent. The covert signal is injected before any parameter estimation is performed, therefore the effect on the incumbent estimators will be tested in this experiment. The power of the covert signal is varied relative to the power of the OTA Wi-Fi signal, and this varies SIR. Finally, we measure the BER performance of the covert signal with and without cancellation as a function of SIR. We ran 35427 covert bits beneath 961 OTA packets over a range of SIR values. The results are plotted in Fig. 2.10. The average cancellation was 19.5 dB with no covert signal present, which is lower than observed in simulation. The covert signal BER dramatically improves when aided by the cancellation algorithm. The range of SIR values in which the covert signal can operate exceeds the requirements established in analysis.

2.9 Conclusion

The work in this chapter measured the impact of a DSSS spread covert signal on an OFDM signal. We have shown that SIR, not SNR, is the primary limitation of the covert signal sharing the link with the OFDM signal. When high modulation orders are used in the OFDM signal, the covert signal must operate at high SIR values in order to prevent the interference from causing easily identified anomalies in loss of throughput for the incumbent signal. A higher SIR at a fixed SNR impacts the BER of the covert signal as the power of the covert signal power is being reduced with respect to noise. One means of mitigating this problem is to reduce the data rate of the covert signal. Instead of reducing the bandwidth of the covert signal,

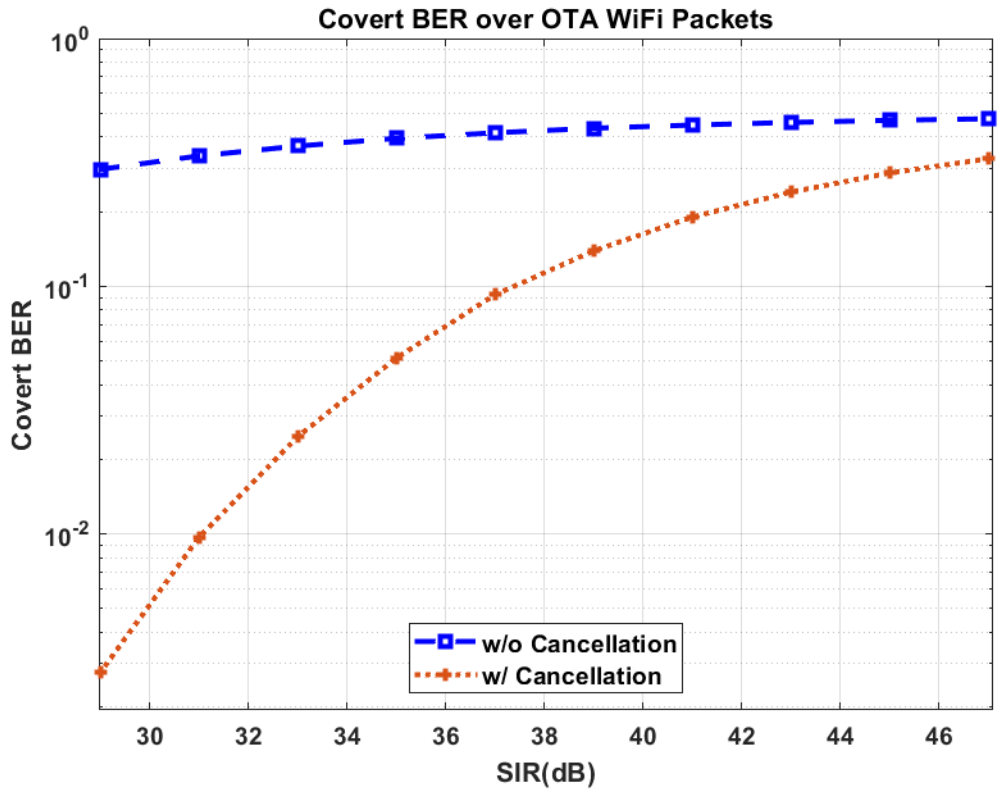

Figure 2.10: Covert BER using OTA Data

we implemented imperfect cancellation of the OFDM signal and demonstrated that was sufficient to recover a 156.25 kbaud BPSK signal spread at 10 Mcps from beneath an OFDM signal using 64 QAM. This recovery was demonstrated both in simulation and when using OTA data. We observed that the cancellation on OTA data is lower than that of our simulation, and we attribute that to signal parameters not estimated in our signal model. Our covert signal is capable of operating at SIR values higher than 30 dB, which maintains the throughput of the OFDM signal.

Future research on the covert method described in this work includes but is not limited to increasing the modulation order of the covert signal in order to increase the data rate and more direct comparisons with methods like those described in [6] and [7].

2.10 References

[1] J.G. Andrews. "Interference cancellation for cellular systems: a contemporary overview". In: *IEEE Wireless Communications* 12.2 (2005), pp. 19–29. DOI: 10.1109/MWC.2005.1421925.

[2] Paulo Cardieri. "Modeling Interference in Wireless Ad Hoc Networks". In: *IEEE Communications Surveys Tutorials* 12.4 (2010), pp. 551–572. DOI: 10.1109/SURV.2010.032710.00096.

[3] Daniel Chew, Andrew L. Adams, and Jason Uher. "Secondary Spectrum Usage and Signal Detection". In: *Wireless Coexistence: Standards, Challenges, and Intelligent Solutions*. 2021, pp. 115–154. DOI: 10.1002/9781119584230.ch5.

[4] Daniel Chew et al. "Covert Communications through Imperfect Cancellation". In: *Accepted to Proceedings of the 2022 ACM Workshop on Information Hiding and Multimedia Security*. IH&MMSec '22. New York, NY, USA: Association for Computing Machinery, 2022.

[5] Jiska Classen, Matthias Schulz, and Matthias Hollick. "Practical covert channels for WiFi systems". In: *2015 IEEE Conference on Communications and Network Security (CNS)*. 2015, pp. 209–217. DOI: 10.1109/CNS.2015.7346830.

[6] Aveek Dutta et al. "Secret Agent Radio: Covert Communication through Dirty Constellations". In: *Information Hiding*. Ed. by Matthias Kirchner and Dipak Ghosal. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 160–175. ISBN: 978-3-642-36373-3.

[7] Salvatore D'Oro, Francesco Restuccia, and Tommaso Melodia. "Hiding Data in Plain Sight: Undetectable Wireless Communications Through Pseudo-Noise Asymmetric Shift Keying". In: *IEEE INFOCOM 2019 - IEEE Conference on Computer Communications*. 2019, pp. 1585–1593. DOI: 10.1109/INFOCOM.2019.8737581.

[8] Szymon Grabski and Krzysztof Szczypiorski. "Steganography in OFDM Symbols of Fast IEEE 802.11n Networks". In: *2013 IEEE Security and Privacy Workshops*. 2013, pp. 158–164. DOI: 10.1109/SPW.2013.20.

[9] Zaid Hijaz and Victor S. Frost. "Exploiting OFDM systems for covert communication". In: *2010 - MILCOM 2010 MILITARY COMMUNICATIONS CONFERENCE*. 2010, pp. 2149–2155. DOI: 10.1109/MILCOM.2010.5680484.

[10] "IEEE Standard for Information Technology–Telecommunications and Information Exchange between Systems - Local and Metropolitan Area Networks–Specific Requirements - Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications". In: *IEEE Std 802.11-2020 (Revision of IEEE Std 802.11-2016)* (2021), pp. 1–4379. DOI: 10.1109/IEEESTD.2021.9363693.

[11] Kyouwoong Kim et al. “Cyclostationary Approaches to Signal Detection and Classification in Cognitive Radio”. In: *2007 2nd IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks*. 2007, pp. 212–215. DOI: 10.1109/DYSPAN.2007.35.

[12] Negar Kiyavash et al. “A Timing Channel Spyware for the CSMA/CA Protocol”. In: *IEEE Transactions on Information Forensics and Security* 8.3 (2013), pp. 477–487. DOI: 10.1109/TIFS.2013.2238930.

[13] Józef Lubacz, Wojciech Mazurczyk, and Krzysztof Szczypiorski. “Principles and overview of network steganography”. In: *IEEE Communications Magazine* 52.5 (2014), pp. 225–229. DOI: 10.1109/MCOM.2014.6815916.

[14] David M Pozar. “Chapter 10: Noise and Active RF Components”. In: *Microwave engineering; 3rd ed.* Hoboken, NJ: Wiley, 2005, pp. 493–497.

[15] Kunal Sankhe et al. “Impairment Shift Keying: Covert Signaling by Deep Learning of Controlled Radio Imperfections”. In: *MILCOM 2019 - 2019 IEEE Military Communications Conference (MILCOM)*. 2019, pp. 598–603. DOI: 10.1109/MILCOM47813.2019.9021079.

[16] Matthias Schulz et al. “Shadow Wi-Fi: Teaching Smartphones to Transmit Raw Signals and to Extract Channel State Information to Implement Practical Covert Channels over Wi-Fi”. In: *Proceedings of the 16th Annual International Conference on Mobile Systems, Applications, and Services*. MobiSys ’18. Munich, Germany: Association for Computing Machinery, 2018, 256–268. ISBN: 9781450357203. DOI: 10.1145/3210240.3210333. URL: https://doi.org/10.1145/3210240.3210333.

[17] C.M. Spooner and W.A. Gardner. “The cumulant theory of cyclostationary time-series. II. Development and applications”. In: *IEEE Transactions on Signal Processing* 42.12 (1994), pp. 3409–3429. DOI: 10.1109/78.340776.

Chapter 3

Image Classification for Signal Detection

3.1 Introduction

Given the near-insatiable demand for wireless connectivity, the 2.4 GHz ISM band has become heavily congested with competing signals, including WiFi, Bluetooth, and some Zigbee waveforms. As unlicensed and shared spectrum becomes more crowded, *spectrum sensing* becomes increasingly necessary to determine if a channel is occupied before using it. In *dynamic spectrum access*, sometimes called opportunistic spectrum sharing, the spectrum is monitored for available spectral slots, or *white space*. Upon detecting an available slot, the system then transmits in that slot. One example of dynamic spectrum access is in accessing TV white space [16]. Another example of opportunistic spectrum sharing is the Citizens Broadband Radio Service (CBRS) [11], [14], which uses a frequency band that had, until recently, been reserved for the use of radar systems. Due to this increasing demand, improved methods of spectrum sensing are being explored. Spectrum sensing in this context is to classify a set of input data as either having a signal present or having no signal present.

Classification techniques can be generalized as having two stages, a feature

extractor and a classifier. A feature of the signal is a quantitative measurement of some attribute of that signal. Some examples of features of signals are energy, bandwidth, center frequency, duration in time, and symbol rate [1]. It is sometimes useful to think of *raw features* which comes directly from the data source and *derived features* which come from some manipulation of that data [9]. The feature extractor maps the input data to points in a set of features, and the classifier uses those features to discern between classes. Feature selection is an important part of the classification process. Features for classification are intended to minimize the difference within a class and maximize the differences between classes. The exact boundaries between classes within the chosen feature space is determined in the classifier. An Artificial Neural Network (ANN) provides a trainable classifier which can learn how to draw those boundaries. Rather than use hand-engineered features to discern between classes, Deep Learning can be used to converge on a set of features to be extracted from the input data. This chapter explores the use of Deep Learning to detect the presence of signals in noise. A convolutional neural network (CNN) is an example of Deep Learning that has been used to distinguish between classes of images [12].

Taking the spectrogram [1] of the received data allows the signal detection problem to become an image-recognition problem. To this end, a CNN was modified and repurposed through the process of "fine-tuning" [6]. The concept of fine-tuning was used to re-purpose a CNN, which had been pre-trained for image recognition, by replacing the classification layers and training those new layers on new images without causing a large effect to the feature extraction gained from the pre-training. This process is also called "transfer learning." The AlexNet CNN [12] was chosen for this research due to the demonstrated resilience to noise [7]. Transfer learning enables this task to make use of AlexNet with only a few hundred training samples,

[1] A spectrogram is a two dimensional frequency vs time image of a signal.

as opposed to the over one million training samples originally required. After fine-tuning the CNN to distinguish between spectrograms in which a signal is present or not, the modified CNN is referred to in this work as the CNN detector. The system performance is characterized from experimental outcomes and compared to the classic energy detector [18].

Five experiments were performed in this work. The first three experiments aim to detect a primary user in noise and interference. The primary user experiments assess the ability of the CNN detector to detect signals in: (i) a fixed noise floor; (ii) across variable noise floor; and (iii) with interference. The experiments were separated into training and testing phases, and the signal is BPSK in all the primary user experiments.

In the noise floor experiment, the CNN detector was tested across different SNRs where the absolute value of the noise power was fixed and only the relative value of the signal power was allowed to vary. In the variable noise floor experiment, the CNN detector was again tested across a range of SNR values, but this time the absolute value of noise power varied. The goal of both noise experiments was to demonstrate that the CNN detector is a viable signal detection method that operates without a noise floor estimate. In the interference experiment, a variable power, random frequency Continuous Wave (CW) interference signal was added to the received signal.

The first three experiments were published in [2] and are focused on users with exclusive access to a channel, i.e., primary users. The last experiment in this work tests the AlexNet detector's ability to detect the covert signal as defined in chapter 2 and [3]. That experiment is generalized to other secondary users as defined in chapter 1.

The secondary-user AlexNet Detector is trained on residue impaired by errors in the center frequency, channel, and payload estimates as described in chapter 2

and [3].These secondary-user detectors are then tested in imperfect cancellation residue. For completeness, the primary-user detectors (those trained on AWGN) are also tested in the secondary-user experiments. Using these different detectors trained in different types of noise answers the research questions: (i) how well can the detectors perform in the imperfect cancellation residue; (ii) how well do the detectors transfer across those distributions; and (iii) whether or not the Constant False Alarm Rate (CFAR) behavior of the detectors is maintained between operating in AWGN and the cancellation residue.

The structure of this chapter is as follows: Section 3.2 provides a brief literature survey into examples of neural networks being employed for similar or related tasks. Section 3.3 describes the operation of the classic energy detector. Performance characteristics and closed form solutions for that classic detector are provided. Section 3.4 describes the creation and operation of the CNN detector. Section 3.5 discusses training the CNN detector. Section 3.6 provides the test results for the experiments in AWGN and interference. Section 3.7 concludes the primary-user experiments in AWGN and represents the end of the work published in [2]. Section 3.8 extends the work in [2] by describing experiments in cancellation residue. Section 3.9 provides an analysis of the experiments in cancellation residue. Lastly, section 3.10 provides directions for future work.

3.2 Neural Networks for Signal Detection

For signal detection, the classifier will distinguish between two classes, Signal Present and Signal Not Present. The classifier presented in this chapter will perform this function in the presence of AWGN and CW interference. One of the earliest references to the use of machine learning in signal detection comes from [20] and [15]. In [20], a multilayer perceptron was used for detecting the presence of a target signal

corrupted by bandlimited AWGN. In both [20] and [15] it is shown that a multilayer perceptron can converge to a minimum mean square error approximation of the Bayes optimal discriminant. This concept was further tested in [13] where it was demonstrated that for a single signal the multilayer perceptron did converge to the optimal classifier. As more types of signals were added, increasing uncertainty, the multilayer perceptron did not converge to the optimal detector. It was concluded in [13] that the nonlinearities (activation functions) used in the neurons could not approximate the optimal likelihood ratio test in the presence of multiple classes of signals. The nonlinearity of the multilayer perceptron was replaced with a radial-basis function in [5] and used to detect signals in non-Gaussian noise that followed either a double exponential, contaminated Gaussian, or Cauchy distribution.

Reference [19] describes the use of a Support Vector Machine (SVM) to determine the threshold for a CFAR energy detector. An SVM is a classifier, and the experiment in [19] used the SVM to classify the environment. From that classification of the environment, a CFAR threshold was selected. In reference [17], an ANN was employed for signal detection using multiple features combining energy and cyclostationary properties. The test signal in reference [17] was analog Amplitude Modulation, oversampled by 10 with 1024 samples.

Deep learning is a subset of machine learning where the feature extractor is trained along with the classifier. The advantage is that the deep learning classifier can, through training, identify a feature set which distinguishes between the desired classes as opposed to designing those features by hand. Deep learning was used for signal detection in [8]. Reference [8] uses a CNN to detect a Binary Phase-Shift Keying (BPSK) signal oversampled by 100. Reference [8] provides cyclostationary properties as the input to the CNN.

3.3 The Energy Detector

The classic energy detector [18], also called a radiometer, computes the sum of the squares of the magnitude of each sample over a vector of observed samples. This sum of squared-magnitude values represents the energy of the signal. The energy detector uses this measure of energy to perform a binary hypothesis test. If the measured energy exceeds a predetermined threshold, then the detector determines that a signal is present, however, this decision may be incorrect and therefore a *false alarm*. If the energy falls short of the threshold, then the detector will determine there is no signal present, however this decision may be incorrect and therefore a *missed detect*. The incoming signal will be complex-valued and at baseband; therefore the decision regions shown in Fig. 3.1 are on the complex plane. There are two error categories, *false alarm* and *missed detection*. A false alarm (*fa*), occurs when the detector erroneously determines that a signal is present. A missed detect (*md*), occurs when the detector fails to determine that a signal is present. The probabilities P_{fa} and P_{md} of false alarm and missed detection, respectively are not symmetric. The energy detector is a parametric detector as it requires knowledge of the noise power.

Expressions for P_{fa} and P_{md} in AWGN are given in equation 3.1 and 3.3. Proofs are provided in reference [4]. $\Gamma(\cdot,\cdot)$ is the normalized upper incomplete gamma function as shown in (3.2) where the normalization is $\frac{1}{\Gamma(N)} = (N-1)!$. σ is the standard deviation of the AWGN. λ is the tuneable threshold for the energy detector. The Q function is the Marcum-Q function. N is the number of complex samples used to measure the energy. γ is the Signal to Noise ratio (SNR).

$$P_{fa} = \Gamma\left(N, \frac{\lambda}{2\sigma^2}\right) \tag{3.1}$$

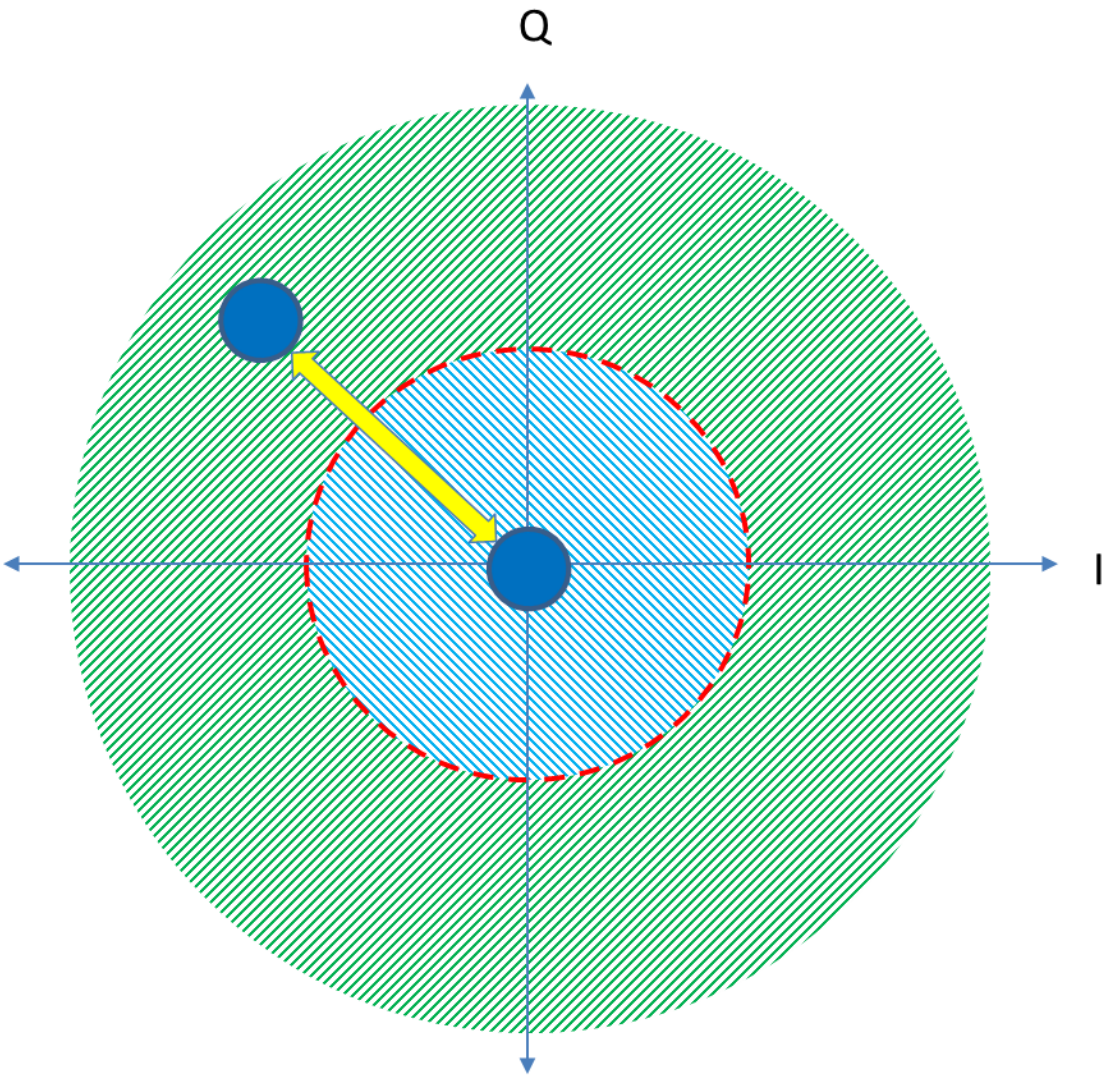

Figure 3.1: Energy Detector Decision Regions

$$\Gamma(N,b) = \frac{1}{(N-1)!}\int_b^\infty x^{N-1}e^{-x}dx \tag{3.2}$$

$$P_{md} = 1 - Q_N\left(\sqrt{\frac{\gamma}{\sigma^2}}, \sqrt{\frac{\lambda}{\sigma^2}}\right) \tag{3.3}$$

Using (3.1), one can solve for a threshold to given a desired probability of false alarm and estimate of the noise variance. That solution is shown in (3.4) where

Γ^{-1} is the inverse normalized upper gamma function. This function is available in Matlab.

$$\lambda = \sigma^2 \Gamma^{-1}(N, P_{fa}) \tag{3.4}$$

The P_{fa} does not vary with signal power, only with noise power and the threshold. As a result of this phenomenon, a common method for determining where to place the threshold is CFAR. In CFAR, the threshold is chosen for an acceptable P_{fa} with respect to some measured noise power at the receiver. The noise at the receiver is expected to be wide-sense stationary. P_{md} depends upon the power of the received signal and the power of the noise; and the power of the received signal will vary. This concept is illustrated in Fig. 3.2.

