ABSTRACT

Title of dissertation: EXTRINSIC CHANNEL-LIKE FINGERPRINTING FOR TRANSMITTER AUTHENTICATION IN WIRELESS SYSTEMS

Nathan Goergen, Doctor of Philosophy, 2011

Dissertation directed by: Professor K. J. Ray Liu
Department of Electrical and
Computer Engineering

EXTRINSIC CHANNEL-LIKE FINGERPRINT EMBEDDING FOR TRANSMITTER AUTHENTICATION IN WIRELESS SYSTEMS

by

Nathan Goergen

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2011

Advisory Committee:
Professor K. J. Ray Liu, Chair/Advisor
Professor Richard La
Professor Min Wu
Professor T. Charles Clancy
Professor Lawrence Washington, Dean's Representative

Dedication

To my wife Katie and my parents. Their patience has made this possible.

Acknowledgments

I would like to thank my advisor, Professor K. J. Ray Liu, for his guidance and advice. His understanding and extreme patience has made this dissertation possible. I would also like to thank my co-author and lab partner, Wan-Yi Lin, for her painstaking proofreading help, and for her excellent ideas. Perhaps now, she will finally have the chance to get to bed at a decent hour.

I would also like to thank my employer, the Laboratory for Telecommunication Sciences (LTS), for their support throughout my graduate career.

Thanks are also in order for Professor Wu, Professor La, Professor Washington, and Dr. Clancy for agreeing to serve on my thesis committee and for sparing their invaluable time reviewing this manuscript and the works herein.

I owe my deepest thanks to my family - my wife, mother, father, brother, and sister, who have always stood by me. They have supported me through the toughest times. Words cannot express my gratitude.

I would like to acknowledge the financial support of the Science, Mathematics, And Research for Transformation (SMART) program, and Dr. Clancy's help, for work presented here.

Table of Contents

List of Figures

List of Abbreviations

3GPP	3rd Generation Partnership Project
ATSC	Advanced Television Systems Committee
AWGN	additive white Gaussian noise
BER	bit error rate
BPSK	binary phase-shift keying
CDMA	code division multiple access
CR	cognitive radio
CSI	channel state information
DBPF	data-bearing pilot frameworks
DFT	discrete Fourier transform
DLL	digital locked loop
DSA	dynamic spectrum access
DTV	digital television
FCC	Federal Communications Commission
FEC	forward error correction
FFT	fast fourier fransform
FIR	finite impulse response
IDFT	inverse discrete Fourier transform
IEEE	Institute of Electrical and Electronics Engineers
ISI	inter-symbol interference
LLMSE	linear least mean-squared error
LMS	least mean-squared
MAC	media access control
MIMO	multiple-input multiple-output
MMSE	minimum mean-squared error
MRC	maximum-ratio combining
MSE	mean-squared error
OFDM	orthogonal frequency division multiplexing
OSI	Open Society Institute
PAM	pulse amplitude modulation
PEP	pairwise error probability
PHY	physical
PLL	phase locked loop
PSAM	pilot signal assisted modulation
PUE	primary user emulation
QPSK	quadrature phase-shift keying
SDR	software defined radio
SER	symbol error rate
SFB	space-frequency block
SHA	secure hash algorithm
SISO	single-input single-output
SNR	signal-to-noise ratio
STC	space-time code

STF	space-time-frequency
SVD	singular value decomposition
TFB	time-frequency block
TM	time-multiplexed
VSB	vestigial sideband
WRAN	Wireless Regional Area Networks
WSSUS	wide-sense-stationary uncorrelated scattering

Chapter 1

Introduction

The security of wireless systems has become an extensively researched topic due to the inherent broadcast nature of the medium. Transmissions sent over wireless links must be protected from malicious message interception, forgery, modification, deletion, and replay. The security mechanisms of the system, such as user *authentication*, help mitigate acts of message forgery and the nefarious actions of impostors. However, when the authentication mechanisms of a wireless system fail or are compromised the system becomes vulnerable to a number of identity-based attacks, like those explored by our work in [14, 9], and in the works of [14, 48, 53, 12, 11].

To prevent such attacks, authentication messages have been widely used in upper-layer network protocols. However, when the authentication message is embedded in an upper-layer protocol, a node must perform a significant amount of processing before the message can be checked for validity. Unauthenticated messages must be systematically processed and decoded without error at every layer of the protocol stack, before the authentication message can be extracted and the credentials of the message verified. These sophisticated upper-layer processing tasks expose the receiver to potentially malicious transmissions designed to exploit vulnerabilities in upper-layer protocols and the implementation thereof. Additionally, when the authentication message is conveyed using an upper-layer protocol, it is received with a quality no greater than the quality of the trans-

mission's payload message. Wireless transmissions that can be authenticated immediately at the physical (PHY) layer, however, help protect upper-layer protocol processing algorithms from these malicious transmissions.

PHY-layer authentication offers a number of computational benefits in addition to the security considerations mentioned above. When a wireless signal can be authenticated using only its PHY-layer characteristics, the source of the signal can be verified before the signal is demodulated or decoded. Therefore, wireless nodes that must verify the authenticity of a signal, but are otherwise uninterested in the content of the signal, can avoid wasteful processing steps when unintended, uninteresting, or maliciously fabricated Denial of Service style transmissions are received. This allows nodes to more quickly authenticate legitimate users within cooperative communication systems, and implicate the charlatans of the system. Additionally, PHY-layer authentication approaches allow for a completely independent authentication mechanism that is decoupled from upper-layer system design. This allows designers of wireless systems to create authentication devices that are invariant of upper-level protocol mutability.

While there are a number of advantages when leveraging PHY-layer signal authentication methods over traditional 'bit-level' authentication, in this work we discuss an authentication messaging system that does not necessarily replace traditional messaging used in higher layer protocols. Instead, we consider a method for embedding an authentication message into wireless transmissions that is capable of transmitting a digital authentication message not unlike the authentication messages sent by traditional protocols. The embedded message is completely independent from the upper-layer protocols, and thus, independent authentication messages sent via both methods may be used to redundantly

authenticate the transmission.

The independence offered by our fingerprint embedding scheme provides another advantage, in terms of signaling performance. Since the specifications of the fingerprint signaling scheme can be defined independently from the specification of original message in which the authentication message is embedded, key signaling parameters that define the synchronization mechanism, for example, can be relaxed to mitigate the effects of noise and various types of interference. Very little system bandwidth is required for the authentication message in most practical communication systems, because the fingerprint message is typically selected to be very small compared to the bandwidth required by the original transmission to minimize system overhead. This allows a number of signal spreading and coding schemes to be leveraged when designing fingerprint. Thus, robust fingerprint signaling designs can be devised that allow for transmission authentication in scenarios where original signal itself is unrecoverable due to low signal-to-noise ratio (SNR) or fading conditions.

1.1 Prior Work

Message fingerprinting techniques, which append or embed a message conveying the credentials of a data source, have been successfully applied to a number applications such as multimedia systems. One such application providing secure transmission of multimedia content is described in [46].

In addition to multimedia applications, a number of PHY-layer fingerprinting approaches for wireless communications have been investigated using rudimentary signal

superposition methods. In [61] the superposition of low-power pseudo random sequences onto digital television transmissions is discussed. In [47], multi-resolution approaches are considered, where a low-power fingerprint constellation is superimposed onto the original signal constellation. Multi-resolution approaches use a low-power *high resolution* signal to carry the authentication message, while the main transmission is conveyed by a *low resolution* signal. The use of wavelet transforms have also been explored to achieve a multi-resolution composite signal. PHY-layer superposition approaches have been studied for the specific case of telephone signals [56], where an encrypted hash of the conversation is superimposed onto the original signal. In [37] a low-power spread spectrum watermarking scheme is discussed, where the watermark is directly superimposed onto the signal. Physical layer approaches to authentication have also been explored for the case of wired Ethernet signals [39]. This work attempts to identify unique network adapters in a wired network using PHY-layer signatures. General signal fingerprinting through superposition methods are discussed in [66].

While these previous works use superposition to embed the authentication fingerprint signal onto the original transmission, there are a number of drawbacks to this approach. The main disadvantage of basic superposition techniques is that the fingerprint signal will appear as additional noise when receiving the original signal, since the authentication signal is fully present when the signal is decoded. Thus, superposition-based methods inject the fingerprint signal as interfering noise which reduces the SNR of the original signal.

To improve on the drawbacks of blind superposition approaches, we propose a fingerprint embedding approach that exploits typical receiver preprocessing algorithms such

as channel equalization, signal phase recovery, and timing recovery mechanisms, with a design methodology closely resembling the Category 2 and Category 3 fingerprints discussed in [17]. These fingerprints are designed according to anticipated channel distortions, and with consideration of how the original signal will be *perceived* by the receiver. Using these extra pieces of information, the undesirable effects associated with blind superposition approaches, representing Category 1 designs [61, 47, 56, 37, 39, 66], are partially removed by the receiver through traditional channel equalization and signal synchronization practices. This filtering of the fingerprint signal before the original signal data signal is perceived represents a substantial performance improvement over prior works.

1.2 Relationship between Intrinsic and Extrinsic Fingerprints

In the previous section we discussed the advantages of PHY-layer authentication, we presented a number of prior works in the area of message fingerprinting, and we described how we intend to leverage the perceptual model of wireless receivers and side information of the anticipated channel distortions that the fingerprint signal will encounter, to improve on the results presented in prior works. In this section we will briefly describe related works in which natural channel distortions are leveraged to authenticate transmissions, and we will describe how the fingerprinting work presented here can augment and improve upon these schemes.

A number of bodies of work have demonstrated how channel distortions themselves can be leveraged for signal authentication purposes. In [64] it was demonstrated that

robust PHY-layer fingerprints can be obtained from intrinsic features characteristic of wireless channels, such as unique scattering environments, spatial variability, and time-varying channel distortions. This work demonstrated that transmitters may be validated when the multipath channel profiles for each transmitter are unique and sufficiently stationary, and provides the initial edifice for a PHY-layer authentication approach that is cognizant of channel distortions. Intrinsic fingerprinting mechanisms have also been considered to mitigate Sybil attacks [65]. However, when channel conditions are not conducive to intrinsic fingerprint recognition, due to either highly correlated multipath profiles between transmitters or rapidly varying channel conditions, a more robust PHY-layer signature is required to authenticate wireless nodes.

To overcome the limitations of authentication schemes that leverage only intrinsic channel information, in this dissertation we consider augmenting intrinsic channel-based authentication approaches, such as [64], with an extrinsic synthetically generated fingerprint signal applied by the transmitter, in pursuit of a channel-aware PHY-layer authentication framework encompassing both intrinsic and extrinsic fingerprint features. Specifically we are interested in PHY-layer fingerprints designed in the spirit of the Category 2 and Category 3 approaches in [17] exhibiting *channel-like* properties commensurate with the natural channel distortions mitigated by traditional receiver perceptual models. Specifically, we extend fingerprinting approaches that leverage intrinsic channel features alone, i.e. [64] and [65], by augmenting these approaches with an extrinsic synthetically-generated signal.

In this dissertation we describe how the extrinsic, synthetically generated signal is perceived by receivers in a way that is similar to the perception of natural time-varying

channel distortions. Thus, we use the term *channel-like* to denote a class of extrinsic fingerprint signals designed to manipulate parameters of the transmitted signal via methods that may be modeled as time-varying channel distortions. We will demonstrate how these distortions are subsequently corrected by the receiver through traditional channel equalization and synchronization practices. These signals, both the natural intrinsic channel distortion and the synthetic extrinsic signal, are thus filtered from the primary signal in pragmatic receiver designs. In other words, both the extrinsic channel-like fingerprint signal and real-world intrinsic time-varying channel distortions become an aggregate signal, and will be jointly mitigated by the receiver before the primary signal is decoded (i.e. *perceived*.)

Because both the intrinsic channel distortion and the extrinsic signal will be perceived by the receiver as an aggregate signal, we demonstrate how the intrinsic channel distortions become interference when recovering the extrinsic fingerprint signal. While channel estimation aids, such as pilot signals, help mitigate the time-varying channel distortions experienced by the the original signal, intrinsic channel distortions will always be present in the aggregate signal. Therefore the authentication signal must be carefully designed to overcome these distortions. Multipath fading, fluctuations in gain and/or phase, symbol timing, center frequency offsets, and Doppler effects are all distortions indicative of real, intrinsic channels. Therefore, a fingerprint signal that manipulates one or more of these parameters to convey an authentication message constitutes a channel-like design.

We have mentioned that the authentication devices employed by a system must secure against forgery and replay. Because our fingerprint message is conveyed as an independent digital message, the bit-level cryptographic primitives typically used to secure

authentication messages in upper-layer protocols may also be leveraged in the design of the authentication message. By leveraging proven, best security practices in the design of the digital message conveyed by the fingerprint, the forgery and replay resistance of the fingerprint message will be at least commensurate with state-of-the-art authentication messages used in upper layers.

While many fingerprinting schemes consider homogeneous systems, where all nodes within the system can modified to implement the authentication signal, we instead consider a heterogeneous wireless broadcast system with two classes of receivers:

- The *unaware receiver*: regular, unmodified, receivers that will ignore the fingerprint signal and employ traditional channel equalization and data detection
- The *aware receiver*: receivers designed to detect and decode the fingerprint in addition to the primary signal

By considering heterogeneous systems, were *unaware* receivers will also receive fingerprinted messages, we ensure that our fingerprint designs do not degrade the performance of legacy receivers. In the coming sections we will describe how our channel-like fingerprinting scheme emulates distortions typically observed in wireless channels, allowing unmodified receivers to operate unhindered when the fingerprint message is present.

Now that we have defined classes for the type of receivers within our system, we now define two classes of fingerprint designs according to the amount of information available to the transmitter. The **Open-Loop** class of fingerprint designs include applications where channel state information (CSI) in unavailable to the transmitter. Fingerprint designs in this class must be created and embedded blindly into the original signal, since

side information on channel distortions that the transmission may encounter are unknown. Fingerprints in the Open-Loop class must be designed such that their probability of detection is optimized over the entire range of anticipated time-varying channel conditions. Careful consideration of transmission signaling parameters, anticipated channel distortions, and the perceptual model of the receiver must be taken to develop optimal designs.

In contrast to the Open-Loop class of fingerprint designs, the **Closed-Loop** class of fingerprints includes applications where partial CSI is known by the transmitter. Since the extrinsic fingerprint and intrinsic channel distortion will form an aggregate signal, these two signals could potentially interfere with one another. A transmitter in a Closed-Loop system can leverage this additional channel side information to create fingerprint designs that dynamically change according time-varying channel state. Because of the close relationship between intrinsic channel distortions and the channel-like fingerprint signal, in this dissertation we pay close attention to the channel estimation mechanisms comprising the perceptual model of a receiver, to ensure that channel state information is effectively applied when designing the fingerprint.

1.3 Brief Overview of Contributions

Now that we have introduced the problem space for fingerprinting wireless signals, we now briefly discuss the main contributions of this work.

In [24], an extrinsic channel-like fingerprint is considered for narrowband single-input-single-output (SISO) systems using pulse-amplitude modulation (PAM) signaling, where the fingerprint message is embedded at the transmitter by synthetically applying

nominal multipath channel responses. This work considers blind channel estimation, where pilot signals are not present, to aid in the channel estimation process. An authentication signal BER less than the primary signal is achieved, and nearly zero degradation to the original signal, due to the presence of the embedded fingerprint, is demonstrated. These results are briefly discussed as a motivating channel-like fingerprinting example in section 1.5.

In Chapter 2 we extend our work in the narrowband SISO case to designs that operate in multiple dimensions. In [27] and [26], a multiple-input-multiple-output (MIMO) Closed-Loop system is considered, and a framework for extrinsically fingerprinting space-time coded (STC) transmissions is presented. A fingerprint function jointly manipulating both the time and spacial domains is explored, and we discuss how the fingerprint distortions are removed by the channel equalization device employed by the receiver. An authentication signal BER less than the primary signal is achieved, with no impact to the original signal.

In Chapter 3 we consider another multi-dimension fingerprinting scenario, extending our work in Chapters 2 and 3 to Open-Loop systems. In [29] and [28], orthogonal frequency division multiplexing systems (OFDM) are considered, and a number of fingerprint designs are discussed that jointly manipulate the fingerprinting space in both time and frequency via an "overlay" signal. Knowledge of time-varying channel state is leveraged in these designs to create improved fingerprint signals. A taxonomy of Open-Loop designs are presented that leverage various degrees of channel state knowledge to design the embedded fingerprint signal overlay. An authentication signal BER less than the primary signal is once again achieved, with no perceived impact to the original signal.