When comparing the CNN detector against the classic energy detector, it is assumed the energy detector enjoys a perfect noise-floor estimation. In practice, the performance of the energy detector would heavily depend on the accuracy of the noise-floor estimate.

For an energy detector taking 453 observations, with noise at unit variance, and no normalization of the sum of energy, P_{fa} and P_{md} are shown in Table 3.1. The length of the measurement is set to 453 samples because that is the length of the measurements used by the AlexNet detector, and will be explained in section3.4. The "P_{fa}" column in Table 3.1 is the false alarm probility used to compute the threshold. The values in Table 3.1 are calculated directly from equations 3.1 and 3.3. Simulation results of the energy detector are shown in Fig. 3.2. For the simulation, SNR is calculated by adjusting signal power with respect to the noise power. Fig. 3.2 shows the probabilities P_{fa} and P_{md} as a function of the detector threshold for different Signal to Noise Ratios (SNRs) while holding noise power at the receiver constant. In the simulation, the square-root of the average energy was used in order

to compress the x-axis of the Fig. 3.2. The overlap between the P_{fa} and P_{md} curves increases as SNR decreases.

Table 3.1: Missed Detect Rate for the Energy Detector in AWGN

P_{fa} \ SNR	3 dB	0 dB	-3 dB	-6 dB	-9 dB
1e-3%	0.00%	0.00%	0.00%	25.72%	94.90%
0.25%	0.00%	0.00%	0.00%	2.12%	58.91%
1.50%	0.00%	0.00%	0.00%	0.42%	34.80%
8.00%	0.00%	0.00%	0.00%	0.00%	12.90%

Fig. 3.3 shows how the energy detector responds to different levels of SIR. When an interfering signal is added to the received signal, the total power of that received signal increases. The result is that the P_{fa} and P_{md} curves move to the right. The energy detector threshold is sensitive to absolute power, and that threshold will need to be recalculated for every power level of the CW interferer that is encountered in order for the energy detector to maintain a constant false alarm rate or overall accuracy.

The values in Table 3.1 are dependent on absolute power. If the absolute power of the received signal were to change, as would happen if there were an adjustment in the automatic gain control of the analog front-end of the receiver, the P_{fa} and P_{md} curves would shift much like is shown in interference example in Fig. 3.3.

3.4 The CNN Detector

The CNN detector is obtained by retraining the AlexNet CNN, to classify 227×227 sized spectrograms as representing signal plus noise or noise only. The spectrogram is calculated from the signal, and given to the CNN Detector for classification. Each spectrogram is created for complex-valued baseband samples. Those samples are entered into a length 227 delay line. Every new sample in pushes one sample

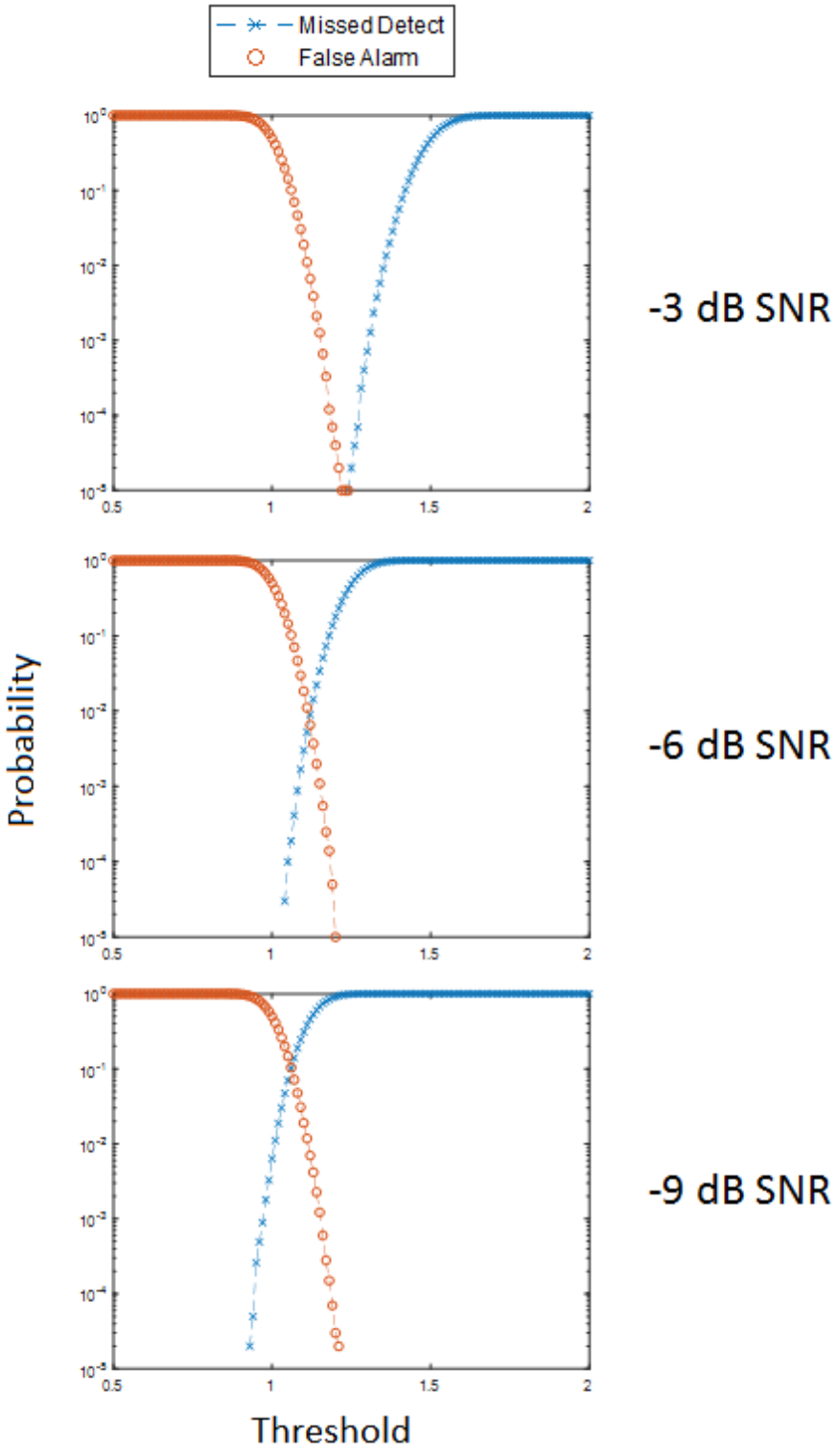

Figure 3.2: Energy Detector Performance

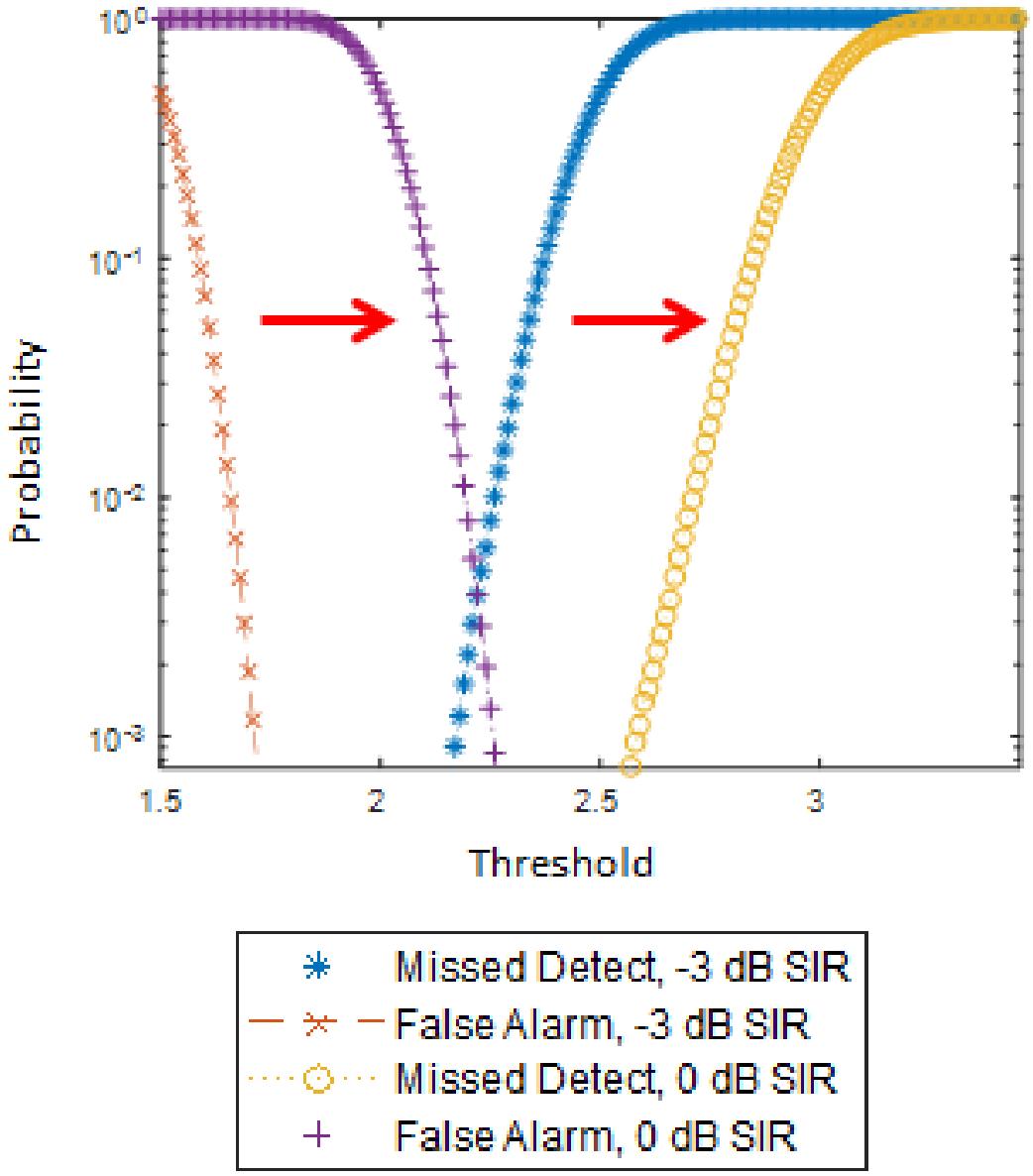

Figure 3.3: Energy Detector Performance in the Presence of Interference

out. The delay line begins full. For each new sample, a 227-length Discrete Fourier Transform (DFT) is taken. The magnitude-squared of each sample from the resulting vector output by the DFT is then taken. That vector of magnitude-squared values produces one row in the 227×227 spectrogram. In total, 453 samples are used in the creation of a spectrogram. The spectrogram is then converted to decibels (dB). The

smallest bin in dB in the spectrogram is set to pixel intensity 0. All dB bin values are then normalized by the largest dB bin value. All normalized dB bin values are then scaled by 255. This yields a spectrogram images with pixel intensities that always range from 0 to 255.

Given that the goal is to detect images in AWGN with low SNRs, it is important to select a CNN that shows some resilience in the face of significant AWGN. Reference [7] compares numerous CNNs against different impairments. The CNNs which did well against AWGN were VGG-Faces, SqueezeNet, and AlexNet. Reference [7] tested other impairments such as brightness. Brightness is a concern in this work due to the limited dynamic range provided by the 8 bits available for pixel intensity. SqueezeNet lagged behind AlexNet when the images were impaired by brightness. After all considerations, AlexNet was chosen for these experiments.

3.5 Training Results in AWGN (Primary-User Detectors)

All training used mini-batches and stochastic gradient descent with momentum. Momentum was set to 0.9. The mini-batch size was set to 10. All training sets are split such that 10% of the training set is reserved as a validation set.

For the two noise experiments, the CNN detectors were trained against images that were created at four SNR values (-9 dB, -6 dB, -3 dB, 0 dB) using AWGN and a BPSK signal oversampled at 4 samples per symbol. For each SNR case, one CNN detector was created and trained. The CNN detector only needed to be trained once in order to be tested in both experiments. Both noise experiments use the same trained CNN detector for a given SNR, but use two different sets of testing images. When training for the two noise experiments, training was substantially slower in the low SNR cases as compared to the high SNR cases. When trained against 0 dB

SNR spectrograms, a total of 800 training images were used in 5 epochs resulting in a 100% validation at the end of training, while training at -9 dB SNR required 1600 training images and 24 epochs with an 82.5% validation accuracy. The CNN detector at -3 dB was trained in 6 epochs settling on 100% validation using 800 training images. The CNN detector trained at -6 dB SNR was trained in 9 epochs settling on 93.75% validation with 800 training images. These values are shown in Table 3.2. Given the falling validation accuracy, -9 dB SNR was taken as the lower limit for these experiments.

For the interference experiment, the CNN detector was trained against 0 dB SNR and a random Signal-to-Interference ratio (SIR) where the CW interferer had a random center frequency for each spectrogram. When training for the interference experiment, 1600 training samples were used and 100% validation accuracy was achieved in 3 epochs. The CNN detector was retrained for 3 epochs on a 0 dB SNR spectrogram with a CW interferer. The power of the CW interferer is a uniformly distributed random variable ranging between 0 (no interferer) and 0 dB SIR (equal to the signal power). The frequency of the CW interferer is a uniformly distributed random variable between -0.5 and +0.5 of the sample rate.

Table 3.2: Training Results

Trained SNR	Epochs	Training Samples	Validation Accuracy
0	5	800	100.00%
-3	6	800	100.00%
-6	9	800	93.75%
-9	9	1600	82.50%

3.6 Testing Results in AWGN (Detecting a Primary-User)

For all experiments, each test case was conducted with 800 unique test images. In each test case, 400 of the images contained a signal and 400 images did not. The SNR or SIR of the images used in a test case are specific to that case. The signal was a BPSK signal oversampled at 4 samples per symbol. The noise, and in some cases random CW interference, was added to the signal before the spectrogram was taken. Images without a signal were produced from noise and in some cases random CW interference.

3.6.1 Detection in the Presence of a Fixed Noise Floor

The first test for the CNN Detector was to characterize the performance of the detector in AWGN. In the training phase, four separate CNN detectors were trained, each focused on a different target SNR. Each of those CNN detectors was tested with new spectrograms. The tests measured the rate of missed detects and false alarms across SNR values of 3 dB, 0 dB, -3 dB, -6 dB and -9 dB.

The results for the P_{fa} are shown in Table 3.3. Each row is a *trained value* (Train) meaning that specific CNN detector was trained against spectrograms with that SNR. The column headers are for *test values* (Test) meaning that a specified CNN detector was tested against spectrograms with that SNR. The CNN detector behaves as a CFAR detector where the false alarm rate is determined by the SNR against which it was trained.

The results for P_{md} are shown in Table 3.4. Comparing Table 3.4 and Table 3.1 shows that the perfect energy detector exhibits better performance at a threshold related to the CFAR exhibited by the CNN detector. Much like the energy detector, P_{md} at low SNRs improves as P_{fa} is allowed to worsen.

Table 3.3: False Alarm Rate for the CNN Detector in AWGN

Train \ Test	3 dB	0 dB	-3 dB	-6 dB	-9 dB
0 dB	0.00%	0.00%	0.00%	0.00%	0.00%
-3 dB	0.75%	0.00%	0.25%	0.25%	0.00%
-6 dB	1.75%	1.00%	1.25%	2.00%	1.50%
-9 dB	6.50%	7.75%	8.00%	8.25%	8.50%

Table 3.4: Missed Detect Rate for the CNN Detector in AWGN

Train \ Test	3 dB	0 dB	-3 dB	-6 dB	-9 dB
0 dB	0.00%	0.25%	12.50%	69.00%	96.25%
-3 dB	0.00%	0.00%	0.25%	30.00%	79.75%
-6 dB	0.00%	0.00%	0.00%	10.75%	55.50%
-9 dB	0.00%	0.00%	0.00%	1.25%	21.50%

3.6.2 Primary-User Detection in the Presence of a Variable Noise Floor

In this experiment, the CNN detector was tested on spectrograms resulting from signals with the same SNR but varying absolute values of noise power resulting in different noise floor levels. The goal of this experiment was to determine if the AlexNet detector could maintain the same performance across different absolute values of noise without an explicit noise estimate. No retraining was performed. Because the SNR was constant, a change to the absolute value of the noise requires the same change in dB to the absolute value of the power of the signal. The CNN detector has no such dependency on absolute values and continues to operate on the spectrograms. The results of this experiment are shown in Table 3.5. The CNN detector used for this experiment was the one trained against -3 dB SNR. The spectrograms for testing where all at 0 dB SNR.

The CNN detector results in Table 3.5 demonstrate near identical performance to

Table 3.4 and Table 3.3. This is quite different from the Energy Detector which must change the threshold in order to maintain an optimal overall accuracy or CFAR to accommodate a changes to the absolute power levels. It is clear that the CNN detector has no such dependency on absolute power.

Table 3.5: Variable Noise Power Results

Noise Variance	Missed Detect	False Alarm
0.5 (-3 dB)	0.00%	0.00%
1 (0 dB)	0.00%	0.00%
2 (3 dB)	0.25%	0.50%
4 (6 dB)	0.50%	0.00%

3.6.3 Primary-User Detection in the Presence of Noise and Interference

In the last experiment, the CNN detector was tested against an environment with interference. Spectrograms without a signal were created from a CW tone that was added to AWGN. Three cases of CW interferer power were tested, those being CW interferer power equal to the signal (0 dB SIR), random CW interferer power, and no CW interferer. Fig. 3.4 shows two conditions side by side. One spectrogram in Fig. 3.4 has only noise and the other has noise and a CW tone, which is the white line on the left. The center frequency of the CW interferer was random, ranging across the entire band in the spectrogram. For any one spectrogram, the CW interferer would be of one absolute power and one center frequency. For a spectrogram with a signal, a CW tone and AWGN were added to the BPSK signal (oversampled by 4).

The results are shown in Table 3.6. An SIR of ∞ dB indicates there was no interferer in that test case. One CNN detector, trained on a variety of CW interferer powers and frequencies, can distinguish spectrograms containing signals whether the interferer is present or not. The CNN detector does not need to be retrained if

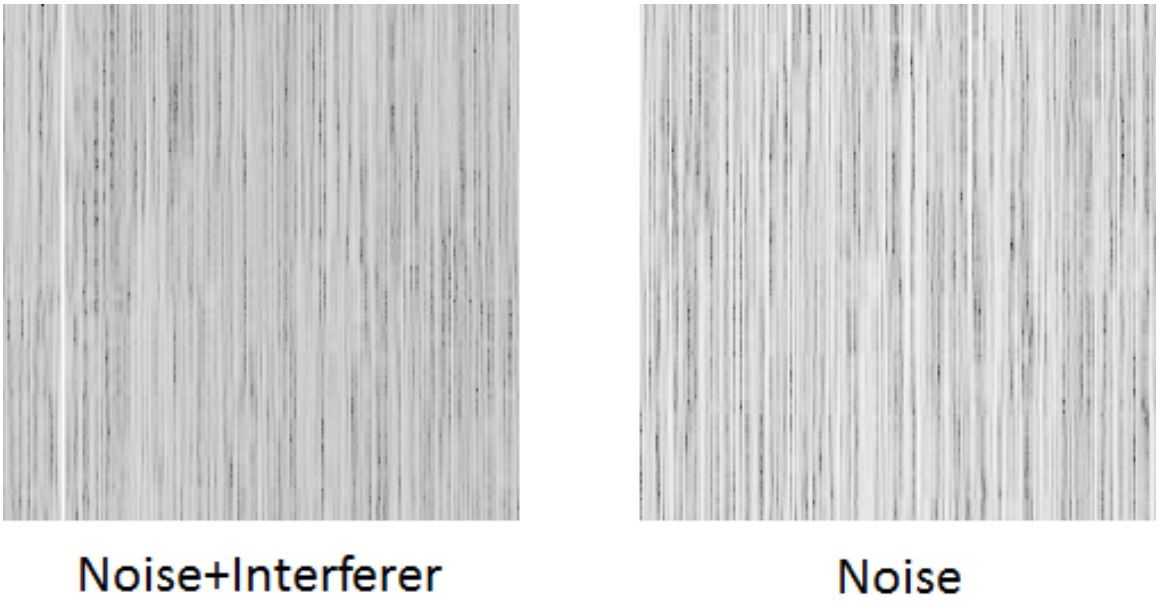

Figure 3.4: Noise and Interference

the CW interferer goes away or re-appears.