While not presented in this dissertation, in [25] we improve upon our designs in [29] and [28] by applying predictive filtering to reduce one type of fingerprint detection error. In this work, predicted channel-state information (CSI) is leveraged when creating the fingerprint, to increase the robustness of the fingerprint to interference caused by time-varying channel distortions. A fingerprint BER improvement of nearly 8-9dB over [29] and [28] is demonstrated.

Now that we have discussed the main contributions of the work in this dissertation, we will briefly discuss the key assumptions that have been made to tie these works together. All fingerprinting approaches in this dissertation leverage the key assumption that consecutive channel estimates are correlated in time, to some degree. Specifically, when channel conditions are slowly changing, the sequence of channel estimates produced by the receiver using the pilot signals embedded in the transmission contain redundant information. Channel-like fingerprinting exploits this correlation to transmit a new piece of information to the receiver, the fingerprint signal, via explicit manipulation of the transmitted signal. The fingerprint signal is typically detected by comparing time-varying channel estimates, i.e. estimates derived from embedded pilot signals. Thus, one may also think of channel-like fingerprinting approaches as explicit manipulation of the transmitted pilot signals, where pilot signal distortions are merely duplicated onto the payload signal to keep these two signals synchronized.

While the focus of this work is to apply this signal fingerprinting technique to wireless authentication problems, in general our fingerprint signal is not limited to authentication applications. Indeed, messages of all sorts may be transmitted by the fingerprint signal, to suit a number of purposes. This being said, at first glance our work in [30] and

[31], presented in Chapter 5, might seem very removed from the fingerprint embedding discussions of Chapters 1 through 4. However, we will explain how these bodies of work share some interesting similarities.

The works of [30] and [31] presented in Chapter 5 consider a four node relay system where the relayed signal is transmitted using a *best-effort* delivery policy, with energy reallocated from the pilot signals of the transmission. The power allocation parameters of the new relayed signal change in response to the time-varying channel. Upon closer inspection we see that this body of work leverages the same channel stationarity assumptions as our fingerprinting scheme, to transmit the new relay transmission signal. In short, the relay signal is created through manipulating embedded pilot signals in response to time-varying channel conditions. We note that the fingerprint signal discussed in Chapters 1 through 4 can be described using a very similar description.

Through this argument, we see an ulterior connection between these two bodies of work: the fingerprinting method discussed in Chapters 1 through 4 and best-effort delivery method discussed in Chapter 5. When the amount of new information in a sequence of channel estimates decreases due to stationary channel conditions, the information contained in the pilot signals from which the channel estimates are derived decreases accordingly. While bandwidth resources devoted to pilot signals have historically been considered useless overhead in the transmission, conveying zero information to the receiver, in both bodies of work we will show that through strategic manipulation of the pilot signals we can actually convey additional information to the receiver. Pilot signals convey information about time-varying channel distortions to the receiver, and when little new information can be gleaned from pilot signals, the information conveyed by these

signals can be increased by manipulating the transmission to modulate a new signal. In our relay work the new signal is the relayed message, while in our fingerprinting work this new signal is the fingerprint message.

Now that we have discussed how the best-effort relaying work in Chapter 5 is similar to the fingerprinting methods of Chapters 1 through 4, we will briefly discuss how a similar argument can be made to relate the works in Chapters 1 through 4 to the work in Chapter 5. While many time-varying channel models exist to help explain the behavior of mobile wireless channels, true channel state in a real world system can be a very unpredictable indeed. Thus the channel stationary assumptions used in this dissertation will not accurately describe all types of channel scenarios that a wireless node will encounter. In very unpredictable channels the channel estimation aids, i.e. the pilot signals, are exercised to their fullest to convey timely channel state information to the receiver. In unpredictable and rapidly-varying channel conditions the new information conveyed by the transmitter, be it a fingerprint signal or a relayed signal, will be severely degraded or lost entirely as the intrinsic channel distortion information measured by pilot signals strongly interferes with the new signal. By this definition we conclude that the channel-like fingerprint signals, like the relay signals of [30] and [31], are transmitted using a best-effort delivery model as well.

Now that we have discussed how our channel-like fingerprinting approach will leverage pilot-signal resources to convey new information to the receiver, we would like to briefly discuss the amount of overhead that pilot signals impose on modern mobile waveforms. Pilot signals or training data, used primarily for channel estimation and synchronization purposes, represent a significant resource and energy overhead in modern

waveforms [35], [1], [2], [10]. In IEEE 802.16e for example, the mobile version of the WiMax specification, 11 percent or more of bandwidth resources are reserved for pilot signals in the downlink signal while 33 percent or more of the bandwidth of the uplink signal is consumed for the purpose of channel estimation [35]. As pilot signals typically do not convey any useful payload data to the receiver, these bandwidth resources are generally wasted outside of their primary purpose of channel estimation. Thus it is obvious that any method endowing pilot signals with the capability of transmitting useful payload data, while also allowing these signals to fulfill their original purpose in conveying timely channel state information to the receiver, demonstrates a substantial reduction in system overhead and an accompanying increase in system efficiency. This potential reduction in system overhead represents an ancillary contribution of this work that is most interesting.

1.4 Overall Problem Formulation and Design Goals

Now that we have given an overview of our fingerprinting approach, we will now formally state the problem space and goal of this work. In this dissertation, we aim to design channel-like fingerprint signaling schemes according to three design goals:

- Design Goal 1: The fingerprint signal should cause minimal degradation to the primary data signal and the detection thereof, when present.
- Design Goal 2: The fingerprint signal should achieve near *fingerprint channel* capacity, where in this work the "fingerprint channel" is band-limited by the characteristics of the channel estimation mechanism employed by the transmitter/receiver.
- Design Goal 3: The broadcast authentication message conveyed by the fingerprint

function should be resistant to common attacks, including forgery, modification, and replay.

- Design Goal 4: Optionally, we can also consider an equi-energy transmission constraint, where fingerprinted transmissions are sent according to the original transmission energy budget.

We make special note of Design Goal 2, which will be clarified in Chapter 3. Because our channel-like fingerprinting mechanism uses embedded pilot signals to perform channel estimation and convey fingerprint information as an aggregate transmission, in Chapter 3 we will demonstrate how the time-varying intrinsic channel response and projection of receiver noise onto the received pilot signals become interference during fingerprint detection, resulting in decreased *fingerprint channel* capacity.

1.5 Motivating Work

We will now briefly discuss our work in [24], which describes a pragmatic channel-like fingerprinting scenario applicable to the IEEE 802.22 Digital Television Whitespace FCC ruling. These results will serve as a motivating example for the work to be presented in the coming chapters. This work will also provide an example for scenarios where wireless nodes must detect and decode the fingerprint transmission to operate correctly, yet may be uninterested in the original signal's payload content.

With the ubiquitous adoption of wireless communications, access to the electromagnetic spectrum has become increasingly competitive. To facilitate efficient access to shared spectrum, an arbitration method known as Dynamic Spectrum Access (DSA) was

recently proposed [4]. Motivated by the FCC ruling regarding Wireless Regional Area Networks (WRAN), IEEE 802.22 [20], interest in spectral sensing and shared spectrum technologies has dramatically increased. Under this standard, access to the unused space between Digital Television (DTV) channels, or the white space spectrum, is granted to next-generation wireless broadband equipment. In shared spectrum scenarios such as those proposed by 802.22, licensed DTV stations are considered primary users and are given explicit first-right-of-access to the television spectrum, while broadband users are allowed access to the shared spectrum as secondary users in the absence of primary user signals only.

To ensure efficient use of white space spectrum under IEEE 802.22, spectrum allocations must first be sensed for primary users such as DTV and wireless microphone signals before secondary users exhibiting varying bandwidths and modulation types are granted access to an allocation. Since classification accuracy is exigently required for correct spectral usage, a number of methods have been proposed for the detection and classification of signals in DSA environments. Traditional approaches to unknown signal identification involving the computation of various statistical properties [44] or cyclostationary features [22, 57] have been proposed. Further detection and classification of signals using these features has been discussed, including machine learning and policy-based classification engines. In DSA environments such as the DTV spectrum, robust spectrum sensing devices are required for the interoperability, and correct operation of smart radios.