Table 3.6: Interference Results

SIR	Missed Detect	False Alarm
0 dB	1.00%	0.00%
Random	0.50%	0.00%
∞ dB	1.50%	0.00%

3.7 Conclusion of the Primary-User Detection Tests

The CNN detector demonstrates strong advantages over the classic energy detector. The energy detector's threshold can be tuned for a specific set of absolute power levels, but for a range of absolute power levels, the CNN detector shows better performance. The energy detector's threshold can be tuned to a constant false alarm rate corresponding to a specific noise floor measurement, but this requires an accurate noise floor measurement. Because the received signal will vary in

power, attempting to normalize the variance of the input to the energy detector will cause a change to the absolute power of the noise floor and thus break the constant false alarm rate. The CNN detector has no such dependency on absolute power at the receiver. The CNN detector achieves CFAR performance by being trained on specific SNRs, and does not require an estimate of the noise floor.

The CNN detector demonstrated superior performance as compared to the energy detector in the presence of an interferer, and did not require any parameterization of that interferer. The threshold for the energy detector moved with the power of the interferer whether the threshold be set for overall accuracy or constant false alarm.

When compared the results of these experiments to those reported in [8], calculations converting SNR to energy per sample over the power spectral density of noise, E_s/N_0, must take oversampling into account. Without band-limiting after digitization, E_s/N_0 will be 10log10 (OSR) dB higher than the reported SNR; where OSR is the rate by which the data is oversampled. This is due to the increased noise power commensurate with the wider Nyquist bandwidth resulting from oversampling. The CNN used in [8] detected the BPSK signal with 70% accuracy at -20 dB SNR while oversampling by 100. This means the detector had a 70% probability of signal detection at 0 dB E_s/N_0. The CNN detector oversampled the BPSK signal by 4, therefore an SNR of -6 dB corresponds to an E_s/N_0 of 0 dB, at which point the CNN detector had a 98.75% probability of signal detection when trained against -9 dB SNR.

In addition to the advantages in performance as a signal detector, the CNN detector demonstrates how fine tuning can quickly re-purpose an existing CNN to perform a new task seemingly unrelated to the original task. With very few samples, and on readily available desktop computing hardware, a CNN as complex as AlexNet can be redeployed to numerous other tasks by way of fine tuning.

3.8 Detection in Cancellation Residue

The following sections will detail the training and testing of the AlexNet detector for detecting the presence of secondary users in cancellation residue. The residue is created as described in chapter 2 and [3]. An Energy detector and a Cyclostationary detector establish a means by which to compare the performance of the AlexNet detectors. Two of the AlexNet detectors from the preceding sections are re-used for these experiments, those being the detectors trained at -3 and -6 dB SNR. In addition, two secondary-user detectors are trained on cancellation residue impaired by imperfections in the frequency, channel, and payload estimates as originally described in chapter 2 and [3].

The secondary-user detectors were each trained on 1600 training samples. As in section 3.5, the training used mini-batches and stochastic gradient descent with momentum. Momentum was set to 0.9. The mini-batch size was set to 10. The sampling rate was 20 MSPS, and thus the covert signal was sampled at 2 samples per chip. Both detectors were trained for 6 epochs. Two detectors were trained, one with residue produced with an OFDM signal at 25 dB SNR and the other at 27 dB SNR. Both had the covert signal set to 31 dB SNR. The AlexNet detector trained on the 25 dB SNR residue had a final validation accuracy of 80%. The AlexNet detector trained on the 27 dB SNR residue had a final validation accuracy of 88.75%

3.8.1 Determining the Range of Test Cases

The goal of these measurements is to find the effective ranges of the detectors. Three test cases are presented in this section and are defined by pairs of SNR and SIR values. SNR refers to the Signal to Noise ration of the incumbent OFDM signal. SIR refers to the Signal to Interfere Ratio measured as the ratio of the power of the OFDM signal to the power of the covert signal. The Interference to Noise Ratio

(INR) is the ratio of the power of the covert signal to the power of the noise and that is calculated in decibels as $INR = SNR - SIR$. The values in Table 3.7 are provided for convenience.

Because the energy detector does not normalize the power of the input, this experiment was designed in a manner similar to section 3.6.2 where the power of the input is varied. As in [3] and [2], a noise estimate is not being generated for the energy detector.

The cyclostationary detector relies on the cycle-frequency domain profile which is extracted from the spectral coherence as described in [10]. Therefore the features used by the cyclostationary detector in this experiment are normalized.

In these tests, ROC curves are produced by varying the detector threshold across the signal-present and signal-not-present distribution of measurements. These tests were run using a packet of 1000 bytes for a total of 3440 samples.

Table 3.7: INR values for SNR,SIR pairs

SNR (dB)	SIR (dB)	INR (dB)
25	31	-6
25	33	-8
23	31	-8

In the first test case with an SNR of 25 dB and an SIR of 31 dB, as shown in Fig. 3.5, the cyclostationary detector with cancellation performs well yielding a ROC curve with a considerable area under the curve. This means that a low probability of false alarm can be selected and still maintain a high probability of detection. The energy detector does not show much promise without or without the cancellation, although cancellation does yield some improvement to the ROC curve of the energy detector.

The second test case uses an SNR of 25 dB and an SIR of 33 dB.The INR of this

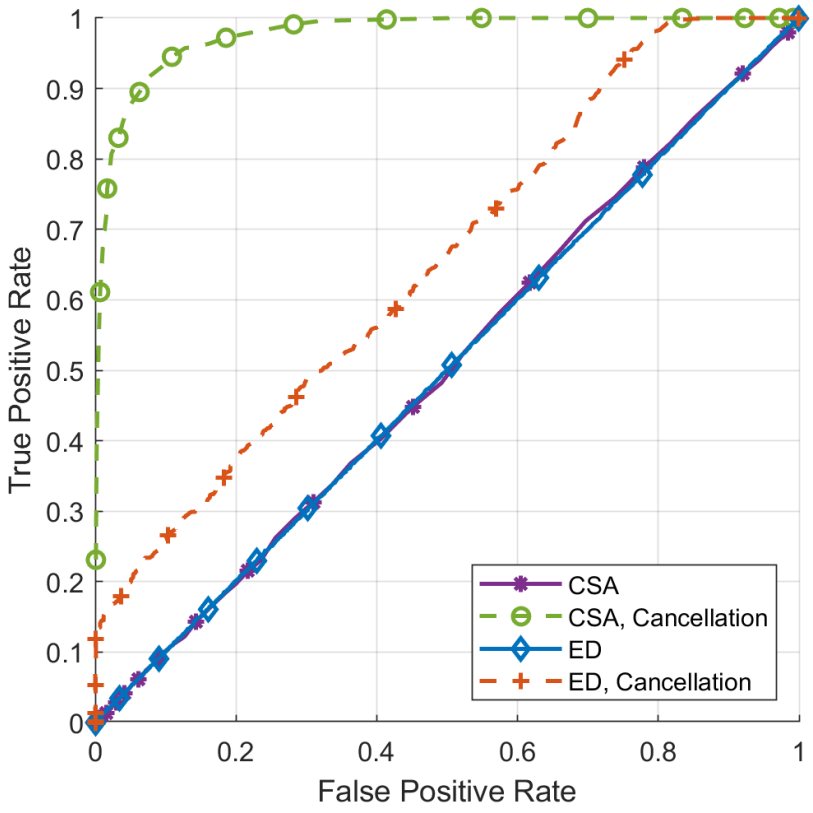

Figure 3.5: ROC for SNR 25 dB, SIR 31 dB

test case is -8 dB, meaning that the covert signal will be 8 dB below the noise floor as opposed to the INR of -6 dB in the previous test case. The results are shown in Fig. 3.6. The ROC curve of the cyclostationary detector with cancellation begins to fall back toward the diagonal line. This is due to the reduction in power of the covert signal. The covert signal is 8 dB below the noise floor.

In the third test case the SIR is restored to 31 dB but the SNR of the OFDM signal is reduced to 23 dB. The INR is -8 dB just like in the second test case, but now bit and packet errors are more likely due to the lower SNR of the OFDM signal. The ROC curve is plotted in Fig. 3.7 and that result is comparable to the second test case plotted in Fig. 3.6.

These experiments have shown that the energy detector fails to detect the covert

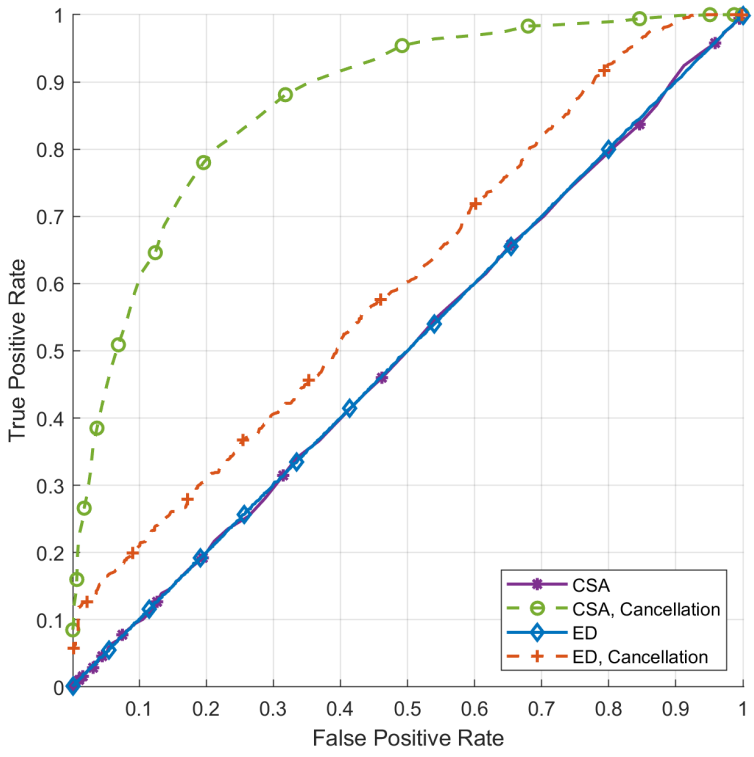

Figure 3.6: ROC for SNR 25 dB, SIR 33 dB

signal at all, and that the performance of the cyclostationary detector begins to plummet around -8 dB INR. These results provide a range of test cases for the AlexNet experiments. The AlexNet test will focus on testing at -6 dB INR with a 25 dB SNR and a 31 dB SIR and -8 dB INR with a 23 dB SNR and a 31 dB SIR.

3.8.2 AlexNet Detector Performance in Imperfect Residue

In this experiment, the AlexNet detectors were run against residue from *imperfect cancellation*. What makes the cancellation imperfect is there are errors in the estimation of the exact waveform to be canceled. These errors can be bit errors from incorrect demodulation, errors in the frequency and phase estimates, and errors in the channel estimates.

The results for False Alarm are shown in Table 3.8, and the results for Missed

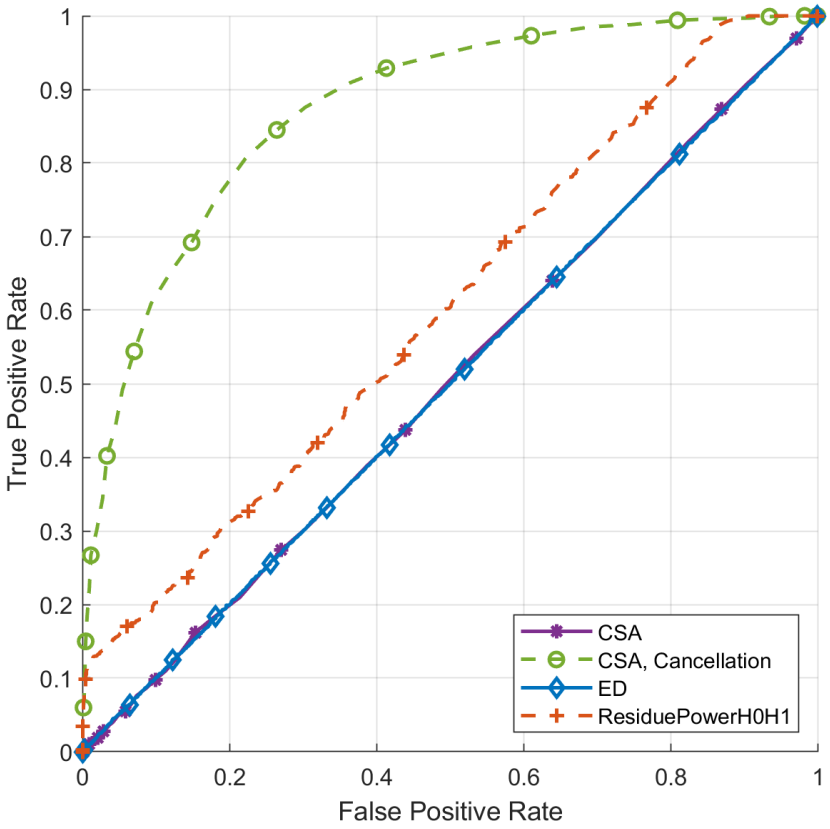

Figure 3.7: ROC for SNR 23 dB, SIR 31 dB

Detect are shown in Table 3.9. Each row in the tables gives results for one of the previously defined AlexNet detectors. The columns indicate the SNR and SIR of the test case where 25/31 is an SNR of 25 dB and an SIR of 31 dB.

Table 3.8: False Alarm Rate for the AlexNet Detector in Imperfect Residue

Label \ Test	25/31	23/31	25/33
-3 dB	0.50%	1.125%	0.75%
-6 dB	0.125%	0.125%	0.50%
25 dB Imp	8.375%	8.5%	7.875%
27 dB Imp	7.375%	7.00%	6.75%

Table 3.9: Missed Detect Rate for the AlexNet Detector in Imperfect Residue

Label \ Test	25/31	23/31	25/33
-3 dB	82.63%	91.88%	92.50%
-6 dB	84.50%	93.00%	93.75%
25 dB Imp	35.375%	49.25%	50.5%
27 dB Imp	36.625%	47.25%	49.75%

3.9 Analysis of the Secondary-User Detection Results

The experiment with the secondary-user detectors was designed to answer several research questions: (i) how well can the detectors perform in the imperfect cancellation residue; (ii) how well do the detectors transfer across those distributions; and (iii) whether or not the CFAR behavior of the detectors is maintained between operating in AWGN and the cancellation residue.

The secondary-user AlexNet detectors outperformed the energy detector with cancellation, even though the energy detector had 3440 samples and the AlexNet detector can only use 453 samples. The cyclostationary detector outperformed the AlexNet detector, however it was able to use far more samples. This is one area where the AlexNet detector does not do well: it is not as easily scalable in input size as are the energy and cyclostationary detectors. The detectors do not transfer well outside of the distribution of the noise in which they were trained. The two AlexNet detectors trained in AWGN did not perform well in the residue. This shows that an AlexNet detector trained in a distribution of noise must remain in that distribution of noise. The AlexNet detectors did not require a noise floor estimate in cancellation residue. The AlexNet detectors maintained their CFAR-like performance in cancellation residue.

3.10 Future Work

It may be possible to repeat these successes in signal detection with a CNN less complex than AlexNet. Reducing the computational complexity of the CNN detector is one avenue of future work. For example, the cyclostationary detector outperformed the AlexNet detector in cancellation residue, however the cyclostationary detector is computationally costly. If a simpler network could be employed, that network could outperform the cyclostationary detector from the perspective of computational cost. A CNN detector could be expanded to tackle other interference sources and jamming threats. This would require identifying the interference, simulating the interference, and training the CNN detector to ignore the interference.

3.11 References

[1] Daniel Chew, Andrew L. Adams, and Jason Uher. "Secondary Spectrum Usage and Signal Detection". In: *Wireless Coexistence: Standards, Challenges, and Intelligent Solutions*. 2021, pp. 115–154. DOI: 10.1002/9781119584230.ch5.

[2] Daniel Chew and A. Brinton Cooper. "Spectrum Sensing in Interference and Noise Using Deep Learning". In: *2020 54th Annual Conference on Information Sciences and Systems (CISS)*. 2020, pp. 1–6. DOI: 10.1109/CISS48834.2020.1570617443.

[3] Daniel Chew et al. "Covert Communications through Imperfect Cancellation". In: *Accepted to Proceedings of the 2022 ACM Workshop on Information Hiding and Multimedia Security*. IH&MMSec '22. New York, NY, USA: Association for Computing Machinery, 2022.

[4] Fadel F. Digham, Mohamed-Slim Alouini, and Marvin K. Simon. "On the Energy Detection of Unknown Signals Over Fading Channels". In: *IEEE Transactions on Communications* 55.1 (2007), pp. 21–24. DOI: 10.1109/TCOMM.2006.887483.

[5] P.P. Gandhi and V. Ramamurti. "Neural networks for signal detection in non-Gaussian noise". In: *IEEE Transactions on Signal Processing* 45.11 (1997), pp. 2846–2851. DOI: 10.1109/78.650111.

[6] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. MIT Press, 2016.

[7] Klemen Grm et al. "Strengths and weaknesses of deep learning models for face recognition against image degradations". In: *IET Biometrics* 7.1 (2018), pp. 81–89. DOI: https://doi.org/10.1049/iet-bmt.2017.0083. eprint: https://ietresearch.onlinelibrary.wiley.com/doi/pdf/10.1049/iet-bmt.2017.0083. URL: https://ietresearch.onlinelibrary.wiley.com/doi/abs/10.1049/iet-bmt.2017.0083.

[8] Dong Han et al. "Spectrum sensing for cognitive radio based on convolution neural network". In: (2017), pp. 1–6. DOI: 10.1109/CISP-BMEI.2017.8302117.

[9] John D. Kelleher, Brian Mac Namee, and Aoife D'Arcy. *Fundamentals of Machine Learning for Predictive Data Analytics: Algorithms, Worked Examples, and Case Studies*. The MIT Press, 2015. ISBN: 0262029448.

[10] Kyouwoong Kim et al. "Cyclostationary Approaches to Signal Detection and Classification in Cognitive Radio". In: *2007 2nd IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks*. 2007, pp. 212–215. DOI: 10.1109/DYSPAN.2007.35.

[11] J.D. King. "Fixing spectrum auctions". In: *IEEE Spectrum* 55.39 (2018).

[12] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. "ImageNet Classification with Deep Convolutional Neural Networks". In: *Commun. ACM* 60.6 (2017), 84–90. ISSN: 0001-0782. DOI: 10.1145/3065386. URL: https://doi.org/10.1145/3065386.

[13] Z.H. Michalopoulou, L.W. Nolte, and D. Alexandrou. "Performance evaluation of multilayer perceptrons in signal detection and classification". In: *IEEE Transactions on Neural Networks* 6.2 (1995), pp. 381–386. DOI: 10.1109/72.363473.

[14] Marko Palola et al. "The first end-to-end live trial of CBRS with carrier aggregation using 3.5 GHz LTE equipment". In: *2017 IEEE International Symposium on Dynamic Spectrum Access Networks (DySPAN)*. 2017, pp. 1–2. DOI: 10.1109/DySPAN.2017.7920763.

[15] D.W. Ruck et al. "The multilayer perceptron as an approximation to a Bayes optimal discriminant function". In: *IEEE Transactions on Neural Networks* 1.4 (1990), pp. 296–298. DOI: 10.1109/72.80266.

[16] Carl R. Stevenson et al. "IEEE 802.22: The first cognitive radio wireless regional area network standard". In: *IEEE Communications Magazine* 47.1 (2009), pp. 130–138. DOI: 10.1109/MCOM.2009.4752688.

[17] Yu-Jie Tang, Qin-Yu Zhang, and Wei Lin. "Artificial Neural Network Based Spectrum Sensing Method for Cognitive Radio". In: *2010 6th International Conference on Wireless Communications Networking and Mobile Computing (WiCOM)*. 2010, pp. 1–4. DOI: 10.1109/WICOM.2010.5601105.

[18] H. Urkowitz. "Energy detection of unknown deterministic signals". In: *Proceedings of the IEEE* 55.4 (1967), pp. 523–531. DOI: 10.1109/PROC.1967.5573.

[19] Leiou Wang, Donghui Wang, and Chengpeng Hao. "Intelligent CFAR Detector Based on Support Vector Machine". In: *IEEE Access* 5 (2017), pp. 26965–26972. DOI: 10.1109/ACCESS.2017.2774262.

[20] J.W. Watterson. "An optimum multilayer perceptron neural receiver for signal detection". In: *IEEE Transactions on Neural Networks* 1.4 (1990), pp. 298–300. DOI: 10.1109/72.80267.

Chapter 4

OFDM Window Estimation for Interference Cancellation

4.1 Introduction

Windowing at the transmitter is a popular means to control the bandwidth of an OFDM signal. This windowing is typically implemented by extending the cyclic prefix is into the previous OFDM symbol and adding a cyclic suffix that extends into the next OFDM symbol. These two extensions are tapered by multiplying the samples with a windowing function. The effect is a tapered transition from one OFDM symbol into another. The benefit of these tapered extensions is to limit the bandwidth of the OFDM signal in a way that is less expensive than direct filtering. The benefits of windowing at the transmitter are explored in detail in [3]. Windowing at an OFDM transmitter is often optional, not required, by a wireless standard. For example, the IEEE 802.11 standard suggests windowing to limit the OFDM bandwidth and provides a window definition, but does not mandate the use of it [6]. The effects of the recommended IEEE 802.11 window are documented in [15]. The use, shape, and length of a window at the transmitter is left to individual vendors in the 802.11 standard. The effects of windowing at the transmitter are kept in check by the modulation accuracy requirements of the standard.

4.1.1 Effects of OFDM Windowing at the Transmitter on PER

The downside of these extensions into adjacent symbols is that they constitute self-imposed inter-symbol interference, reducing the resilience of the OFDM signal to multipath channels. The effect of OFDM windowing at the transmitter as self-interference on multipath immunity was explored in [5] and [7] which propose "orthogonality restoration" to remove the self-interference by cancellation and improve system performance. In order for this cancellation to work, the windowing term is included when reconstructing the interference. Both [5] and [7] assume the windowing function is known in advance, thus the window is not estimated.

4.1.2 Multi-User Interference Cancellation Considerations for OFDM Windowing at the Transmitter

Multi-User Interference Cancellation (MUIC) is a well-known concept that enables wireless users to address interference caused by multiple users accessing a single spectrum resource concurrently. MUIC is the process of modeling and reproducing a signal from one particular user for the purposes of removing that signal from the summed ensemble of all received signals. Successive Interference Cancellation (SIC) [9] is a category of MUIC implementations in which cancellation is applied sequentially over multiple users in the ensemble.

The use of MUIC requires that the receiver be able to cancel a portion of the received signal. That assumes that the receiver has a sufficient model of the signal, and can estimate the parameters of that model with sufficient accuracy. Several models appear in literature. All the models begin with demodulating the signal and then remodulating. The remodulated signal is then augmented with estimated impairments for a better match at cancellation. This process is described in [9] but it does not define the parameters to be estimated. The signal model employed for cancellation in [2] estimates the channel coefficients. The signal model in [10]

employs estimates for channel coefficients and "inter-channel interference" (ICI) meaning wireless impairments have caused the subcarriers of the OFDM symbol to no longer be orthogonal. The signal model in [8] estimates amplitude and phase. These signal models for cancellation may be insufficient. The work in this chapter shows that OFDM windowing at the transmitter has a significant impact on cancelling the OFDM signal.

4.1.3 Contributions of this Chapter

The work in this chapter shows that the window applied at the OFDM transmitter can be estimated from the received samples. That estimate can then be used to remove the aforementioned self-interference or improve over-all cancellation for MUIC applications. Despite uncertainty in the estimate of the window, it is shown that this signal parameter can offer significant SIC performance improvements. The goal is to estimate a sufficient number of signal parameters to create a copy of the OFDM with enough accuracy to provide significant suppression of the OFDM signal.

OFDM symbols from IEEE 802.11 are used as a working example without loss of generality. Multiple examples of windows are used following the recommendation in the 802.11 standard. The estimation of the window is detailed and applied to improve the PER as selectively applied cancellation can remove self-interference. The estimate of the window is also applied to cancel whole OFDM packets. The novel contributions of this chapter are as follows:

- The work in this chapter shows that the windowing parameter in OFDM has a significant effect on cancellation performance.
- The estimation of the OFDM window is detailed in this chapter. The OFDM window is estimated without assuming window symmetry. That window

estimate is then be used to cancel the self-interference caused by windowing.

- In this chapter it is not assumed that the OFDM transmitter has special implementation, meaning one cannot rely on any symbol alignment, common OFDM numerology, superposition coding as in [14], or some other means implemented to the ease retrieval of lower power users.

4.1.4 Organization of this Chapter

The structure of this chapter is as follows: Section 4.2 develops a signal model for cancellation similar to that in [2] using estimates of the channel coefficients and the center frequency offset (CFO), without accounting for windowing. Section 4.3 develops a signal model for OFDM with windowing. Section 4.4 develops an algorithm to estimate the window which includes the presence of a multipath channel. The RMS error of the algorithm is plotted as functions of the number of OFDM symbols used to generate the estimate and as a function of the power of noise. Section 4.5 applies the algorithms from in sections 4.2 and 4.4 to synthetic OFDM packets and shows that estimating the window provides significant improvement in cancellation. Section 4.6 describes an experiment applying the algorithms in sections 4.2 and 4.4 to cancel OFDM signal data collected over-the-air. It is shown that the new method provides a 5.3 dB improvement over the method in section 4.2. The cancellation is then applied to the covert signal established in chapter 5. It is shown that the cancellation algorithm enhanced with a window estimate decreases the BER of the covert signal, outperforming the original design. Lastly, it is shown that using the enhanced cancellation algorithm to selectively cancel the self-interference of the cyclic extensions results in an improved PER for the OFDM waveform itself. Section 4.7 summarizes the results.

4.2 OFDM Signal Model without Windowing

An OFDM symbol is created using an M-length Inverse Fast Fourier Transform (IFFT) represented in (4.1) where M represents the total number of subcarriers including all pilots and nulls. An OFDM packet is a concatenation of N OFDM symbols in time, where each individual symbol is denoted $o_p[k]$. The subscript p indicates the placement of an OFDM symbol in the packet and k indexes the individual samples which comprise that OFDM symbol. The data frame d_p in (4.1) represents all data to go into the IFFT for any one OFDM symbol. A single OFDM symbol is length K and is at least $M + L$ long where L is the mandatory cyclic prefix length and $L < M$. As an example, an 802.11g OFDM symbol uses a 64-length IFFT (M=64) and an OFDM symbol length of 80 (K=80) after a cyclic prefix of 16 (L=16). The range $-L \leq k \leq -1$ represents the mandatory cyclic prefix and the range $0 \leq k \leq M - 1$ represents the original length-M IFFT.

$$o_p[k] = \sum_{m=0}^{M-1} d_p[m] e^{j\frac{2\pi mk}{M}} \tag{4.1}$$

An OFDM packet is a concatenation of N OFDM symbols in time, where each individual symbol is denoted $o_p[k]$ as represented in (4.2). The subscript p indicates the placement of an OFDM symbol in the packet and k indexes the individual samples which comprise that OFDM symbol.

$$\vec{s} = \{\vec{o}_0 || \vec{o}_1 ... || \vec{o}_{N-1}\} \tag{4.2}$$

4.2.1 Channel and Frequency Offset Impairments

The OFDM packet as seen by the receiver $\vec{r}$ after timing offset correction is shown in (4.3) where matrix H represents the multipath channel, the diagonal matrix Λ_Θ represents the carrier frequency offset, and $\vec{n}$ noise at the receiver. The carrier offset

diagonal matrix Λ_Θ contains phase offsets for each sample of $\vec{s}$ as shown in (4.4). If the carrier frequency offset is constant then when integrated in time it will produce a phase ramp in Λ_Θ. The magnitude of each element is unity. Each k^{th} element on the diagonal represents a phase offset value $\Theta[k]$ shown as a phase ramp in (4.5).

$$\vec{r} = \Lambda_\Theta H\vec{s} + \vec{n} \tag{4.3}$$

$$\Lambda_\Theta = \text{diag}(e^{j\Theta}) \tag{4.4}$$

$$\Theta[k] = \omega k + \phi \tag{4.5}$$

In order to correct the impairments, the corrections derived from estimates of those impairments must be applied to the received samples. Once the first sample of the 802.11 packet has been determined, estimates are needed for Λ_Θ and H, those being $\widehat{\Lambda_\Theta}$ and $\widehat{H}$. After those parameters are estimated, the conjugate and inverse of those estimates $\widehat{\Lambda_\Theta}^*$ and $\widehat{H}^{-1}$ will be found. However, these impairment estimations are imperfect and there will be some error after the corrections are applied.

The first impairment to be estimated is the CFO impairment defined in (4.4). The frequency correction estimate $\widehat{\Lambda_\Theta}^*$ is applied to the received samples $\vec{r}$ before an estimate of the channel matrix can be calculated. Therefore, this error propagates into the estimation of the channel impulse response. The estimate of the channel impulse response is initially calculated using the known sequence in the 802.11 preamble. This only provides estimates for the 52 non-zero subcarriers. The channel estimate is linearly interpolated over the null-subcarriers in phase and magnitude. The inverse of the channel estimate $\widehat{H}^{-1}$ is calculated and used to equalize the received signal. The channel estimate is imperfect and therefore the equalization

will be imperfect.

4.2.2 Applying Cancellation

The re-modulation produces an OFDM waveform $\widehat{s}$ representing the estimate of the OFDM signal at the transmitter with no impairments. Applying the corrections to the received waveform may shape the noise and may adversely affect signals from other users. Therefore, the estimated channel $\widehat{H}$ and the estimated carrier frequency offset $\widehat{\Lambda_\Theta}$ are applied to $\widehat{s}$ as shown in (4.6).

$$\vec{u} = \Lambda_\Theta H\vec{s} + \vec{n} - \widehat{\Lambda_\Theta}\widehat{H}\widehat{s} \tag{4.6}$$

Assuming the demodulation process saw no bit errors, then $\widehat{s} = \vec{s}$ and the residue reduces to (4.7).

$$\vec{u} = (\Lambda_\Theta H - \widehat{\Lambda_\Theta}\widehat{H})\vec{s} + \vec{n} \tag{4.7}$$

The signal-error-term of the residue is $(\Lambda_\Theta H - \widehat{\Lambda_\Theta}\widehat{H})\vec{s}$. As the estimation functions improve, the estimated parameters approach the actual parameters, $\widehat{\Lambda_\Theta}\widehat{H} \rightarrow \Lambda_\Theta H$. As that happens, the signal-error term goes to zero. This would leave only the noise term in the residue. It is not expected that the parameter estimates will be perfect. The cancellation technique is expected to be imperfect and it is expected that some remainder of the OFDM signal to be present in the residue. Furthermore, these two impairments may be insufficent as will be shown in the subequent sections.

4.3 OFDM Windowing at the Transmitter

Different windowing schemes have been explored in literature, such as [3] [11] [4] [13] and others. Here the general case is analyzed. A generalized model for the windowing function w_i is provided in (4.8) where i is the sample index, L is the length of the cyclic prefix and suffix in samples, and M is the IFFT length in samples. The variable i is being used instead of k because the application of the window to the OFDM signal is periodic. The length L is generally set the longest multipath channel in which the OFDM signal is expected to operate. Note that the window definition in (4.8) does not assume that the windowing is symmetric. There are two transition regions in the window definition defined by two separate sets of coefficients α_i and β_i. The transition from 1 to 0 at the end of a symbol, including the cyclic suffix, is represented by the coefficients α_i. The transition from 1 to 0 in the cyclic prefix, including any extension thereof, is represented by the coefficients β_i.

The window definition in (4.8) does set some practical limits on the window, without loss of generality. The window is only given nonzero values for $-2L \leq i \leq M + L - 1$. The window values are 1 for $0 \leq i \leq M - L - 1$. The indices for α and β range over $0 \leq i \leq 2L - 1$. The length of the α and β coefficients is set to $2L$; however, some of these coefficients may be zero or one.

$$w_i = \begin{cases} 0 & \text{for } i > M + L - 1 \\ \alpha_{i-M+L} & \text{for } M - L \leq i \leq M + L - 1 \\ 1 & \text{for } 0 \leq i \leq M - L - 1 \\ \beta_{-1-L-i} & \text{for } -2L \leq i \leq -1 \\ 0 & \text{for } i < -2L \end{cases} \tag{4.8}$$

The definition in (4.8) limits the nonzero values in a window applied to any one OFDM symbol to a range $-2L \leq i \leq M - L - 1$. This imposes a maximum length of $M + 3L$ on the extended OFDM symbol where $M + L$ represents the standard

length of the OFDM symbol, L represents the maximum extension of the cyclic prefix, and the final L term represents the maximum length of the cyclic suffix. The index into α is $i - M + L$ where the offset $-M + L$ represents the start of that transition. The index into β is $-1 - L - i$ where the term $-i$ reverses the order of the coefficients and where the offset $-1 - L$ represents the optional extension of the cyclic prefix.

Windowing is illustrated in Fig. 4.1. The figure shows three copies of an OFDM symbol $o[k]$ each of M samples. The repetitions on either end form the cyclic prefix and suffix. The cyclic prefix is longer than the suffix because the cyclic prefix has a mandatory minimum length of L. The cyclic suffix and extended cyclic prefix must protrude into the adjacent OFDM symbol.

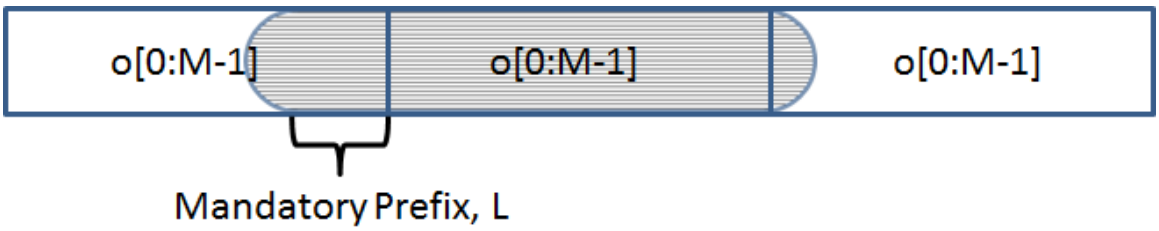

Figure 4.1: Cyclic Prefix and Suffix as Repeating Symbols and Window

The IEEE 802.11 standard contains a windowing recommendation that defines a window transition period spanning equally across the boundary between two OFDM symbols adjacent in time. One half of that transition window extends the cyclic prefix past 16 samples and into the previous OFDM symbol. The other half of the transition window creates a cyclic suffix which extends into the cyclic prefix of the next OFDM symbol. This is illustrated in Fig. 4.2 where a transition region is shown between the two sequential and overlapping OFDM symbols.

The windowed OFDM symbols are defined in (4.9). The sample index k is relative to the current OFDM symbol o_p. This is why the previous and next OFDM symbols, $o_{(p-1) \bmod N}$ and $o_{(p+1) \bmod N}$ require offsets in the sample index. The

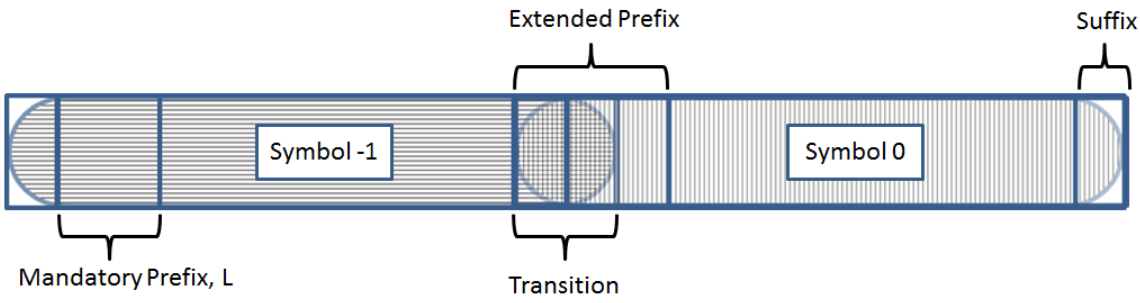

Figure 4.2: Cyclic Prefix and Suffix Overlapping

window coefficients used on an OFDM symbol are selected by the sample index plus offset applied to that OFDM symbol. Three OFDM symbols are windowed and then summed to create a combined symbol v_p. With no windowing v_{-1} and v_1 are always zero, and thus $\vec{v} = \vec{s}$ where $\vec{s}$ is defined in (4.2). This process introduces self-interference from $o_{(p-1) \bmod N}$ and $o_{(p+1) \bmod N}$ into o_p. The summation in (4.9) can be compressed into (4.10) where q is an integer $-1 \leq q \leq 1$ representing the previous, current, and next OFDM symbol.

$$\begin{aligned} v_p[k] = & \\ & w_{k+M+L} o_{(p-1) \bmod N}[k+M+L] \\ & + w_k o_p[k] \\ & + w_{k-M-L} o_{(p+1) \bmod N}[k-M-L] \end{aligned} \tag{4.9}$$

$$\begin{aligned} v[k] = \sum_{q=-1}^{1} \{ & w[k-q(M+L)] \\ & o_{(p+q) \bmod N}[k-q(M+L)] \} \end{aligned} \tag{4.10}$$

Fig. 4.3 illustrates the creation of v_p and the resulting self-interference. The cyclic prefix of o_{p+1} extends into the end of o_p. The cyclic suffix of o_{p-1} extends into the cyclic prefix of o_p. Therefore OFDM symbol o_p has interference from both

adjacent symbols.

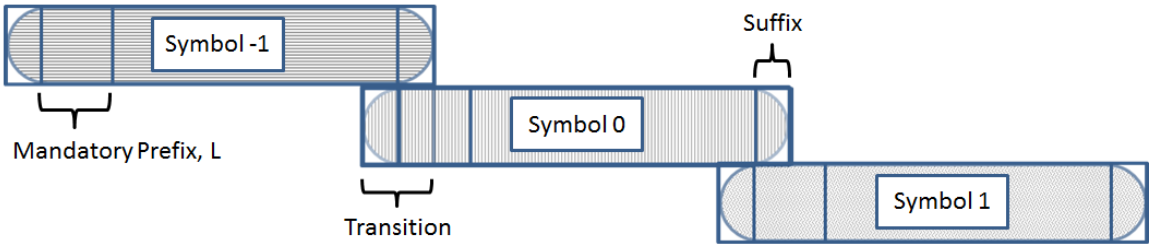

Figure 4.3: Overlaying the Extended OFDM Symbols

4.4 Estimating the OFDM Window

While using a cyclic suffix and windowing an OFDM symbol is well documented, window estimation for OFDM signals is rarely found in the literature. Estimating the window function is not necessary for demodulation. A cyclic suffix and window are not required in many OFDM standards, such as IEEE 802.11.