Previous work has demonstrated the utility of machine learning approaches to signal classification particularly in DSA applications. However, recent work [14, 48] has

shown potential weaknesses in these approaches. When using unsupervised learning in non-cooperative environments, adversaries may easily manipulate the learning process compromising security and exposing DSA systems to a number of node identity attacks. To prevent such attacks, the proposed physical layer authentication mechanism introduces an unambiguous and explicit feature to the transmitted signal, providing stronger user authentication capabilities to cognitive radios than those afforded by statistical and cyclostationary features alone. Additionally, physical layer approaches facilitate the authentication of unknown signals before higher layer processing, allowing smart radios to quickly identify its source.

In [24] we describe an approach to primary user authentication (PUA) in the DTV spectrum, specifically targeting the Advanced Television Systems Committee (ATSC) broadcast standard. Since malicious nodes are motivated to conduct PUE attacks to obtain increased bandwidth allotments and unfettered use of spectrum, the fingerprint message is used to expose attackers and protect primary user resources. The application of the fingerprint signal to existing ATSC signals is discussed, and simulation results demonstrating minimal impact to DTV coverage area are presented.

We now briefly discuss how our method [24] may be applied to DTV transmissions for use in 802.22 DSA environments, in particular we focus on the ATSC DTV standard used in the United States. The diagram of a smart receiver capable of decoding the authentication signal is given in Figure 1. We note that the smart receiver applies additional signal processing on the equalizer taps, allowing for detection of the fingerprint signal, while legacy receivers, i.e. regular ATSC televisions, have a similar signal processing chain sans the fingerprint signal processing. The additional signal processing required

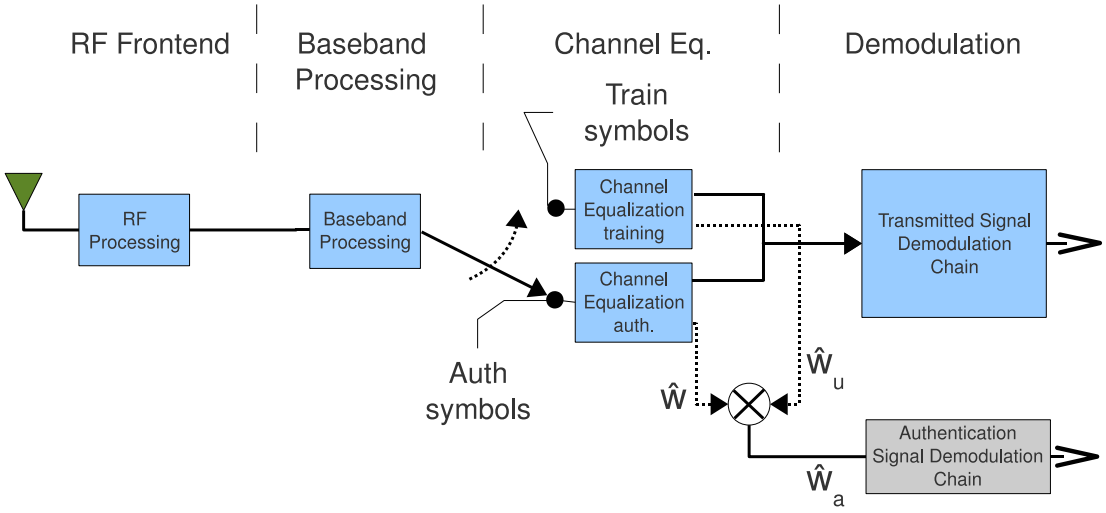


Figure 1: Example receiver system diagram with fingerprint decoder

by PHY-layer fingerprinting methods are one disadvantage to higher-layer authentication approaches.

The ATSC standard [6] specifies an eight level PAM signal with a 10.76 MHz symbol rate. The resulting signal is then filtered with a linear phase root raised cosine filter with rolloff factor $R = .1152$ creating an bandlimited signal with an effective bandwidth of 5.38 MHz. This filtering creates a single sideband signal with a vestigial sideband component still present, thus giving rise to the 8-VSB modulation specified by the ATSC standard.

We now present simulation results for the fingerprinting method proposed in [24]. In each run of the simulation, $80,000$ uncoded symbols of an 8-VSB ATSC signal were generated and filtered according to the VSB filter specifications of the ATSC standard, using 5 samples per symbol. The authentication fingerprint signal consisting of a symbol alphabet with two synthetic FIR channels was then applied at IF onto the transmitted signal at the rate of one authentication symbol for every $10,000$ symbols of the primary

signal, for an authentication signal period ratio of $1/10000$.

The augmented signal was then subjected to a AWGN channel. At the receiver, a decision directed LMS equalizer using a 21 tap transversal filter and step-size parameter $\mu = .002$ was used to reverse the ISI introduced by the fingerprinting signal. Realistic receiver timing recovery was simulated using a DLL timing recovery algorithm, however explicit timing offsets were not introduced in the simulations conducted. Synthetic channels using authentication signal power ratios of $.0015$, $.0025$, and $.004$ were used, using the impulse response alphabet depicted in Figure 2. The responses in Figure 2 are shown with an exaggerated authentication signal power ratio of $.15$ to show secondary impulse tap definition.

System BER results for the authentication signal and the primary signal were obtained via Monte Carlo simulation. A second, independent simulation was also conducted, producing BER results for a primary signal where the authentication fingerprint was not present. For each value of channel SNR, 50 sub-experiments were conducted and bit error results accumulated. Each sub-experiment consisted of repeated runs of the $80,000$ symbol Monte Carlo experiment described above, terminating when 100 authentication signal bit errors were accumulated or $8,000$ total authentication symbols were received, whichever came first. Bit error rate results are presented in Figure 3 for the authentication signal, the primary signal, and for the original primary signal with the fingerprint not present.

From these results we see that the authentication signal is received with a slight BER improvement over the primary VSB signal, for authentication signal power ratios greater than about $.002$. Additionally, we notice that the primary signal BER is negligibly

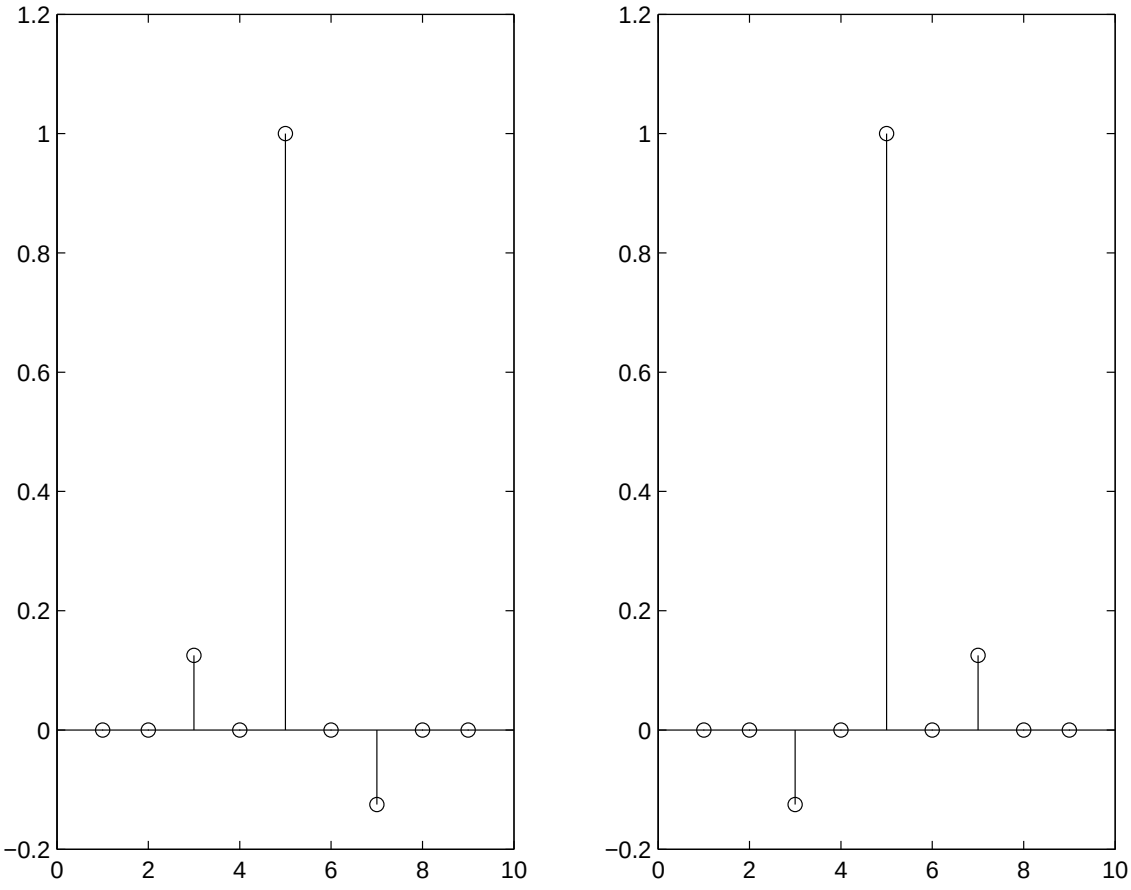


Figure 2: Synthetic channel impulse responses (exaggerated)