A preamble is included at the start of many OFDM packets in order to facilitate estimation of various parameters such as carrier frequency offset and symbol timing phase. That estimate is used to synchronize the received signal. In addition to synchronization impairments, the windowed OFDM signal defined in (4.9) passes through a wireless channel modeled as a linear convolutional impairment shown in (4.11). The term h represents the channel impulse response. An estimate of the channel impulse response can be obtained from the aforementioned preamble. The channel impulse response is estimated in the frequency domain and provides estimates for non-zero subcarriers only. In the case of IEEE 802.11g this represents 52 out of 64 possible subcarriers. The channel estimate is linearly interpolated over the null-subcarriers in phase and magnitude as described in [12]. The channel estimate is then converted to the time domain. The sample index $k - \ell$ in (4.11) can take on a value less than $-L$, as allowed by (4.9) for any combined symbol v_p.

$$y_p[k] = \sum_{\ell=0}^{L-1} h[\ell] v_p[k-\ell] \tag{4.11}$$

The receiver generates an estimate of the received signal $\hat{y}_p[k]$ which includes the multipath channel distortion. This estimate is constructed from estimates of the individual OFDM symbols $\hat{o}_p$ and an estimate of the multipath channel $\hat{h}$. A squared-error term is defined in (4.12) using the estimate $\hat{y}_p[k]$. Minimizing the error defined in (4.12) with respect to the window coefficients provides an estimate of the window $\hat{w}_i[Kp+k]$ where K is the length of an OFDM symbol without windowing. Note that the sample index k is relative to the current symbol. For each OFDM symbol p, the sample index k ranges over $-L \leq k \leq M-1$ thus creating $M+L$ error samples for each OFDM symbol used. As an example, for an 802.11 OFDM sample, there would be 80 (64+16) error samples produced for each OFDM symbol. The estimation of the window is performed over N OFDM symbols each of length K samples.

$$|\epsilon[Kp+k]|^2 = \left(y_p[k] - \hat{y}_p[k]\right)\left(y_p^*[k] - \hat{y}_p^*[k]\right) \tag{4.12}$$

Taking the derivative of (4.12) with respect to the estimate of the window $\hat{w}_i[Kp+k]$ where $i = k - q\,(M+L) - \ell$, and substituting $v_p[k-\ell]$ with (4.10), the derivative in (4.13) is found. The derivative in (4.13) reduces to (4.14).

$$\begin{aligned} &\frac{\partial |\epsilon[Kp+k]|^2}{\partial \hat{w}_i[Kp+k]} = \\ &-\epsilon[Kp+k]\hat{h}^*[\ell]\hat{o}^*_{(p+q) \bmod N}[i] \\ &-\epsilon^*[Kp+k]\hat{h}[\ell]\hat{o}_{(p+q) \bmod N}[i] \end{aligned} \tag{4.13}$$

$$\frac{\partial|\epsilon[Kp+k]|^2}{\partial\hat{w}_i[Kp+k]} = \\ -2\,\mathrm{Re}\{\epsilon^*[Kp+k]\hat{h}[\ell]\hat{o}_{(p+q) \bmod N}[i]\} \tag{4.14}$$

The update step is defined in (4.15). For every OFDM symbol there are L updates to each window coefficient estimate as the channel h is L coefficients long. Each OFDM symbol is indexed by p indicating the sequential placement of that OFDM symbol in the OFDM packet. The update function in (4.14) is scaled by a step size μ and accumulated into estimates of all $\hat{w}_i[Kp+k+1]$ for each valid index i. The values for α are the window coefficient estimates $\hat{w}_i[Kp+k+1]$ where $M-L \leq i \leq M+L-1$. The values for β are the window coefficient estimates $\hat{w}_i[Kp+k+1]$ where $-2L \leq i \leq -1$.

$$\hat{w}_i[Kp+k+1] = \hat{w}_i[Kp+k] \\ +2\mu\,\mathrm{Re}\{\epsilon^*[Kp+k]\hat{h}[\ell]\hat{o}_{(p+q) \bmod N}[i]\} \tag{4.15}$$

The process of estimating an OFDM window is as follows: when receiving an OFDM packet, demodulate all symbols. Use the demodulated bits to create a new OFDM packet. The received OFDM packet is $y_p[k]$ and the locally generated one is $\hat{y}_p[k]$. Use no windowing on $\hat{y}_p[k]$. The estimate $\hat{y}_p[k]$ is used as the reference in (4.12) to which the received data measurement $y_p[k]$ is compared. The received OFDM signal $y_p[k]$ and the estimate $\hat{y}_p[k]$ are then used as in (4.14) to create an estimate of the window $\hat{w}$.

Adding noise into the window estimation creates uncertainty not only in the measurement $y_p[k]$ but also in the reference $\hat{y}_p[k]$ to which that measurement is compared because the added noise creates bit errors. Those bit errors cause errors in the estimate $\hat{y}_p[k]$ and that reduces the total accuracy of the window estimation. Fig. 4.4 shows the average RMS error resulting from estimating the window of a

received OFDM packet as a function of SNR. The estimate is performed on an IEEE 802.11 packet using 64 QAM. The channel estimate is perfect for this measurement. Each data point in Fig. 4.4 represents an estimate performed over 148 OFDM symbols using a step size of 0.01 and 20 epochs. The window defined in [6] and the parameter *transition time* serves as a roll-off. Here the transition time is set to 500 ns for ease of visualization. The window error is not linear with SNR because the frequency of bit errors increases as SNR decreases, adding additional uncertainty into the estimate.

Fig. 4.5 shows the actual and estimated window overlaid. The window was estimated with a data symbol E_s/N_0 of 27 dB using 20 epochs. The index i ranges from $-32 \leq i \leq 79$, that being $-2L \leq i \leq M + L - 1$. The regions representing α

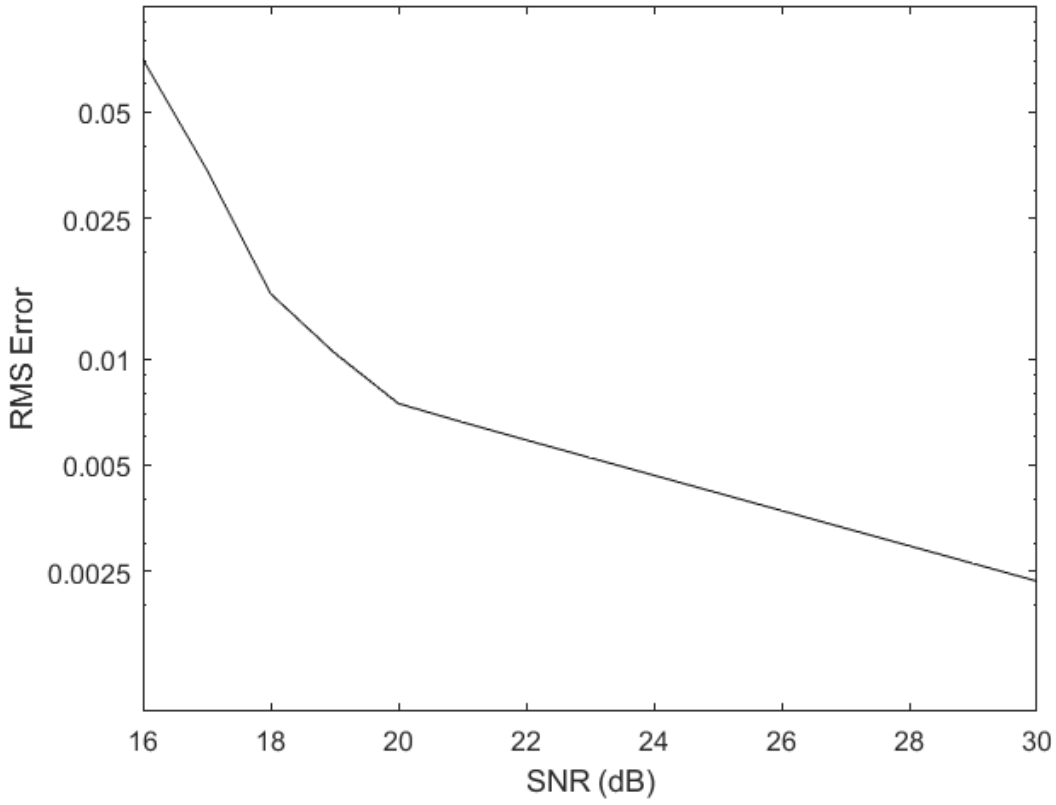

Figure 4.4: Window Error as a function of SNR

and β are shown.

4.5 Cancelling OFDM Signals with Imperfect Window and Channel Estimates

Combining the definition of windowing in (4.10) with that of cancellation in (4.6) results in (4.16). For the purposes of cancellation, windowing is represented as a diagonal matrix Λ_{w_q} estimated as $\widehat{\Lambda_{w_q}}$. Each element is Λ_{w_q} is the window value applied to the signal s_q offset by q as in (4.10). Three copies of the signal $\vec{s}$ without windowing are created, offset in samples, and then windowed by the respective Λ_{w_q}. The channel and CFO estimates will be imperfect as in (4.6).

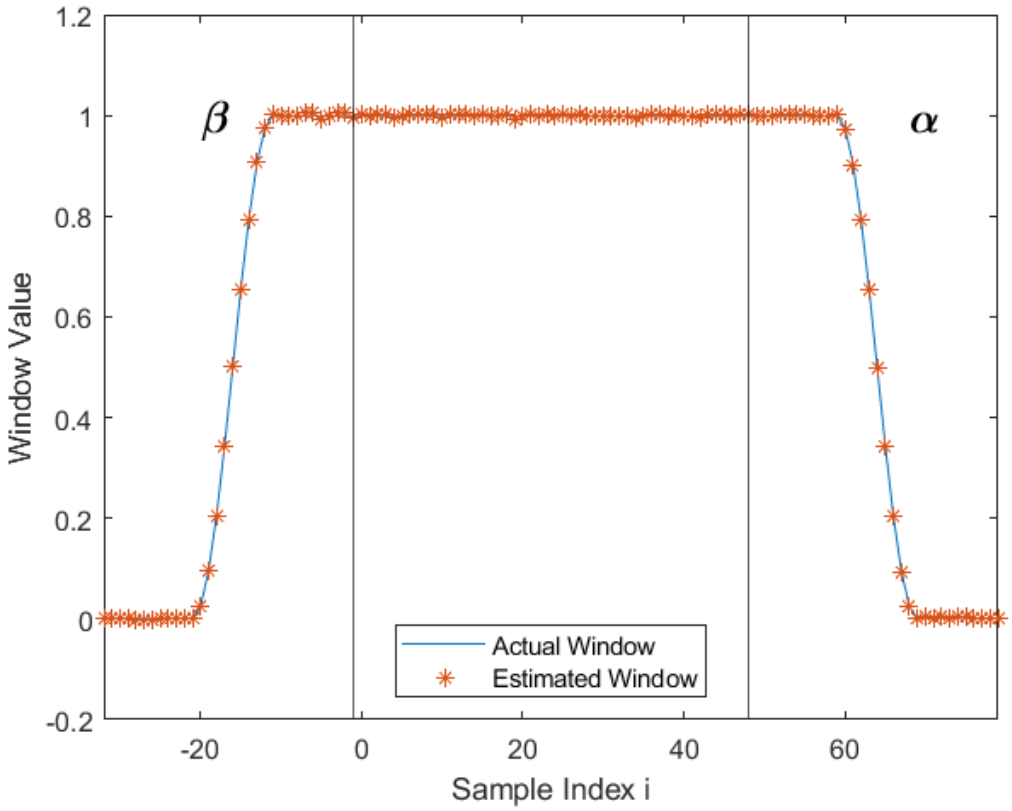

Figure 4.5: Actual and Estimated Window Overlaid

$$\vec{u} = \Lambda_{\Theta} H \sum_{q=-1}^{1} \Lambda_{w_q} s_q + \vec{n} - \widehat{\Lambda_{\Theta}} \widehat{H} \sum_{q=-1}^{1} \widehat{\Lambda_{w_q}} \widehat{s_q} \tag{4.16}$$

In this experiment, 500 synthetic OFDM packets are generated each with a payload of 4000 bytes. The raised cosine window as defined in [6] with a range of transition times to the packets from 100 ns to 1 μs. A window transition time of 100 ns represents the smallest window that can be applied to the 802.11 packet. We execute the window estimation algorithm across a range of SNR from 20 to 30 dB. We then perform cancellation with and without the window estimate following (4.6) and (4.16). The cancellation c is calculated as $c = var(\vec{r})/var(\vec{u})$ that being the ratio of the variance of the received samples $\vec{r}$ to that of the residue $\vec{u}$. The ratio of the cancellation c with windowing to without windowing for six window transition values are plotted in Fig. 4.6. The two methods are equal (0 dB) when no window is present. As soon as even a small window is applied, the cancellation method with windowing falls well short of that without. In Fig. 4.7 the power of the residue resulting from each test case is subtracted from the SNR of that test case. The SNR represents the absolute maximum cancellation, as then the residue would be noise. The results are all negative values demonstrating all test cases fall short of perfect estimates. The test cases employing windowing estimation keep much closer to a ratio 0 dB ratio.

4.6 Over The Air Experiment

4.6.1 OTA Cancellation Residue Reduction

The OTA data from chapter 2 and [1] was used in an experiment to determine the how well the technique described in this chapter would improve cancelling OTA OFDM signals. In this experiment, the two cancellation algorithms were applied to 961 OFDM packets captured OTA. The SNR estimates of the OTA packets varied

from 29 to 31 dB. The modulation scheme used by all captured OFDM packets for the data was QAM64. The average cancellation without estimating the window as in (4.6) was 19.5 dB. The average cancellation with estimating windowing as in (4.16) was 5.3 dB higher. Both the base cancellation without windowing and the windowing improvement fell short of expectations from the experiment with synthetic data, however, this experiments demonstrates unequivocally that the window estimation provided a substantial boost in cancellation. Fig. 4.8 is a spectrum plot showing a single recorded OTA packet and the residue from the cancellation algorithms with and without windowing. This figure illustrates the impact of including windowing in the cancellation algorithm.

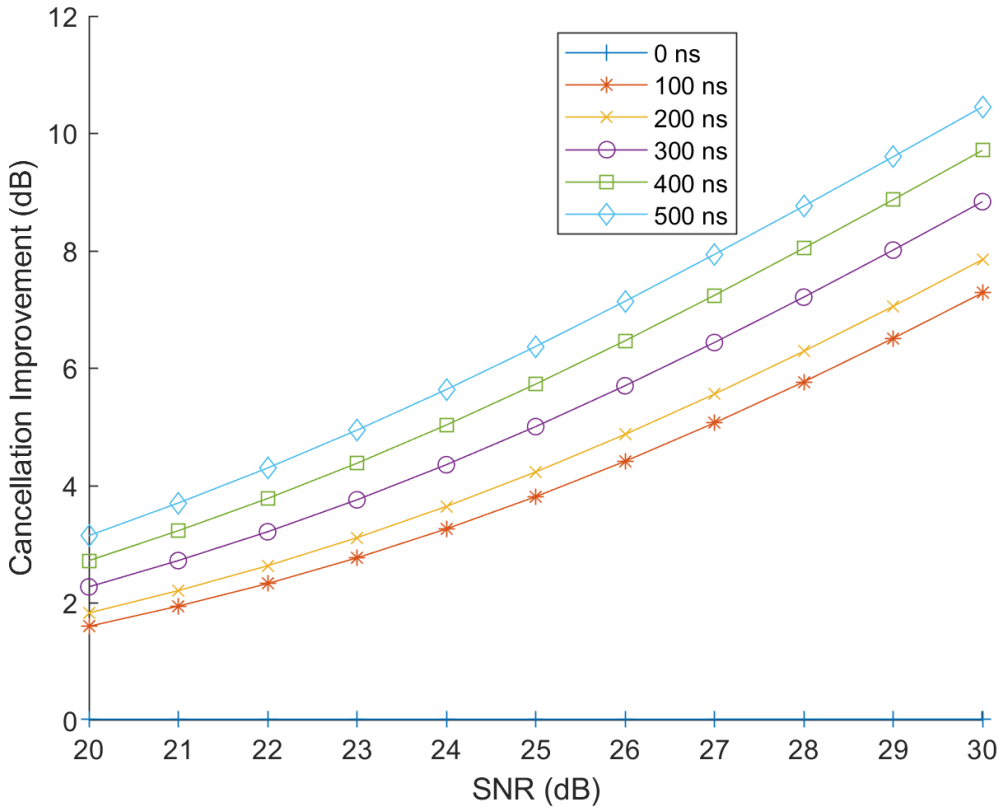

Figure 4.6: Cancellation Improvement as a function of SNR

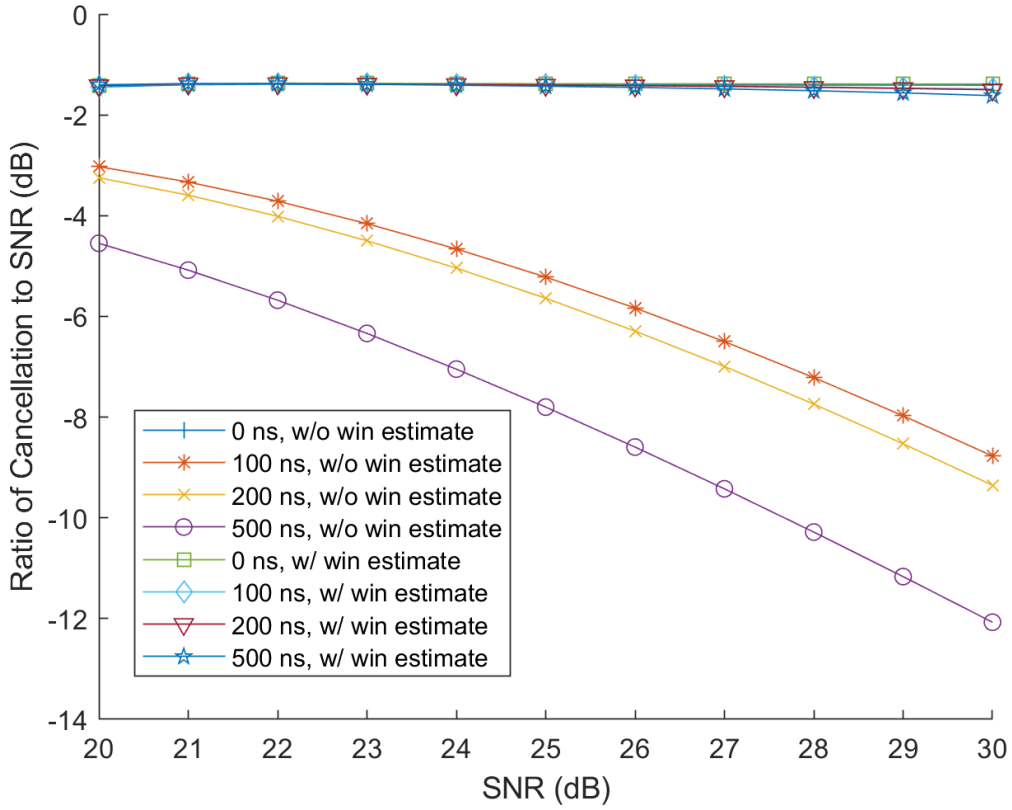

Figure 4.7: Ratio of Cancellation to SNR as a function of SNR

Fig. 4.9 shows the residue of the cancellation in time with and without windowing. The residue without windowing is as it was in chapter 2. The cancellation improves by 5.3 dB. There is still a periodic jump in the residue magnitude that occurs in between OFDM symbols, but it is now diminished.

4.6.2 OTA PER Improvement

The OTA data from chapter 2 and [1] was used in an experiment to determine the how well the technique described in this chapter would improve Packet Error Rates. The improvement was measured over a range of SNR. In order to generate that range, synthetic complex-valued noise was added to the recorded packets. The synthetic noise was scaled relative to the variance of a given OTA recording. If

a given OTA recording had a variance of σ_r^2, the noise would be scaled to have a variance of $\rho\sigma_s^2$ where ρ is the linear value of the desired SNR. The variance σ_r^2 is for a given recording, including the noise in that recording as $r = s + n$. This the actual SNR for any instance in this experiment is actually lower (worse) than the target SNR. That fact is not important to this experiment as the PER resulting from the window cancellation improvement will be compared to the PER at that same simulated SNR without the window cancellation improvement.