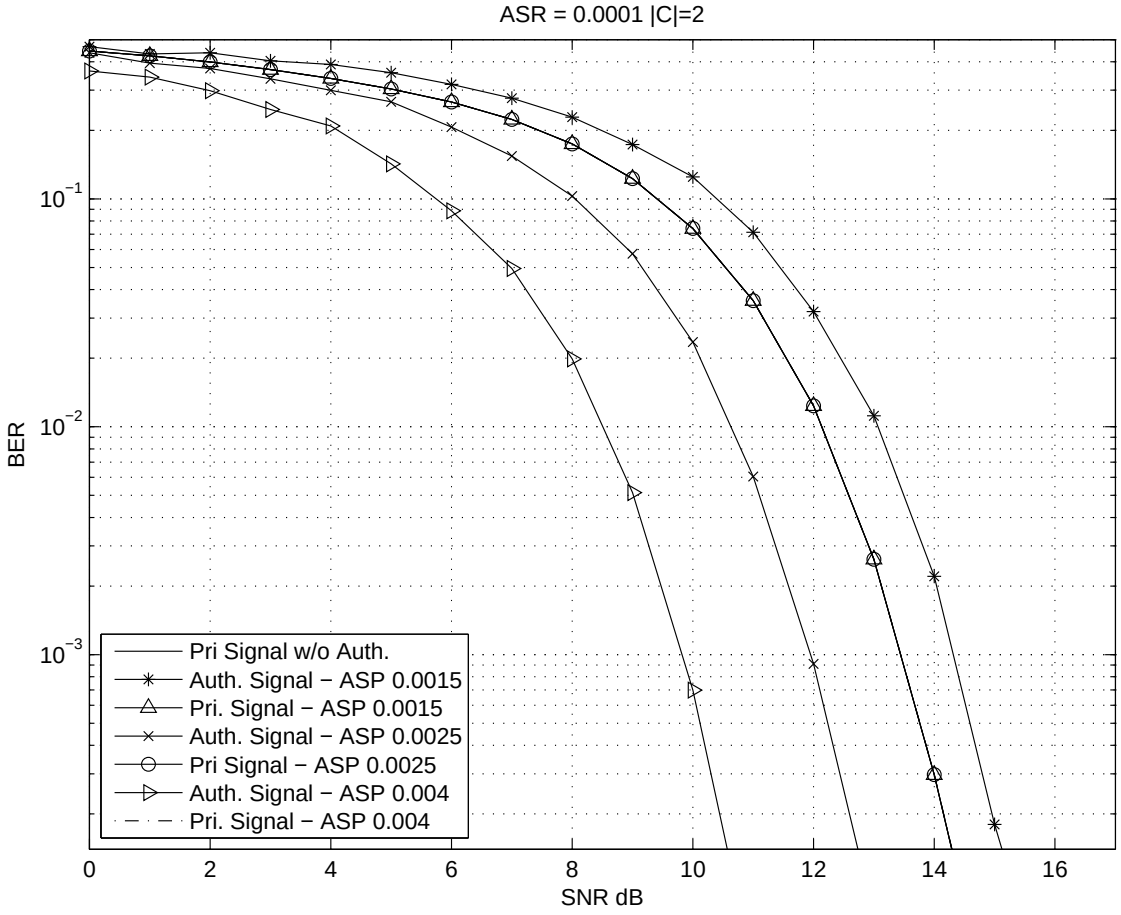


Figure 3: BER Authentication and ATSC Signals (overlapping series)

impacted by the presence of the authentication signal, since the BER curve for the primary signals with and without the authentication signal are indistinguishable, since these series overlap.

We observe that the impact to primary signal BER is negligible for all values of authentication power ratios used, as these series are overlapping in Figure 3. This result implies negligible impact to ATSC coverage area when the authentication signal is present. We note that a naive authentication signal detector was used in this simulation, which weighted every observation of the authentication signal over an entire baud equally. A more correct detector considering the learning curve of an LMS equalizer and employing maximum ratio combining of the observations would yield even better results in practice.

Decreasing the authentication signal period ratio of the authentication signal will improve its BER, at the expense of decreasing capacity. Since the fingerprint signal introduces additional model mismatch in the receiver's channel estimate, we note that decreasing the relative baud rate of the authentication channel has the effect of decreasing the frequency at which synthetic model mismatch error is introduced into the primary VSB signal. Thus the quality of the primary signal increases accordingly. Conversely, an increase in the authentication signal period ratio will result in decreased BER for the authentication signal, however this has the effect of increasing the MSE of the receiver's channel estimate. This increase in model mismatch error results in decreased capacity for the primary signal. Average channel model MSE results for the experiment described above are presented in Figure 4. We note a negligible impact to the the receiver's channel model estimate MSE in the presence of the authentication signal for the chosen system

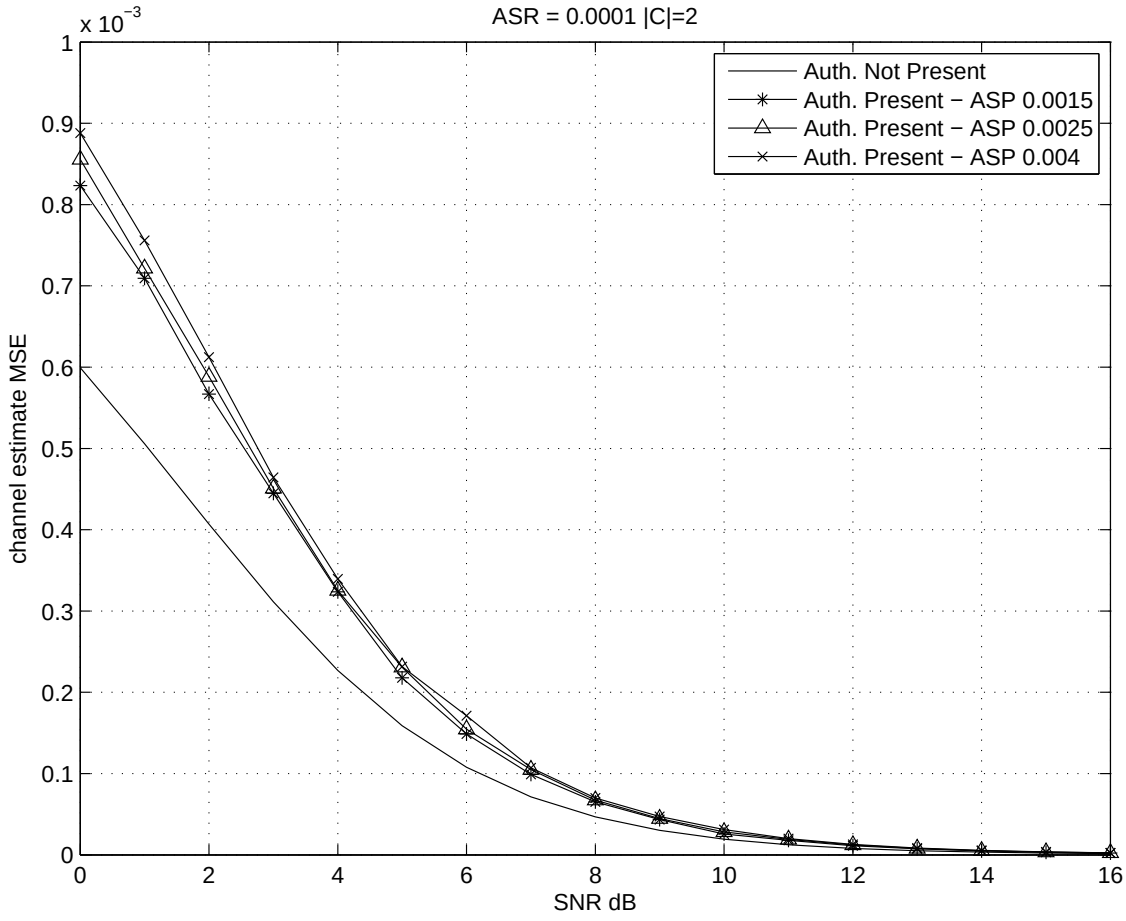


Figure 4: Channel Estimate MSE

parameters, and that increasing authentication signal power ratio has negligible impact to the channel estimate MSE.

In this motivating example, we have presented a simple channel-like fingerprinting scheme and discussed how the fingerprint signal can be used by receivers to authenticate the signal being transmitted. We have demonstrated that cryptographic authentication primitives may be employed in generating the content of the fingerprint, allowing for user authentication at the strength of the primitive. We have also discussed how the fingerprinting signal may be applied to DTV signals adhering to the ATSC standard, in support of the IEEE 802.22 dynamic spectrum access standard, and how a fingerprint authentication device prevents a number of attacks against cognitive radio signal classifiers. In addition to preventing primary user authentication attacks, the fingerprinting method presented may be applied to the primary signal at IF, enabling its use on legacy equipment without modification to the transmitter. In retrofitting existing DTV transmitters for use in 802.22 theaters, the fingerprinting device may be implemented as a preconditioning component in the IF chain of the transmitter, before signal up-conversion and power amplification.

1.6 Outline of Thesis

This dissertation is organized as follows. In Chapter 2 our work in [27] and [26] is presented. These works constitute a "Closed-Loop" approach to channel-like fingerprinting in the time and spacial dimensions, where CSI is not used in the design of the fingerprint. In Chapter 3 our work in [29] and [28] is presented. These works demonstrate "Open-Loop" approaches to channel-like fingerprinting, where partial CSI is available to

the transmitter when designing the fingerprint. In Chapter 4 our work in [32] is presented, where we leverage the PHY-layer fingerprinting approaches discussed in Chapters 1-3 to create a complete authentication system. In this chapter, example cryptographically secure messages are discussed and applied to Dynamic Spectrum Access scenarios. In Chapter 5 our work in [30] and [31] is presented, where channel stationary assumptions similar to the fingerprinting work in Chapters 1-3 are used to create a cooperative relaying system. Our concluding remarks and future work are presented in Chapter 6.

Chapter 2

Extrinsic Channel-Like Fingerprint Embedding for Authenticating MIMO Systems

2.1 Overview

A framework for introducing an extrinsic fingerprint signal to space-time coded transmissions at the physical layer is presented, where the fingerprint signal conveys a low capacity cryptographically secure authentication message of arbitrary length. The multi-bit digital fingerprint message conveyed by the fingerprint signal is available to all users within reception range and is used to authenticate the fingerprinted transmission. A novel approach is discussed where the fingerprint signaling mechanism mimics distortions similar to time-varying channel effects. Specifically, the fingerprint is detectable to receivers considering previous channel state information, but will be ignored by receivers equalizing according to current channel state information. Two example fingerprint signaling mechanisms and detection rules are presented based on pulse-amplitude keying and phase-shift keying approaches. The methods for obtaining the real (intrinsic) channel estimate, the extrinsic fingerprint message, and the primary transmission are analytically demonstrated using general pilot embedding schemes. The worst-case distortions caused by non-ideal equalization of a fingerprinted message are derived using the 2x2 Alamouti code. Simulation results including bit error rate (BER) and model mismatch error using

a maximum-likelihood (ML) receiver are presented for both the primary and fingerprint signal, while authentication signal BERs *lower* than the primary signal are demonstrated.

2.2 Introduction

In our previous work [24], the details of a cryptographically secure arbitrary-length digital signature by using an extrinsic channel-like fingerprint for narrowband single-input single-output (SISO) digital television signals was considered.

In this chapter we extend our work in [26], which considers one fingerprinting function for Space-Time Coded (STC) transmissions.

This chapter is organized as follows. Section 2.3 introduces the multiple-input multiple-output (MIMO) system model and presents a framework for embedding a channel-like fingerprint signal of an arbitrary length in bits. In Section 2.4 the extraction of the intrinsic channel state, the extrinsic fingerprint message, and the primary transmission are demonstrated. Section 2.5 presents two fingerprint signaling functions and accompanying detection rules, and the performance of these functions are derived. In 2.6 we present bit error rate simulations for the example fingerprinting functions, and in Section 2.7 we present our conclusions.

2.3 System Model and Problem Formulation

We assume the transmitter and receiver are MIMO systems with L_t transmit antennas and L_r receive antennas, with a STC transmitted at index t described by matrix $\mathbf{U}[t]$ of size $L_t \times M$. The STC $\mathbf{U}[t]$ transmitted across all L_t transmit antennas in M time slots

is a composite signal composed of both the original STC transmission data, which will be referred to as the primary signal, and pilot signals used for channel estimation. When the fingerprinting function $\mathbf{F}[t]$ is applied by the transmitter to the ST block $\mathbf{U}[t]$ before transmission, the block received at the receiver $\mathbf{Y}[t] \in \mathcal{C}^{L_r \times M}$ expressed in matrix form is

$$\mathbf{Y}[t] = \mathbf{H}[t]\mathbf{F}[t]\mathbf{U}[t] + \mathbf{N}[t], \tag{2.1}$$

where $\mathbf{H}[t] \in \mathcal{C}^{L_r \times L_t}$ is the channel coefficient matrix representing the intrinsic channel conditions experienced by the fingerprinted block at time t, and $\mathbf{F}[t] \in \mathcal{C}^{L_t \times L_t}$ is the fingerprinting function applied to the transmission. The channel noise $\mathbf{N}[t]$ is modeled as complex white Gaussian noise with zero mean and variance $(\sigma^2/2)\mathbf{I}_{(L_r \times M)}$. We assume the elements of $\mathbf{H}[t]$ to be independent Rayleigh fading and block-stationary, where $\mathbf{H}[t]$ remains constant over the block, or M symbols.

We now briefly describe the pilot-embedding framework presented in [50], which provides the edifice for the construction of $\mathbf{U}[t]$. We will demonstrate how our channel-like fingerprinting scheme conveys the fingerprint message through strategic manipulation of the pilot signals used for channel estimation, that are embedded in the transmission.

The transmission $\mathbf{U}[t]$ consists of a ST code data-bearer matrix $\mathbf{D}[t] \in \mathcal{C}^{L_t \times N}$ and data-projection matrix $\mathbf{A} \in \mathcal{R}^{N \times M}$. Here, N is the number of time slots reserved exclusively for data transmission, while time slots $M - N,\ N < M$ are reserved for data mixed with embedded pilot signals. The ST symbol $\mathbf{U}[t]$ with embedded pilots signals, becomes

$$\mathbf{U}[t] = \mathbf{D}[t]\mathbf{A} + \mathbf{P}, \tag{2.2}$$

where $\mathbf{P} \in \mathcal{R}^{L_t \times M}$ is the pilot matrix. The salient point of this data-bearing framework is that most pilot-embedding schemes can be generalized through the superposition of

the data-bearing structure $\mathbf{D}[t]\mathbf{A}$ and the pilot matrix $\mathbf{P}[t]$. The data-projection and pilot matrix satisfy the following properties:

$$\mathbf{AP}^H = \mathbf{0} \in \mathcal{R}^{N \times L_t}, \quad \mathbf{AA}^H = \mathbf{I} \in \mathcal{R}^{N \times N}, \quad \mathbf{PP}^H = \mathbf{I} \in \mathcal{R}^{L_t \times L_t}. \tag{2.3}$$

The properties (2.3) of the data-projection matrix $\mathbf{A}$ and $\mathbf{P}$ essentially allow $\mathbf{A}$ to project the data component $\mathbf{D}[t]$ onto the orthogonal subspace of the pilot matrix $\mathbf{P}$, allowing for signal demodulation by means of a Maximum Likelihood (ML) receiver. The expanded form of the signal at the receiver (2.1), with (2.2) becomes

$$\mathbf{Y}[t] = \mathbf{H}[t]\mathbf{X}[t]\mathbf{A} + \mathbf{H}[t]\mathbf{F}[t]\mathbf{P} + \mathbf{N}[t], \tag{2.4}$$

where $\mathbf{X}[t] = \mathbf{F}[t]\mathbf{D}[t]$ is the fingerprinted data transmission before projection by $\mathbf{A}$.