Fig. 4.10 shows the PER curves as functions of SNR resulting of this experiment. The improvement enabled by the window cancellation is more clear in Fig. 4.11 where the difference between the two PER curves, $PER_{normal} - PER_{cancel}$ is plotted. Note that the difference is always positive meaning the PER of the normal case

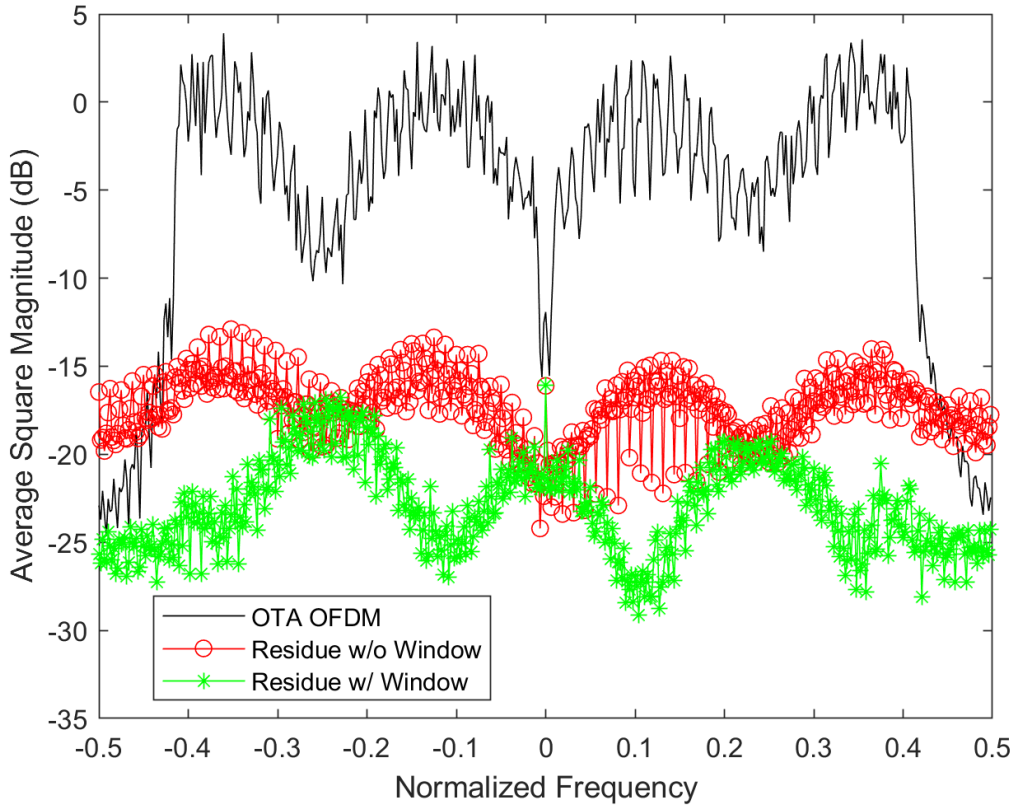

Figure 4.8: Spectrum of OTA Packet and Two Residues

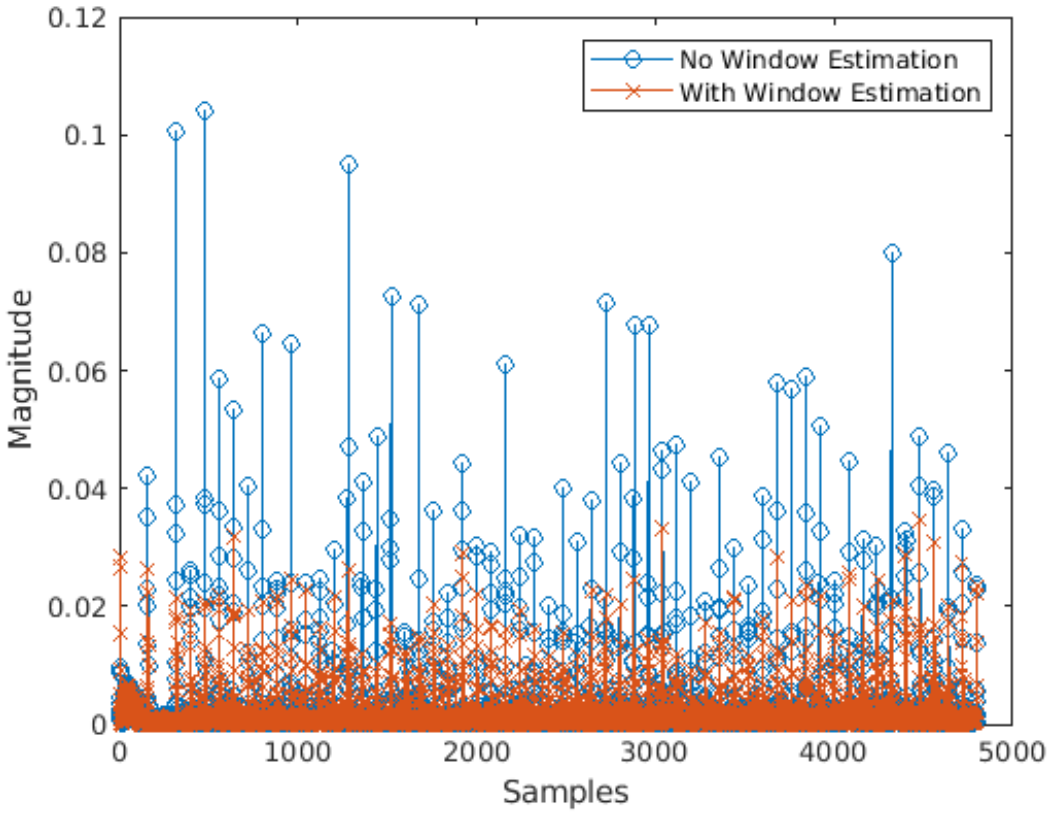

Figure 4.9: OTA Cancellation Residue with and without Windowing

is always worse than the case with window cancellation. The difference in Fig. 4.11 is plotted as an absolute value of increased throughput. The peak of this absolute value occurs in the low 20s of dB SNR where the PER curve for the 64QAM modulation scheme began to accelerate upward in chapter 2. At the peak of the PER improvement, the maximum throughput of the 54 Mbps data rate mode (64QAM) of 802.11a/g increases by 1.428 Mbps representing a 2.65% improvement.

4.6.3 OTA Covert Signal BER Improvement

The OTA experiment from chapter 2 and [1] was repeated to determine if Window estimation improved the covert signal BER. The covert signal is the same as in chapter 2 and [1]. The average cancellation without estimating the window as in (4.6)

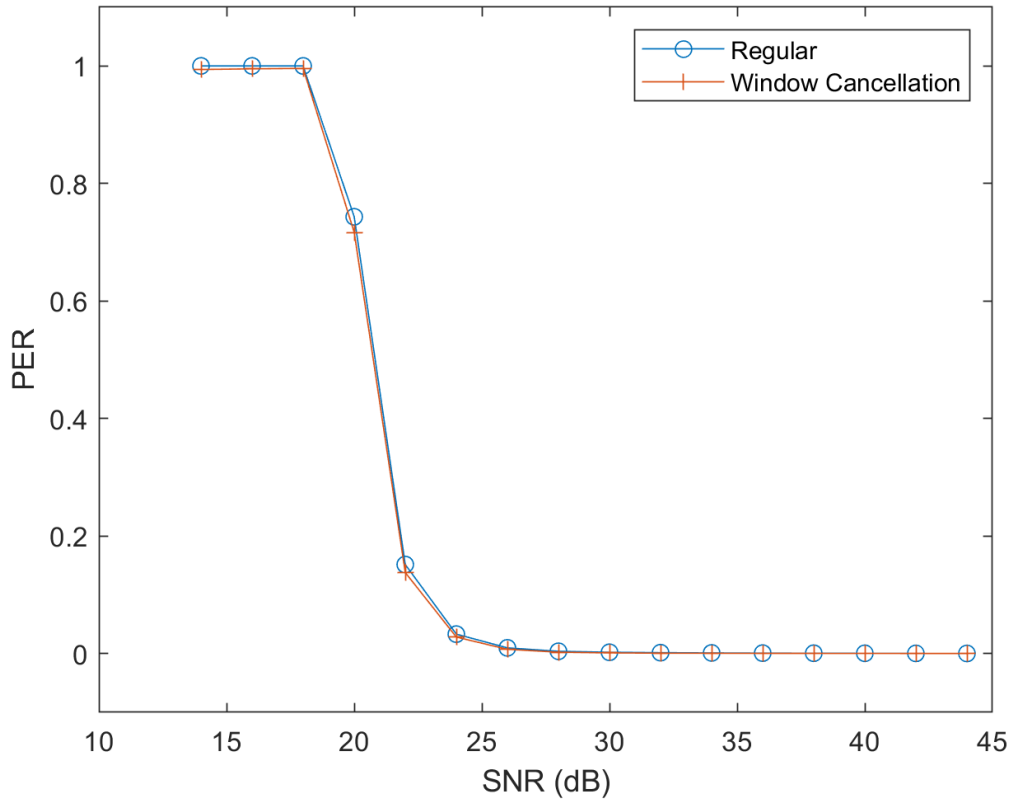

Figure 4.10: Packet Error Rate as a function of SNR

was 19.5 dB with no covert signal present. Fig. 4.12 plots the resulting BER curves. The first curve is the covert BER with no cancellation of the incumbent OFDM signal. The second curve is the covert BER with cancellation but no accounting for the windowing at the OFDM transmitter. The last covert BER employs cancellation of the incumbent with an estimate of the windowing at the OFDM transmitter. The estimated window gives better performance than with no accounting for the windowing at the OFDM transmitter.

4.7 Conclusion

We have presented and evaluated the performance of both an OFDM window estimation method and a technique to use that estimate to cancel the self-interference

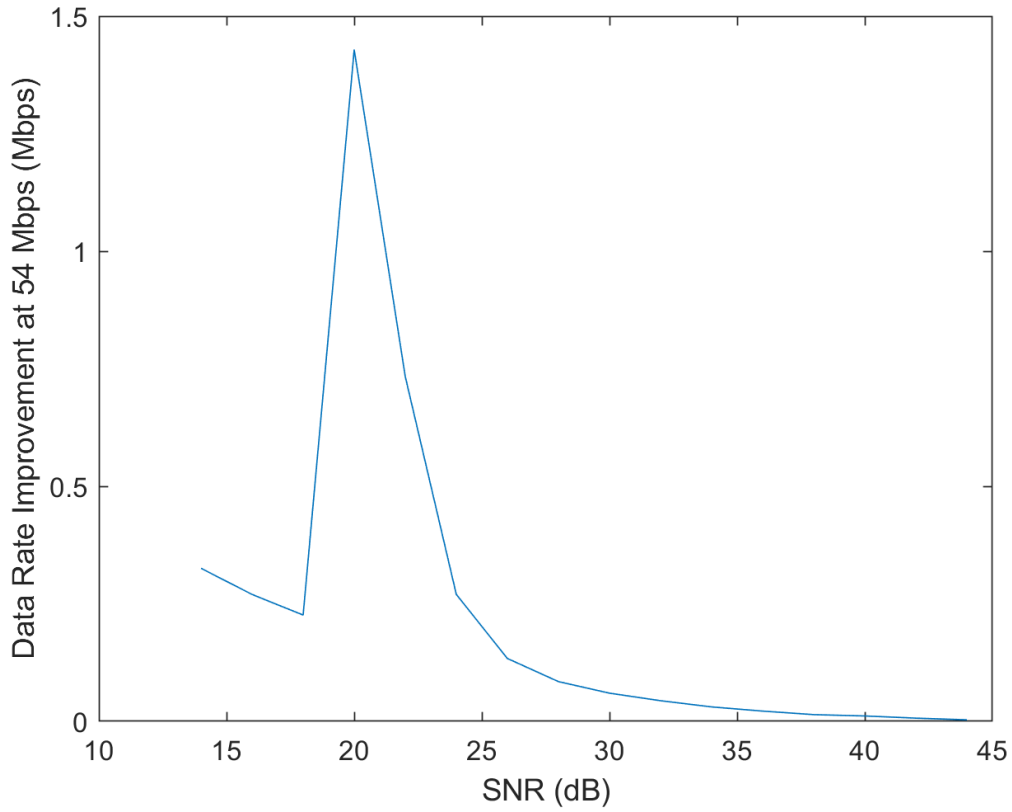

Figure 4.11: Difference in the Packet Error Rate as a function of SNR

resulting from OFDM windowing. We presented the window estimation method in a generalized form that does not require foreknowledge of the window implementation. We use 802.11 as an example application for this method. We show that despite the presence of bit errors which cause uncertainty in the window estimate, the estimation and cancellation methods offer significant performance improvement. Self-interference resulting from OFDM windowing The work in this chapter has developed an algorithm to estimate at the receiver the OFDM window applied at the transmitter. The estimation algorithm was designed to be applied to any OFDM windowing at the transmitter without foreknowledge of the window implementation. The estimation algorithm was evaluated in terms of RMS in AWGN. 802.11 was used as an example application for this method. The estimation algorithm

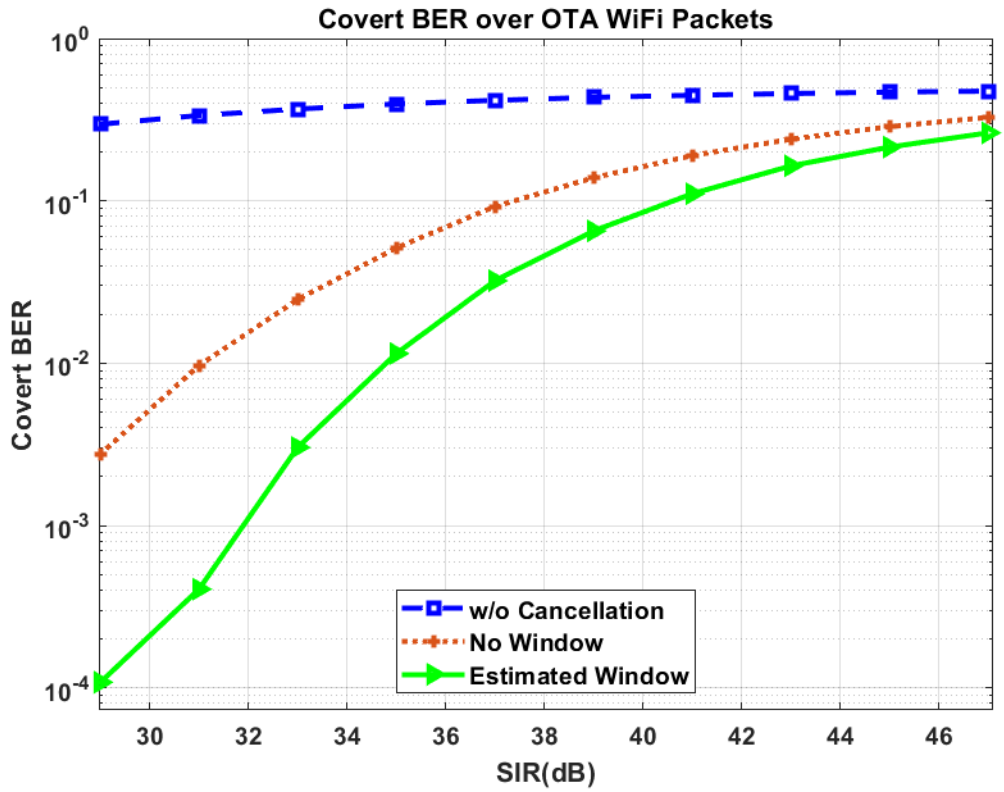

Figure 4.12: Difference in the Packet Error Rate as a function of SNR

was used to create window estimates that were then applied to the cancellation synthetic and OTA OFDM signals. It was shown that including windowing greatly improved the total cancellation of OFDM signals. The targeted cancellation of the extended cyclic prefix and entire cyclic suffix resulted in an improvement in PER.

4.8 References

[1] Daniel Chew et al. "Covert Communications through Imperfect Cancellation". In: *Accepted to Proceedings of the 2022 ACM Workshop on Information Hiding and Multimedia Security*. IH&MMSec '22. New York, NY, USA: Association for Computing Machinery, 2022.

[2] Bruno Clerckx et al. "Rate-Splitting Unifying SDMA, OMA, NOMA, and Multicasting in MISO Broadcast Channel: A Simple Two-User Rate Analysis". In: *IEEE Wireless Communications Letters* 9.3 (2020), pp. 349–353. DOI: 10.1109/LWC.2019.2954518.

[3] Behrouz Farhang-Boroujeny. "OFDM Versus Filter Bank Multicarrier". In: *IEEE Signal Processing Magazine* 28.3 (2011), pp. 92–112. DOI: 10.1109/MSP.2011.940267.

[4] Behrouz Farhang-Boroujeny and Roland Kempter. "Multicarrier communication techniques for spectrum sensing and communication in cognitive radios". In: *IEEE Communications Magazine* 46.4 (2008), pp. 80–85. DOI: 10.1109/MCOM.2008.4481344.

[5] P. Huang and Y. Lee. "Adaptive decision feedback orthogonality restoration filter for windowed OFDM". In: *IEEE 54th Vehicular Technology Conference. VTC Fall 2001. Proceedings (Cat. No.01CH37211)*. Vol. 2. 2001, 1106–1110 vol.2. DOI: 10.1109/VTC.2001.956946.

[6] "IEEE Standard for Information Technology–Telecommunications and Information Exchange between Systems - Local and Metropolitan Area Networks–Specific Requirements - Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications". In: *IEEE Std 802.11-2020 (Revision of IEEE Std 802.11-2016)* (2021), pp. 1–4379. DOI: 10.1109/IEEESTD.2021.9363693.

[7] Y. Lee and P. Huang. "Performance analysis of a decision feedback orthogonality restoration filter for IEEE 802.11a". In: *2002 IEEE Wireless Communications and Networking Conference Record. WCNC 2002 (Cat. No.02TH8609)*. Vol. 1. 2002, 449–453 vol.1. DOI: 10.1109/WCNC.2002.993537.

[8] Talgat Manglayev, Refik Caglar Kizilirmak, and Yau Hee Kho. "Comparison Of Parallel And Successive Interference Cancellation For Non-Orthogonal Multiple Access". In: *2018 International Conference on Computing and Network Communications (CoCoNet)*. 2018, pp. 74–77. DOI: 10.1109/CoCoNet.2018.8476815.

[9] Nikolaos I. Miridakis and Dimitrios D. Vergados. "A Survey on the Successive Interference Cancellation Performance for Single-Antenna and Multiple-Antenna OFDM Systems". In: *IEEE Communications Surveys Tutorials* 15.1 (2013), pp. 312–335. DOI: 10.1109/SURV.2012.030512.00103.

[10] Erdal Panayirci, Habib Senol, and H. Vincent Poor. "Joint Channel Estimation, Equalization, and Data Detection for OFDM Systems in the Presence of Very High Mobility". In: *IEEE Transactions on Signal Processing* 58.8 (2010), pp. 4225–4238. DOI: 10.1109/TSP.2010.2048317.

[11] Alphan Sahin and Huseyin Arslan. "Edge Windowing for OFDM Based Systems". In: *IEEE Communications Letters* 15.11 (2011), pp. 1208–1211. DOI: 10.1109/LCOMM.2011.090611.111530.

[12] C. Shahriar et al. "PHY-Layer Resiliency in OFDM Communications: A Tutorial". In: *IEEE Communications Surveys Tutorials* 17.1 (2015), pp. 292–314. DOI: 10.1109/COMST.2014.2349883.

[13] Tayebeh Taheri, Rickard Nilsson, and Jaap van de Beek. "Asymmetric Transmit-Windowing for Low-Latency and Robust OFDM". In: *2016 IEEE Globecom Workshops (GC Wkshps)*. 2016, pp. 1–6. DOI: 10.1109/GLOCOMW.2016.7848842.

[14] S. Vanka et al. "Superposition Coding Strategies: Design and Experimental Evaluation". In: *IEEE Transactions on Wireless Communications* 11.7 (2012), pp. 2628–2639. DOI: 10.1109/TWC.2012.051512.111622.

[15] Bob Ward. "Non Rectangular Time Windowing Analysis for IEEE 802.11 OFDM System". In: *IEEE P802. 11 Working Group Contribution, IEEE 802.11-99/021* (1999).

Chapter 5

Exploiting Vulnerabilities in Deep-Learning RF Classification using an Interference Signal

5.1 Introduction

Spectrum monitoring is a function performed for a number of commercial and military applications. Spectrum monitors will sense the spectrum for wireless activity and will often attempt some form of classification if activity is found. The scenario is illustrated in Fig. 5.1. Two nodes are communicating, and the spectrum monitor intercepts their communication. The spectrum monitor can then fill the role of “Eve the Eavesdropper" discussed in chapter 1. It is common for spectrum monitors to have a modulation classification capability. CNNs have been used in prior work as a means of modulation classification such as in [11]. One of the things that can be done with the modulation classifier is reverse engineering the intercepted link. For the sake of privacy, and to prevent spoofing, it is advantageous to make such reverse engineering expensive. To that end this chapter will focus on breaking the modulation classification capability in the spectrum monitor.

The covert signal developed in chapter 2 and [3] is a BPSK waveform hidden

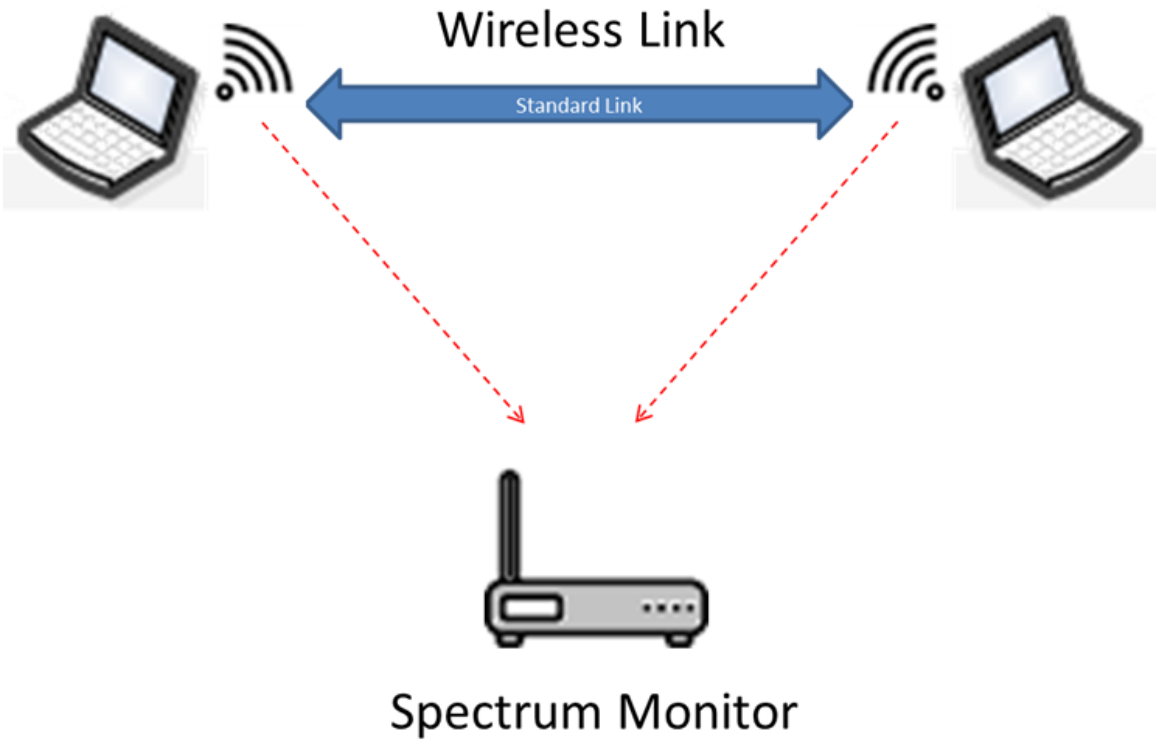

Figure 5.1: Spectrum Monitoring Scenario

inside an OFDM incumbent. In the case of detection, it is beneficial to frustrate classification efforts to potentially convince the monitor the detection was an anomalous false alarm and to frustrate any reverse engineering. Therefore, it is advantageous to arm the covert signal with a means to defeat a modulation classifier.

It is well documented that CNNs are susceptible to adversarial attacks, the concept being first introduced in [13]. Prior work has shown the potential to transfer such attacks from one CNN to another [5]. It was suggested in [5] that the cause of this transferability phenomenon was that these different CNNs are learning similar functions when they are trained to perform the same task. Therefore, it may be possible to develop an adversarial attack against one CNN modulation classifier and deploy it against a different and unknown CNN also used for modulation classification. This will allow the two in-network nodes to employ a means to

frustrate a monitor's modulation classification without knowledge of the internals of the CNN used.

Prior work has shown that modulation classifiers based on CNNs are susceptible to adversarial attacks as shown in [12] [1] [4] [8] [6] and [7]. The work in [12] uses the Fast Gradient Method to develop the adversarial waveform for untargeted attacks and measured the frequency of misclassification. The work in [1] uses the Fast Gradient Sign Method to generate a targeted attack. The work in [8] considers adversarial waveforms for both targeted and untargeted attacks with consideration of the multipath channel. The work in [4] examines the effect of several RF impairments on the application of an adversarial waveform created using the Fast Gradient Sign Method, those impairments being noise, center frequency offset (CFO), sample time offset, and a dynamic channel. In [4] it was found that the adversarial example was less effective as CFO increased. The analysis in [4] has similarities to this work; however, in this work we allow the spectrum monitor to determine the center of the energy detected in the spectrum and allocate a bandwidth for classification. The results in this work show the attack is more effective as the spectrum monitor overestimates or deliberately allows a larger bandwidth. The training method in [6] and [7] is similar to the one employed in this work. In [7] an adversarial waveform is created for the purpose of masking a communication link from a monitor using Projected Gradient Descent. This constrains the L2 norm of the input vectors, thus constraining the power of the adversarial waveform, for the purpose of mitigating adverse effects on the BER of the communication link. The difference in the training method in this work and [7] is that the constraint employed here is imposed on the ratio between the training waveform and the adversarial waveform, thus the power of the adversarial waveform can be scaled freely with respect to the transmitted waveform. All of the cited prior work allows the monitor to intercept the waveform with no uncertainty

in the symbol rate (in terms of samples per symbol). In a deployed spectrum monitoring system, the monitor would not have this information. A common technique would be for the monitor to employ spectrum sensing to detect the waveform, perform bandwidth estimation on the waveform, and then estimate the necessary sample rate based on that information. In this work, we test the transferability of an adversarial waveform with this uncertainty in mind.

In this work, a system consisting of a two communicating wireless network nodes and a spectrum monitor was considered. The spectrum monitor intercepts the wireless link between the two nodes and attempts to determine the modulation being used. The spectrum monitor was modeled as sensing the wireless activity, estimating the bandwidth, and allocating a sample rate to be used for classification based on that bandwidth estimation. The monitor employs a CNN to classify the observed modulation scheme.

In order to frustrate classification, the wireless nodes employed an interference signal to be transmitted concurrently with the communication signal. This specialized interference is referred to as the *adversarial waveform*. This adversarial waveform provided a *targeted attack*, meaning that the misclassification was directed to a chosen false modulation scheme. The adversarial waveform was trained to be a targeted attack and to constrain the Signal-to-Interference Ratio (SIR) to a constant value. That attack was then applied against another CNN modulation classifier with a different but overlapping set of modulation classes, different structure, and trained on different data. Not all misclassifications observed in the second modulation classifier were equal to the intended target modulation. Both targeted and untargeted misclassifications on the second classifier were measured over a range of error of the bandwidth estimate.

The contributions of this work are:

- Create an adversarial waveform that is easily deployable. Because the adversarial waveform and the communication signal are transmitted from the same node concurrently and additively, the monitor will receive them both across the same multipath channel. Therefore, there is no separate multipath channel to consider for the adversarial waveform. To further simplify deployment, the power of the adversarial waveform will be scaled with the power of the communication signal to meet a target SIR. The power of the adversarial waveform need only be set relative to the locally generated transmitted signal. To meet this goal, SIR was chosen as the sole constraint to be applied to the training of the adversarial waveform.

- Create an adversarial waveform that mitigates the adverse effects of the self-interference on the desired communication signal. We trained an adversarial waveform on one CNN, called CNN-A, by constraining SIR. The adversarial waveform in this work is not explicitly constrained to minimize its impact on the BER rate of the communication signal. Our contention is that we can mitigate the adverse effects on the communication signal by constraining SIR.

- Test the transferability of the adversarial waveform to an unknown CNN, representing the monitor and called CNN-B, with symbol rate uncertainty. The two wireless nodes cannot know the details of the modulation classifier used by the monitor. The development of the adversarial waveform is deprived of any knowledge of the training set or modulation classes employed by the monitor. Therefore the adversarial waveform must be developed on a separate modulation classification CNN. Additionally, the monitor does not have exact knowledge of the symbol rate. The modulation classification will be performed over a range of symbol rate estimation error.

Two different CNNs were used for modulation classification which will be

referred to as CNN-A and CNN-B. CNN-A and CNN-B were trained to classify different, but overlapping, sets of modulation classes and at a different number of samples per symbol. The two CNNs were trained on different data.

After training the two CNNs on their respective training sets, an adversarial waveform was created using CNN-A. The adversarial waveform was created by applying a waveform to the input layer of the CNN, but calculating the loss function against a deliberately incorrect label. When creating the adversarial waveform all layers of CNN-A are left constant and not updated as part of back-propagating the error all the way to the input layer. The error at the input layer is accumulated over multiple training iterations and constrained to meet a constant SIR. Instead of adjusting the network weights, the accumulated error is used to bias the input waveform so as to reduce the loss on the next training iteration. This accumulated error becomes the adversarial waveform. The effect of the adversarial waveform on the BER of the original communication system was measured over a range of E_b/N_0 and SIR. The transferability of the adversarial waveform to another CNN was tested by deploying it against CNN-B over a range of SIR values and symbol rate uncertainty. CNN-B is used to test the adversarial waveform. Therefore, the internals are unknown to the in-network nodes deploying the adversarial waveform.

This chapter is organized as follows: Section 5.2 details the CNNs under test and the methods by which those CNNs were trained. The differences between the CNNs in terms of samples per symbol, modulation classes, and input layer size are detailed. Section 5.3 details the creation of the adversarial waveform. Section 5.4 details the effects of the adversarial waveform on the communication signal intended to be masked by the adversarial waveform. Section 5.5 details how well the adversarial attack transferred between the CNNs under test. Conclusions and future work are discussed in Section 5.6.

A portion of this chapter has been accepted to be published in published in [2]

5.2 Convolutional Neural Networks Under Test

5.2.1 CNN-A

For CNN-A we used the CNN model from the MATLAB example "Modulation Classification with Deep Learning" [10]. The layers of CNN-A are shown in Table 5.1, and are as specified in [10]. CNN-A has an input layer which takes in 1024 input time-domain complex samples processed as an image of size 2x1024 at 8 samples per symbol (sps) thus there are 128 symbols in each input vector. Each of the convolutional layers use rectified linear activation units. The first five convolutional layers in CNN-A end with max pooling. The sixth convolutional layer of CNN-A ends in average pooling.

5.2.1.1 Implementation

Table 5.1: Structure of CNN-A

Layer	**Dimensions**
Input	2×1024
Conv 1	$16 \times 8 \times 2$
Conv 2	$24 \times 8 \times 16$
Conv 3	$32 \times 8 \times 24$
Conv 4	$64 \times 8 \times 48$
Conv 5	$96 \times 8 \times 64$
Softmax	11

CNN-A supports the following 11 modulation classes:

- Binary phase shift keying (BPSK)
- Quadrature phase shift keying (QPSK)

- 8-ary phase shift keying (8-PSK)
- 16-ary quadrature amplitude modulation (16-QAM)
- 64-ary quadrature amplitude modulation (64-QAM)
- 4-ary pulse amplitude modulation (PAM4)
- Gaussian frequency shift keying (GFSK)
- Continuous phase frequency shift keying (CPFSK)
- Broadcast FM (B-FM)
- Double sideband amplitude modulation (DSB-AM)
- Single sideband amplitude modulation (SSB-AM)

5.2.1.2 Modulation Classification Training for CNN-A

The training information is provided in [10]. Here an overview is provided: CNN-A was trained using synthetic data representing the 11 different modulation classes. A square root raised cosine (RRC) filter is applied to the symbols of the linear digital modulation schemes (BPSK, QPSK, PSK, QAM, PAM). The excess bandwidth of the RRC filter is 0.35, the filter spans 4 symbols at 8 sps. The GFSK modulation uses a bandwidth-time product of 0.35. The CPFSK modulation uses a modulation index of 0.5.

5.2.2 CNN-B

CNN-B is based on "CNN2" from [11]. The layers of CNN-B are shown in Table 5.2. CNN-B has an input layer which takes in 128 input time-domain complex samples processed as an image of size 2x128 at 4 samples per symbol.

Table 5.2: Structure of CNN-B

Layer	**Dimensions**
Input	2×128
Conv 1	$256 \times 1 \times 3$
Conv 2	$80 \times 1 \times 3$
Dense ReLU	256
Softmax	7

CNN-B supports the following 7 modulation classes:

- Binary Continuous Phase Frequency Shift Keying (labeled “2CPM”)
- 4-ary Continuous Phase Frequency Shift Keying (labeled "4CPM")
- Binary Phase Shift Keying (BPSK)
- Quadrature Phase Shift Keying (QPSK)
- 16-ary quadrature amplitude modulation (16-QAM)
- 4-ary pulse amplitude modulation (PAM4)
- Gaussian minimum shift keying (GMSK)

5.2.2.1 Modulation Classification Training for CNN-B

Time error was introduced as uniformly random sample shifts in the range of [0,OSR], where OSR represents the oversampling rate and was set to 4 sps. That there are 4 sps means that each length 128 input vector represents 32 symbols. Phase error is introduced as a static rotation that is randomly selected for each set of 128 samples from the range of $[0,2\pi]$. Frequency offsets were induced to emulate

tuning error in a standard software-defined radio (SDR). Therefore, we defined the upper and lower bounds of our frequency error to be 10% of our symbol rate; the frequency error applied to each set of 128 samples is sampled from the uniform distribution with the bounds specified by ±10% f_{SYMB}. The range of signal-to-noise (SNR) ratio used for training and validation was [6 dB, 30 dB]. Finally, each input of 128 samples was normalized to 0 dB average power. Training was conducted with a batch size of 1024, using negative log likelihood loss (NLLL) and Adam optimizer [9] with a learning rate of 10^{-3}.

5.3 Creating the Adversarial Waveform

The adversarial waveform was created by training on CNN-A. For each training iteration, a batch of 128 random BPSK waveforms was generated. Each waveform was complex valued and 1024 samples long. Signal power was kept constant for each input waveform during adversarial waveform training. The label of the training data is set to CPFSK instead of BPSK. Because the training data is deliberately mislabeled to a specific modulation scheme, the resulting adversarial waveform will be a targeted attack. The loss function for the training was cross entropy. The layers of CNN-A were prevented from updating any weights during the training. The error gradient was back propagated to the input layer where it was accumulated as a bias that becomes the adversarial waveform. The adversarial waveform was initially set to random values following a Gaussian distribution. After each training iteration, the power of the adversarial waveform was measured. The power of the adversarial waveform was compared to the constant waveform power to measure SIR. The accumulated adversarial waveform is then scaled such that the SIR was kept constant across all training iterations. This scaling imposes a constraint on the power of the adversarial waveform relative to the power of the training samples. The update equation is shown in (5.1) where k is an iteration

index, $\vec{x}$ is the current vector of BPSK samples to be input, $\vec{a}[k]$ is the current vector representing the adversarial waveform, y_{target} is the target class, μ is the step size, $P_{\vec{u}[k]}$ is the average power of the intermediate vector $\vec{u}$, and P_{SIR} is the desired SIR.

$$\vec{a}[k] = \frac{\vec{u}[k]}{P_{\vec{u}[k]}} P_{SIR} \tag{5.1}$$

$$\vec{u}[k] = \vec{a}[k-1] - \mu \nabla_{\vec{x}} (L(\theta, \vec{x}[k-1] + \vec{a}[k-1], y_{target})) \tag{5.2}$$

A target SIR of 7 dB was found to provide consistent misclassification after 100 training iterations. The spectrum of the pulse shaped BPSK signal overlaid with the spectrum of the adversarial waveform is shown in Fig. 5.2.

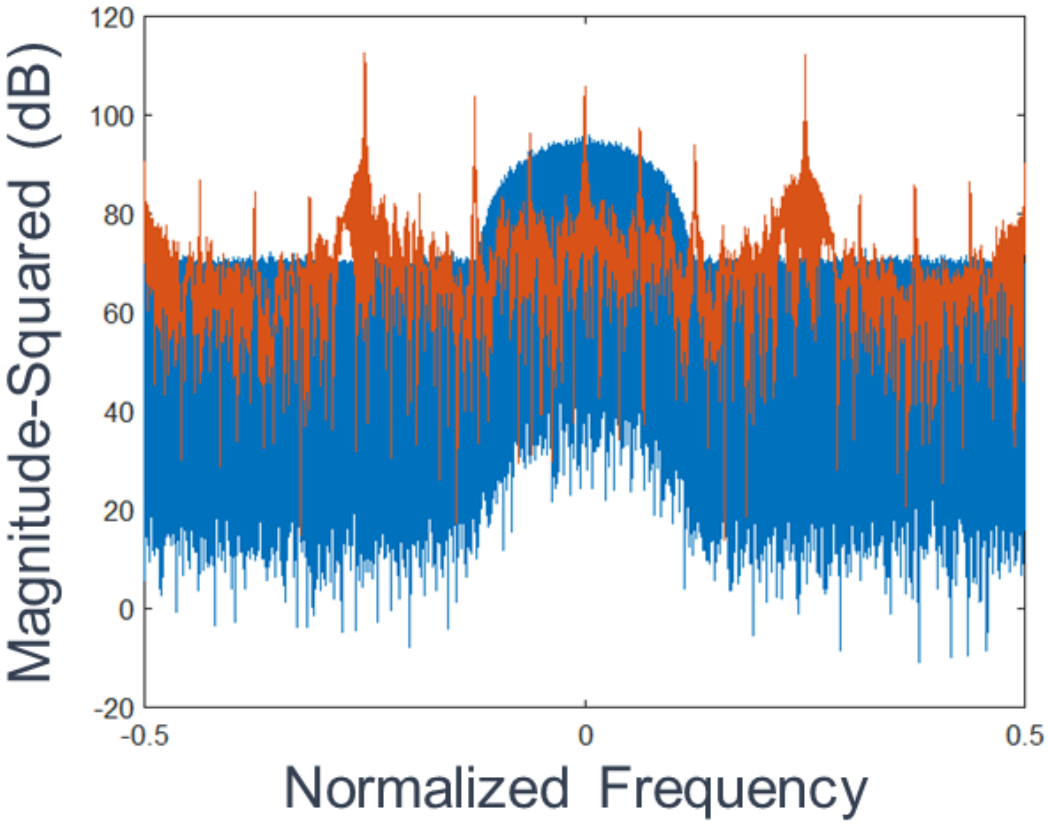

Figure 5.2: Spectrum of RRC BPSK and Adversarial Waveform

After the adversarial waveform was trained, CNN-A was tested against all modulation types with the adversarial waveform added. The confusion matrix is shown in Fig. 5.3. The adversarial waveform pushed the classification of linear modulation schemes toward CPFSK. CPFSK was unaffected. Angular modulation schemes such as B-FM and GFSK were affected much less than the linear modulation schemes.