To the *unaware* receiver the distortions introduced by the fingerprinting function $\mathbf{F}[t]$ can be combined with the channel distortions $\mathbf{H}[t]$ and will be subsequently removed through equalization. This is because we consider the case where the fingerprinting function $\mathbf{F}[t]$ is applied to both the pilot and data signals of the transmission, consistent with the distortions introduced by the intrinsic channel response. A MMSE equalizer operating on current CSI will reverse both the intrinsic and extrinsic channel-like distortions using the block's pilot signals as reference. This process will be explained analytically in a moment.

The *aware* receiver must detect the fingerprinting signal in the presence of time-variant channel distortions. We consider the case where the intrinsic channel estimate $\mathbf{H}[t]$ is delineated from the extrinsic fingerprinting component $\mathbf{F}[t]$ through periodic omission of the fingerprint signal $\mathbf{F}[t]$, which will serve as the channel sounding mechanism

allowing for estimation of the intrinsic channel state only. Under this assumption, the coherence time of the channel will play an important role in the detection probability of $\mathbf{F}[t]$, since time-varying changes in $\mathbf{H}[t]$ will become noise when detecting $\mathbf{F}[t]$.

Since channel coherence over many blocks is a strong assumption for general time-variant channels [58], especially in high mobility scenarios when channel state is quickly changing, we consider here the most frequent channel sounding case where the fingerprint signal is omitted every even block and present on every odd block, yielding a fingerprint transmission with a 50 percent duty-cycle. With this design, channel coherence over only two blocks is sufficient for detecting our fingerprint message and a channel with less stationary behavior will result in degraded performance. Changing our time index to reflect this design, when $t = 2Mk$, the fingerprint is not present in the transmission and $\mathbf{F}[t]$ is replaced by the identity matrix, $\mathbf{I}$, for the channel sounding block. When $t = 2Mk - M$, $\mathbf{F}[t]$ is transmitted. Thus the received signal with the fingerprinting function applied to every other block transmission becomes

$$\mathbf{Y}[t] = \begin{cases} \mathbf{H}[t]\mathbf{U}[t] + \mathbf{N}[t], & t = 2Mk, \\ \mathbf{H}[t]\mathbf{F}[t]\mathbf{U}[t] + \mathbf{N}[t], & t = 2Mk - M. \end{cases} \tag{2.5}$$

While (2.5) considers a differential modulation where the perceived channel changes every block, in [13] *channel-tracking* equalizers were discussed. This work demonstrated that when equalizers that track channel state are employed, distortion to the primary-signal can be avoided by simply extending the symbol length of the fingerprinting function to be longer than the *forgetting period* of the equalizer. By increasing the length of the fingerprinting symbol, and thus decreasing the authentication symbol rate, (5) can be

extend to any scenario where the equalizer ignores previous channel state beyond some finite duration.

When the coherence time of the channel is large the correlation between $\mathbf{H}[2Mk]$ and $\mathbf{H}[2Mk-M]$ is significant, and the fingerprint function can be decoded correctly with a higher probability. Conversely, as the coherence time of the channel decreases, there is less mutual information between the current and outdated CSI and the performance of fingerprint decoder degrades. The correlation between time-varying channel estimates are discussed in [8], [34] and [60].

To ensure fair analysis of the fingerprinting system, the fingerprinting function is designed according to transmission energy constraint

$$||\mathbf{X}[t]||_F = ||\mathbf{D}[t]||_F = P_o, \tag{2.6}$$

where $||\cdot||_F$ represents the Frobenius norm. Therefore, according to (2.5) the fingerprinting function $\mathbf{F}[t]$ must be designed such that $||\mathbf{F}[t]||_F = \sqrt{L_t}$, maintaining an equi-energy transmission for the period when the fingerprint is present, i.e. during $\mathbf{Y}[2Mk - M]$, and when it is omitted, i.e. during $\mathbf{Y}[2Mk]$.

Extending the time-varying channel model used in [64] to MIMO transmissions, we consider a generalized time-variant channel response matrix for the intrinsic component of the channel $\mathbf{H}[t]$, where each scalar complex gain element $H_{i,j}[t]$ for rows $i = 0, \ldots, L_r - 1$ and columns $j = 0, \ldots, L_t - 1$ is the summation of three model components:

- A fixed time-invariant channel gain denoted $\bar{H}_{i,j} = E\left[H_{i,j}[t]\right]$
- A zero-mean time-variant channel gain component denoted $\mu_{i,j}[t]$

- A zero-mean receiver noise component denoted $N_{i,j}[t]$,

where $\bar{H}_{i,j}$ is the mean of the random variable $H_{i,j}[t]$. Thus, $H_{i,j}[t]$ becomes

$$H_{i,j}[t] = \bar{H}_{i,j} + \mu_{i,j}[t] + N_{i,j}[t]. \tag{2.7}$$

While in general each mean of the channel gains, $\bar{H}_{i,j}$, will be changing in time, we will assume that this component will remain stationary over the duration of the channel sounding symbol and adjacent fingerprinted symbol in (2.5). We obtain the following matrix definition for the time-varying channel

$$\begin{aligned}\mathbf{H}[t] &= \left(\bar{\mathbf{H}} + \boldsymbol{\mu}[t]\right) + \mathbf{N}[t] \\ &= \begin{bmatrix} \bar{H}_{0,0} + \mu_{0,0}[t] & \dots & \bar{H}_{0,L_t-1} + \mu_{0,L_t-1}[t] \\ \vdots & \ddots & \vdots \\ \bar{H}_{L_r-1,0} + \mu_{L_r-1,0}[t] & \dots & \bar{H}_{L_r-1,L_t-1} + \mu_{L_r-1,L_t-1}[t] \end{bmatrix} + \\ &\quad \begin{bmatrix} N_{0,0}[t] & \dots & N_{0,L_t-1}[t] \\ \vdots & \ddots & \vdots \\ N_{L_r-1,0}[t] & \dots & N_{L_r-1,L_t-1}[t] \end{bmatrix},\end{aligned} \tag{2.8}$$

where each element $N_{i,j}[t]$ is zero-mean complex Gaussian noise with variance σ_N^2 representing the normalized receiver noise projected on the pilot signals, assuming the projected noise is uniformly distributed over $\mathbf{P}^H$ (i.e. the pilot signals are optimally embedded into the transmission). We model the time-variant portion of the channel response gain for each element of $\boldsymbol{\mu}[t]$ corresponding as an independent first-order autoregressive (AR-1) model. The AR-1 model has been used to describe time-variant channels in previous works [66], and [36, 64]. Assuming an average AR-1 noise power σ_T^2 over all

time-variant gain elements $\mu_{i,j}[t]$, the AR-1 model is given as

$$\mu_{i,j}[t] = a\mu_{i,j}[t-1] + \sqrt{(1-a^2)}u_{i,j}[t]. \tag{2.9}$$

The AR model coefficient a in (2.9) represents the influence of the previous time-variant channel gain component $\mu_{i,j}[t-1]$ on the current estimate $\mu_{i,j}[t]$.The random component of the time-variant channel $\mu_{i,j}[t]$ is represented in (2.9) by $u_{i,j}[t] \sim \mathcal{CN}(0, \sigma_T^2)$, thus $E\left[\mu_{i,j}[t]\right] = 0, \forall i, j$. We consider the case where the AR model coefficient a, and the noise power σ_T^2 are the same for each independent channel i, j.

2.4 Fingerprint Analysis

Upon receiving the signal, the first step for both aware and unaware receivers is channel estimation. The channel estimation problem is to extract and estimate channel distortions in the received signal (2.5) for performing channel equalization and further recovering $\mathbf{D}[t]$. By post-multiplying both sides of (2.5) by $\mathbf{P}^H$ and using the properties in (2.3), the channel response $\mathbf{H}[t]$ can be estimated from the received signal during the channel-sounding symbol at $t = \tau_0 = 2Mk$

$$\begin{aligned} \mathbf{Y}[\tau_0]\mathbf{P}^H &= \left(\mathbf{H}[\tau_0](\mathbf{D}[\tau_0]\mathbf{A} + \mathbf{P}) + \mathbf{N}[\tau_0]\right)\mathbf{P}^H \\ &= \mathbf{H}[\tau_0] + \mathbf{N}[\tau_0]\mathbf{P}^H, \end{aligned} \tag{2.10}$$

where $\mathbf{N}[t]\mathbf{P}^H$, the channel estimate noise in (2.8), is the projection of the noise vector onto pilot signals and represents noise in the channel estimate.