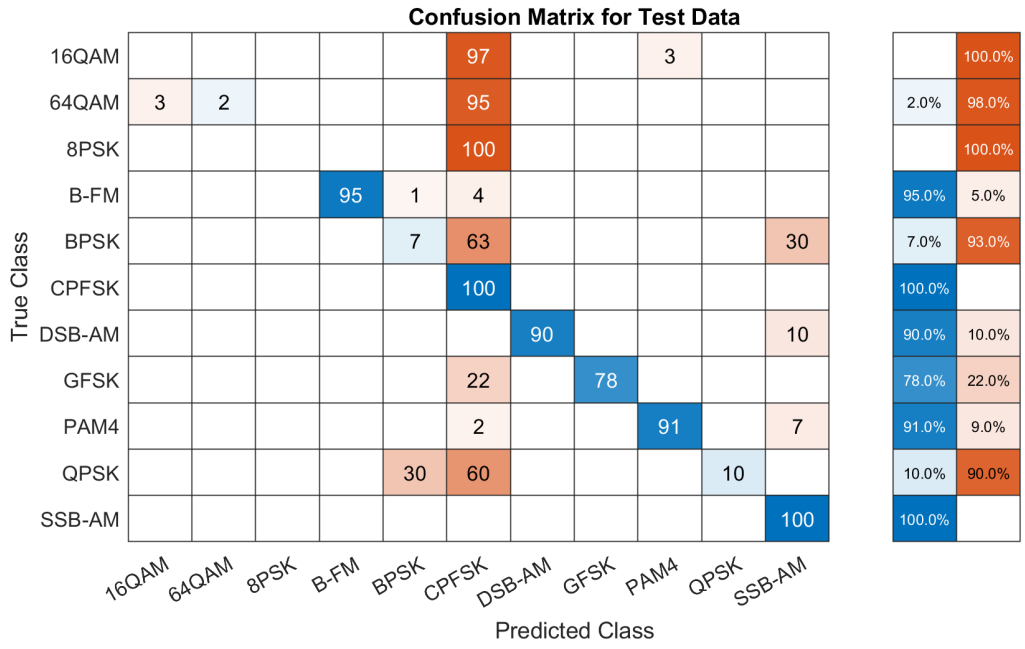

Figure 5.3: Confusion Matrix of CNN-A after Adversarial Waveform Applied

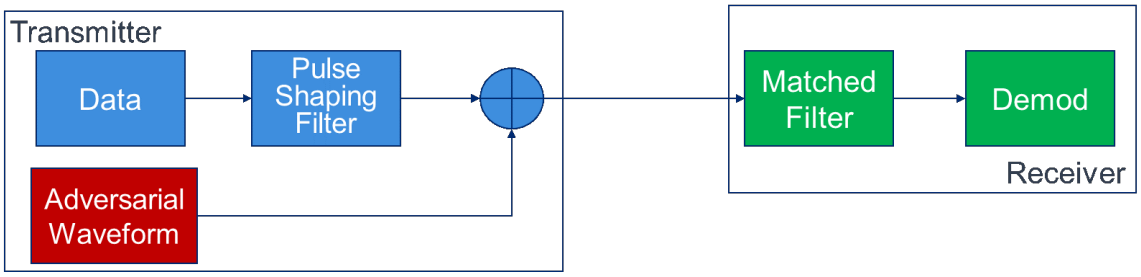

Figure 5.4: Communication System Model

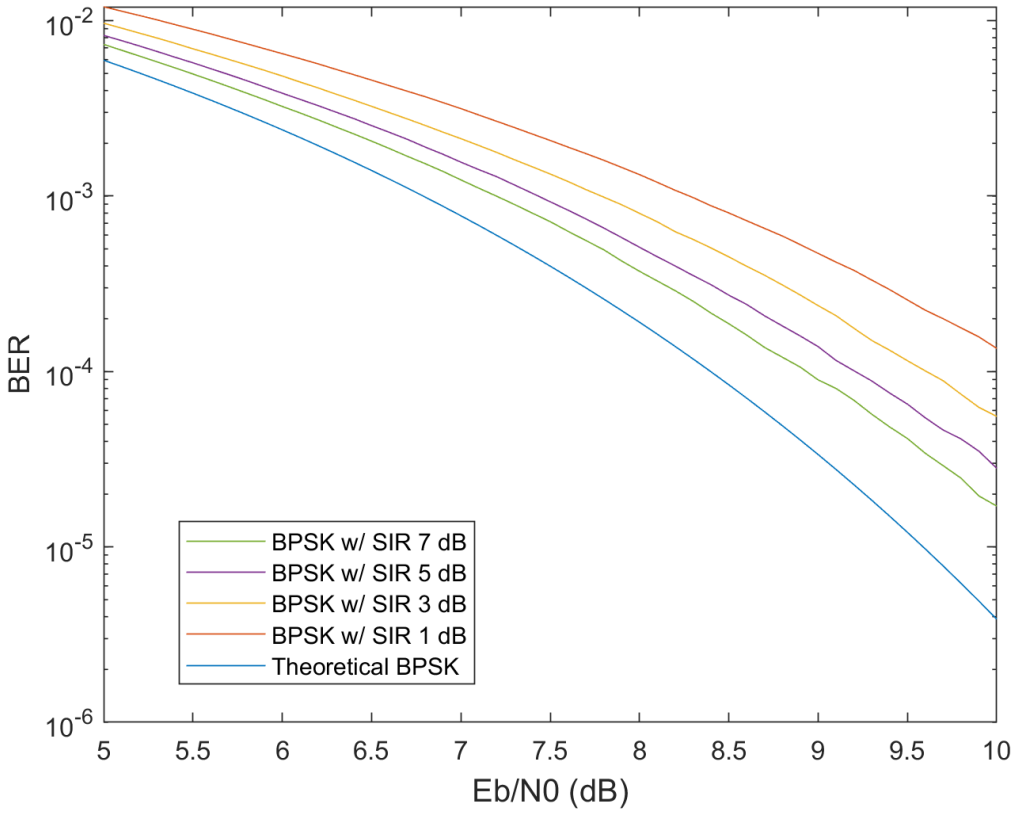

Figure 5.5: BER Loss using the Adversarial Waveform as a function of SIR

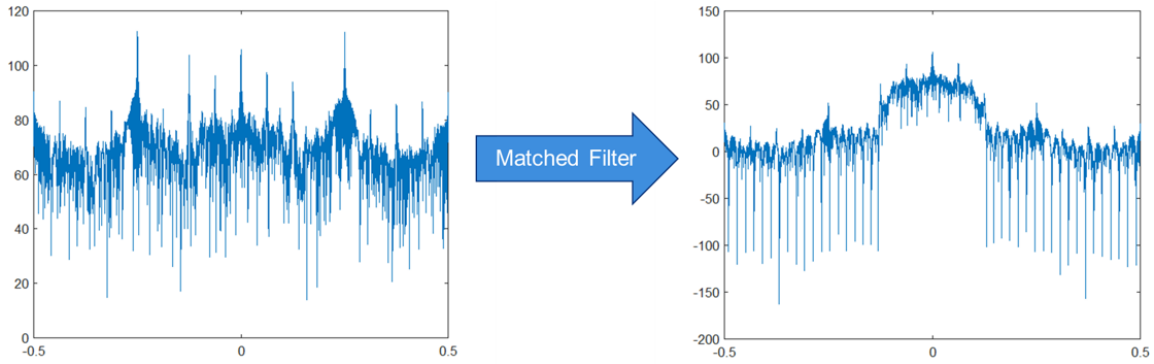

Figure 5.6: Matched Filter Effect on the Adversarial Waveform

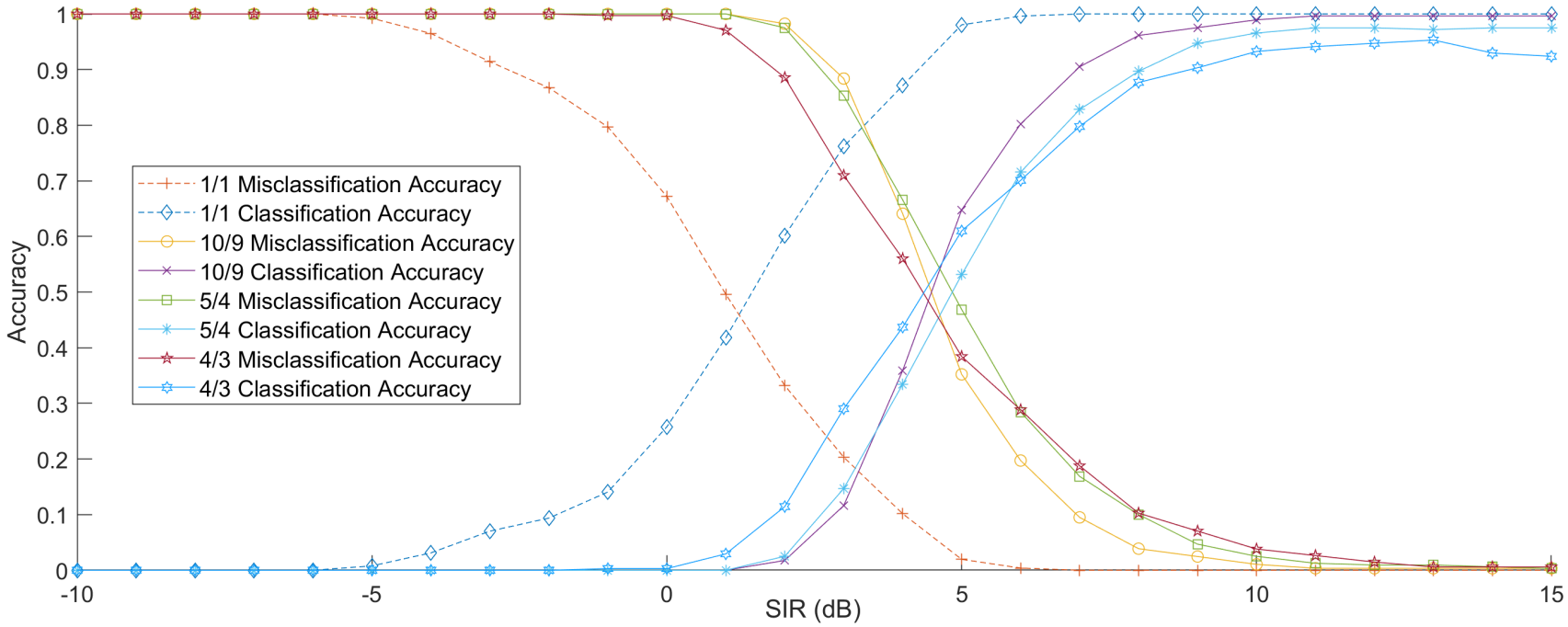

Figure 5.7: Transferability Results

5.4 Impact of the Adversarial Waveform on the Communication System

The impact of the adversarial waveform on the BER of the communication signal was tested. The proposed communication system making use of an adversarial waveform is illustrated in Fig. 5.4. The transmitter sends the communication signal and masks that signal with the adversarial waveform. The receiver of the communication signal uses a matched filter. A significant portion of the power of the adversarial waveform is outside of the bandwidth of the matched filter.

The loss to BER performance as a function of SIR is shown in Fig. 5.5. At an SIR of 5 dB the total loss in BER performance is about 1 dB in E_b/N_0.

The matched filter reduces the effective SIR at the receiver by suppressing a significant portion of the adversarial waveform. Thus the SIR at the receiver is significantly less than the SIR at the transmitter. This is illustrated in Fig. 5.6 which shows the filtered and unfiltered spectrum of the adversarial waveform. The matched filter reduces the impact of the adversarial waveform on the BER performance of the communication signal. The monitor does not have the benefit of this matched filter. The monitor cannot limit the bandwidth of the received signal to the specifications of the receiver node and also provide a generalized modulation classification function. In this work, the transmitter and receiver used RRC filters with an excess bandwidth of 0.35, as per the training data used for CNN-A. After the matched filter, the effective SIR increased by 16 dB.

5.5 Transferability of Attack

The adversarial noise from CNN-A was applied to a set of 512 randomly generated BPSK signals at varying SIRs and supplied them as batch input to the CNN-B

network. 512 randomly generated BPSK signals were generated and the adversarial noise was non-coherently combined at varying SIRs. Because the monitor is intercepting the transmission and does not have knowledge of the modulation scheme, thus the need for modulation classification, there will be uncertainty in the monitor's estimate of the symbol rate. In this experiment, the error between the symbol rate used by the communication system and the monitor was ranged between 1/1, 10/9, 5/4, and 4/3. These resample rates represent an increase in the bandwidth allocated for analysis at the spectrum monitor from 0% to 33%. The assumption is the spectrum monitor will err on the side of making the bandwidth wider rather than risk making the bandwidth too narrow and inadvertently filter the target out. The composite signal to be intercepted by the monitor was resampled to simulate the symbol rate mismatch. Then these resampled composite signals were input into CNN-B. CNN-B was trained on a sample rate of 4 samples per symbol. A resample rate of 1/1 provides that number of samples per symbol. A resample rate of 5/4 provides 5 samples per symbol, thus increasing the bandwidth allowed for analysis at the spectrum monitor.

The results are shown in Fig. 5.7. The classification accuracy of each resampling rate is plotted alongside the misclassification rate as functions of the SIR in dB. The misclassification accuracy measures how often the BPSK signal is misclassified as 2CPM or 4CPM. The best classification accuracy for each symbol rate can be seen at the highest SIR, and there is a decrease in the best accuracy as a function of the resample rate. However, the increase in the analysis bandwidth renders the spectrum monitor far more vulnerable to the adversarial waveform. Assuming the monitor can match the symbol rate of the communication system with 100% accuracy, that being a resample rate of 1/1, the SIR required to reduce classification accuracy to 50% is around 1 dB. This improves to an SIR of 5 dB when the monitor samples the intercepted signal at a rate of 5 samples per symbol, which is a resample

rate of 5/4. The impact of a resampling error is most severe from 1/1 to 10/9, where the SIR required to confuse CNN-B drops the most. This shows that only a small amount of uncertainty is required to increase the vulnerability of CNN-B to the transferred attack. Given that the monitor will not have knowledge of the symbol rate beforehand, this presents a significant vulnerability.

5.6 Conclusion

Our test presents means to attack spectrum monitors using modulation classifiers that rely on CNNs. CNN-A had a different structure, used a different symbol rate, was trained on a different data, and had a different set of modulation classes as compared to CNN-B. An adversarial waveform was generated from CNN-A and effectively transferred to CNN-B despite these differences. This experiment demonstrates the potency of the adversarial waveform attack. We trained the adversarial waveform using a method in which SIR is kept constant such that we could target a low SIR. We then varied the SIR and found that a stronger interferer was needed when transferring the attack against CNN-B.

Though the two CNNs had different modulation classes, the two sets were overlapping. As suggested in [5], the cause of this transferability phenomenon seen in this experiment could be that these different CNNs are learning similar functions as the output does have some overlap. Future research may determine how a transferred adversarial waveform would affect a CNN with a non-overlapping set of classes.

In addition to demonstrating the transferability, we also demonstrated that errors in symbol rate estimation at the monitor renders the modulation classification more susceptible to such attacks. The attack becomes more effective if the spectrum

monitor assigns a higher sampling rate. A small over-estimation of signal bandwidth provides a significant increase to the effectiveness of the attack. Designers of a spectrum monitor may be tempted to increase the bandwidth of the analysis of detected wireless activity; however, this comes at the cost of greater vulnerability to adversarial waveforms.

This technique can be used as obfuscation against classification, frustrating reverse engineering efforts, or potentially allow a signal detection to be discarded as an anomalous false alarm. Therefore this adversarial waveform can aid the covert signal developed in chapter 2 and [3].

It has been demonstrated that in a realistic environment, where the monitor is uncertain of the symbol rate, the monitor is far more susceptible to such attacks. Such attacks make reverse engineering more expensive while maintaining the usability of the communication signal. Future research may be able to use this technique to strengthen classifiers. Additionally, physical layer anomaly detection may reveal the presence of an adversarial waveform masking the modulation type.

5.7 References

[1] Samuel Bair et al. "On the Limitations of Targeted Adversarial Evasion Attacks Against Deep Learning Enabled Modulation Recognition". In: *Proceedings of the ACM Workshop on Wireless Security and Machine Learning*. WiseML 2019. Miami, FL, USA: Association for Computing Machinery, 2019, 25–30. ISBN: 9781450367691. DOI: 10.1145/3324921.3328785. URL: https://doi.org/10.1145/3324921.3328785.

[2] Daniel Chew et al. "Adversarial Attacks on Deep-Learning RF Classification in Spectrum Monitoring with Imperfect Bandwidth Estimation". In: *Accepted to 2022 IEEE Wireless Communications and Networking Conference (WCNC)*. 2022. DOI: 10.1109/CISS48834.2020.1570617443.

[3] Daniel Chew et al. "Covert Communications through Imperfect Cancellation". In: *Accepted to Proceedings of the 2022 ACM Workshop on Information Hiding and Multimedia Security*. IH&MMSec '22. New York, NY, USA: Association for Computing Machinery, 2022.

[4] Bryse Flowers, R. Michael Buehrer, and William C. Headley. "Evaluating Adversarial Evasion Attacks in the Context of Wireless Communications". In: *IEEE Transactions on Information Forensics and Security* 15 (2020), pp. 1102–1113. DOI: 10.1109/TIFS.2019.2934069.

[5] Ian Goodfellow, Jonathon Shlens, and Christian Szegedy. "Explaining and Harnessing Adversarial Examples". In: *International Conference on Learning Representations*. 2015. URL: http://arxiv.org/abs/1412.6572.

[6] Muhammad Zaid Hameed, András György, and Deniz Gündüz. "Communication without Interception: Defense against Modulation Detection". In: *2019 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*. 2019, pp. 1–5. DOI: 10.1109/GlobalSIP45357.2019.8969541.

[7] Muhammad Zaid Hameed, András György, and Deniz Gündüz. "The Best Defense Is a Good Offense: Adversarial Attacks to Avoid Modulation Detection". In: *IEEE Transactions on Information Forensics and Security* 16 (2021), pp. 1074–1087. DOI: 10.1109/TIFS.2020.3025441.

[8] Brian Kim et al. "Over-the-Air Adversarial Attacks on Deep Learning Based Modulation Classifier over Wireless Channels". In: *2020 54th Annual Conference on Information Sciences and Systems (CISS)*. 2020, pp. 1–6. DOI: 10.1109/CISS48834.2020.1570617416.

[9] Diederik Kingma and Jimmy Ba. "Adam: A Method for Stochastic Optimization". In: *International Conference on Learning Representations* (2014).

[10] Mathworks. *Modulation Classification with Deep Learning*. https://www.mathworks.com/help/comm/examples/modulationclassification-with-deep-learning.html, accessed 2020-03-05. 2019.

[11] Timothy J. O'Shea, Johnathan Corgan, and T. Charles Clancy. "Convolutional Radio Modulation Recognition Networks". In: *Engineering Applications of Neural Networks*. Ed. by Chrisina Jayne and Lazaros Iliadis. Cham: Springer International Publishing, 2016, pp. 213–226. ISBN: 978-3-319-44188-7.

[12] Meysam Sadeghi and Erik G. Larsson. "Adversarial Attacks on Deep-Learning Based Radio Signal Classification". In: *IEEE Wireless Communications Letters* 8.1 (2019), pp. 213–216. DOI: 10.1109/LWC.2018.2867459.

[13] Christian Szegedy et al. "Intriguing properties of neural networks". In: *International Conference on Learning Representations*. 2014. URL: http://arxiv.org/abs/1312.6199.

Chapter 6

Conclusion

6.1 Summary, Discussion, and Future Work

This dissertation began with the notion that covert communications can be a **force for good**, that covert communications could protect privacy and security. In order to advance the state-of-the-art in covert communications, this research focused on physical layer steganography. Physical layer steganography represents the cutting edge of covert communications waveforms. This chapter will summarize the work and offers thoughts on the road ahead.

There were several goals for the covert signal developed in this work:

- It is the responsibility of the covert signal to mask itself in the incumbent, no cooperation from the incumbent can be expected,
- The interference from the covert signal must not inflict a noticeably reduce the throughput of the incumbent signal or else the covert signal may be exposed,
- The covert receiver does not have a copy of the incumbent signal in advance,
- The covert link must provide a sizeable throughput, and
- The covert transmitter must embed resistance in the covert signal to classification and reverse engineering in order to frustrate any exploitation efforts in

the event of signal detection.

In each chapter of this dissertation, one or more of these goals was addressed.

In chapter 5 and [3], a covert signal was developed using interference cancellation to remove the incumbent OFDM signal, and the recovery of that covert signal was demonstrated using OTA data. In the context of this interference channel, it was shown that SIR, not SNR, is the primary limitation of the covert signal sharing the link with the OFDM signal. Future research on the covert method described in this work includes but is not limited to increasing the modulation order of the covert signal in order to increase the data rate and more direct comparisons with methods like those described in [4] and [5].

A novel signal detection method was developed that can detect signals in the presence of interference and without an explicit noise floor estimate in chapter 3 and [1]. The covertness of the hidden signal was tested against this and other detectors. The CNN detector achieved CFAR performance by being trained on specific SNRs, and does not require an estimate of the noise floor. The CNN detector demonstrated superior performance as compared to the energy detector in the presence of an interferer. It may be possible to repeat these successes in signal detection with a CNN less complex than AlexNet. Reducing the computational complexity of the CNN detector is one avenue of future work. A CNN detector could be expanded to tackle other interference sources and jamming threats. This would require identifying the interference, simulating the interference, and training the CNN detector to ignore the interference.

In order to facilitate cancelling the incumbent, a parsimonious model of OFDM signalling was developed prioritizing the effect of the model parameters on cancellation in chapter 4. It was demonstrated that OFDM windowing at the transmitter has a significant impact on cancellation applications. This work developed an

algorithm for OFDM window estimation in a generalized form that does not require foreknowledge of the window implementation. This work then evaluated that algorithm and applied it to cancel the self-interference resulting from OFDM windowing, resulting in an improvement in the PER of existing OFDM systems. The estimation of the OFDM window provided significant performance improvement in cancellation applications. The cancellation technique developed in this research is applicable to many situations in which signals are transmitted within the same spectrum resource as an OFDM signal, specifically in the development of new cellphone standards.

Lastly, in chapter 5 and [2], an attack was developed against automated classification tools in order to further protect the privacy and security of the covert signal. This attack can be used as obfuscation against classification, frustrating reverse engineering efforts. The attack became more effective if the spectrum monitor assigns a higher sampling rate. A small over-estimation of signal bandwidth provides a significant increase to the effectiveness of the attack. Designers of a spectrum monitor may be tempted to increase the bandwidth of the analysis of detected wireless activity; however, this comes at the cost of greater vulnerability to adversarial waveforms. Future research may determine how a transferred adversarial waveform would affect a CNN with a non-overlapping set of classes. Future research may be able to use this technique to strengthen classifiers. Additionally, physical layer anomaly detection may reveal the presence of an adversarial waveform masking the modulation type.

With the capabilities of this covert waveform tested and quantified, this research task concludes as a completed case study into physical layer steganography applied to multicarrier transmissions.

6.2 References

[1] Daniel Chew and A. Brinton Cooper. "Spectrum Sensing in Interference and Noise Using Deep Learning". In: *2020 54th Annual Conference on Information Sciences and Systems (CISS)*. 2020, pp. 1–6. DOI: 10.1109/CISS48834.2020.1570617443.

[2] Daniel Chew et al. "Adversarial Attacks on Deep-Learning RF Classification in Spectrum Monitoring with Imperfect Bandwidth Estimation". In: *Accepted to 2022 IEEE Wireless Communications and Networking Conference (WCNC)*. 2020, pp. 1–6. DOI: 10.1109/CISS48834.2020.1570617443.

[3] Daniel Chew et al. "Covert Communications through Imperfect Cancellation". In: *Accepted to Proceedings of the 2022 ACM Workshop on Information Hiding and Multimedia Security*. IH&MMSec '22. New York, NY, USA: Association for Computing Machinery, 2022.

[4] Aveek Dutta et al. "Secret Agent Radio: Covert Communication through Dirty Constellations". In: *Information Hiding*. Ed. by Matthias Kirchner and Dipak Ghosal. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 160–175. ISBN: 978-3-642-36373-3.

[5] Salvatore D'Oro, Francesco Restuccia, and Tommaso Melodia. "Hiding Data in Plain Sight: Undetectable Wireless Communications Through Pseudo-Noise Asymmetric Shift Keying". In: *IEEE INFOCOM 2019 - IEEE Conference on Computer Communications*. 2019, pp. 1585–1593. DOI: 10.1109/INFOCOM.2019.8737581.

Curriculum Vitae

Daniel Chew received his B.E.E. from the University of Delaware in 1998 and an M.S.E.E. in Electrical Engineering from Johns Hopkins University in 2008.

He has held positions at Thales Group, The Boeing Company, and is currently a member of the Principal Professional Staff at the Johns Hopkins University Applied Physics Laboratory. He enrolled in the Johns Hopkins University Whiting School's Doctor of Engineering program in 2019.

He teaches 525.751 "Software Radio for Wireless Communications", 525.752 "Digital Receiver Synchronization Techniques", and 525.201 "Circuits, Devices, and Fields" in the JHU Whiting School Engineering for Professionals program.

He has written two books, The Wireless Internet of Things: A Guide to the Lower Layers (2018) and Wireless Coexistence: Standards, Challenges, and Intelligent Solutions (2021). He has one patent, "Relativistic Wireless Channel Emulator", U.S. Patent 10498470.

His research interests include the Internet of Things, Wireless Communication Systems, Machine Learning, and Software-Defined Radios.