Similarly the joint intrinsic and extrinsic channel distortions, $\mathbf{H}[2Mk]$ and $\mathbf{F}[t]$, can be estimated from the received signal (2.5) during the fingerprinted symbol at $\tau_1 =$

$2Mk - M$

$$\begin{aligned}\mathbf{Y}[\tau_1]\mathbf{P}^H &= \left(\mathbf{H}[\tau_1]\mathbf{F}[\tau_1](\mathbf{D}[\tau_1]\mathbf{A} + \mathbf{P}) + \mathbf{N}[\tau_1]\right)\mathbf{P}^H \\ &= (\mathbf{H}[\tau_1]\mathbf{F}[\tau_1]) + \mathbf{N}[\tau_1]\mathbf{P}^H,\end{aligned} \tag{2.11}$$

Combining results from (2.10) and (2.11), the channel estimate at the receiver, $\hat{\mathbf{H}}[t]$, becomes

$$\hat{\mathbf{H}}[t] = \begin{cases} \mathbf{H}[t] + \mathbf{N}[t]\mathbf{P}^H, & t = \tau_0 = 2Mk, \\ \mathbf{H}[t]\mathbf{F}[t] + \mathbf{N}[t]\mathbf{P}^H, & t = \tau_1 = 2Mk - M, \end{cases} \tag{2.12}$$

where $\mathbf{N}[t]\mathbf{P}^H$ is the normalized projected channel estimate noise. Since $\mathbf{N}[t]$ is uniformly distributed Gaussian noise, and since proper design of $\mathbf{P}$ should ensure that pilot symbols are placed such that channel conditions are uniformly estimated throughout the ST block, then $\mathbf{N}[t]\mathbf{P}^H$ should also have a uniform noise distribution.

2.4.1 Data Recovery

After the channel has been estimated via (2.10) and (2.11), the next step preformed by the receiver is the recovery of the transmitted data $\mathbf{D}[t]$. By post-multiplying both sides of (2.5) by $\mathbf{A}^H$ and using the properties (2.3), the data signal $\mathbf{D}[t]$ can be extracted from the received signal (2.5) during the channel-sounding symbol transmitted at $\tau_0 = 2Mk$, i.e.

$$\mathbf{Y}[\tau_0]\mathbf{A}^H = \mathbf{H}[\tau_0]\mathbf{D}[\tau_0] + \mathbf{N}[\tau_0]\mathbf{A}^H. \tag{2.13}$$

For the sake of exposition, we consider here the case where the number of transmit antenna and the number of receive antenna are equal, or $L_r = L_t$, and that $\hat{\mathbf{H}}[t]$ is invertible, which is the case considered later in simulation. Inversion for the case when $L_r \neq L_t$

is obtainable via a number of methods such as the pseudoinverse, however this topic is beyond the scope of this chapter. An estimate for the intrinsic channel response $\hat{\mathbf{H}}[\tau_0]$ is produced via (2.10), and thus the data signal can be recovered by pre-multiplying (2.13) by the inverse of the normalized channel estimate produced by the MMSE estimator, or $\hat{\mathbf{H}}^{-1}[\tau_0]$. When the channel is perfectly estimated for either the τ_0 or τ_1 block, i.e.

$$\hat{\mathbf{H}}^{-1}[t] = \mathbf{H}^{-1}[t],\ t = \tau_0 \text{ or } \tau_1 \tag{2.14}$$

the extracted data signal at $t = \tau_0 = 2Mk$ is

$$\begin{aligned} \hat{\mathbf{D}}[\tau_0] =& \hat{\mathbf{H}}^{-1}[\tau_0]\mathbf{Y}[\tau_0]\mathbf{A}^H \\ =& \mathbf{D}[\tau_0] + \hat{\mathbf{H}}^{-1}[\tau_0]\mathbf{N}[\tau_0]\mathbf{A}^H. \end{aligned} \tag{2.15}$$

Similarly, when post-multiplying by $\mathbf{A}^H$ for $t = \tau_1 = 2Mk - M$

$$\mathbf{Y}[\tau_1]\mathbf{A}^H = \mathbf{H}[\tau_1]\mathbf{F}[\tau_1]\mathbf{D}[\tau_1] + \mathbf{N}[\tau_1]\mathbf{A}^H, \tag{2.16}$$

an estimate for the intrinsic channel response combined with the extrinsic response $\hat{\mathbf{H}}[\tau_0]\mathbf{F}[\tau_1]$ is produced via (2.11) and the data signal can be recovered by pre-multiplying (2.16) by $\left(\hat{\mathbf{H}}[\tau_1]\mathbf{F}[\tau_1]\right)^{-1}$. For the perfectly estimated channel (2.14) the extracted data signal at $t = \tau_1 = 2Mk - M$ becomes

$$\hat{\mathbf{D}}[\tau_1] = \mathbf{D}[\tau_1] + \left(\hat{\mathbf{H}}[\tau_1]\mathbf{F}[\tau_1]\right)^{-1}\mathbf{N}[\tau_1]\mathbf{A}^H. \tag{2.17}$$

We note that from (2.11) and (2.17) it has been shown that the data signal $\mathbf{D}[\tau_1]$ can be recovered from $\mathbf{Y}[\tau_1]$ in the presence of the fingerprinting distortion $\mathbf{F}[\tau_1]$ without explicitly extracting and detecting the fingerprinting function $\mathbf{F}[\tau_1]$. Thus the primary transmission in the proposed fingerprinting system can be recovered independently from the fingerprint detection by both the aware and unaware receivers.

A further advantage to the proposed system is that channel estimates obtained during (2.10) and (2.11), and subsequent channel equalization steps preformed in (2.15) and (2.17) are identical steps taken by an unmodified/unaware receiver. Thus, we have demonstrated that the fingerprinted signal is received by unaware receivers without modification to the receiver, channel estimation procedure, or equalization device when generalized pilot embedding and channel estimation are employed.

2.4.2 Fingerprint Detection

We now consider two methods for detecting the fingerprint signal given the sequence of channel state information in (2.12).

The first detection rule, also considered in [64], is the differential channel estimate denoted $\mathbf{Z}_{SUB}[\tau_1, \tau_0]$. This detection rule is useful for detecting amplitude differences between the even and odd block transmissions in (2.5), i.e. our differential fingerprint signaling method, and is obtained by subtracting the sounding symbol estimate from the fingerprinted symbol estimate. Under the assumption that the fingerprinting function is transmitted independently from the channel response, their difference becomes

$$\begin{aligned} &E\left[\mathbf{Z}_{SUB}[\tau_1,\tau_0]\right] = E\left[\mathbf{Y}[\tau_1]\mathbf{P}^H - \mathbf{Y}[\tau_0]\mathbf{P}^H\right] \\ =&E\left[\bar{\mathbf{H}}\mathbf{F}[\tau_1]\right] + E\left[\boldsymbol{\mu}[\tau_1]\mathbf{F}[\tau_1]\right] + E\left[\mathbf{N}[\tau_1]\right] - E\left[\bar{\mathbf{H}}\right] - E\left[\boldsymbol{\mu}[\tau_0]\right] - E\left[\mathbf{N}[\tau_0]\right] \\ =&\bar{\mathbf{H}}\mathbf{F}[\tau_1] - \bar{\mathbf{H}}. \end{aligned} \tag{2.18}$$

From (2.18) we note that this detector is unbiased, since only the means $\bar{\mathbf{H}}$ and $\mathbf{F}[t]$ are present.

We also consider the Hadamard product, or element-wise product between two ma-

trices, for detecting fingerprinting functions perturbing signal phase. Denoted $\mathbf{Z}_{HAD}[\tau_1, \tau_0]$, this detection rule is the element-wise product between the channel sounding estimate and the conjugate of the fingerprinted channel estimate, and is given as

$$\begin{aligned}
&E\left[\mathbf{Z}_{HAD}[\tau_1,\tau_0]\right] = E\left[\left(\mathbf{Y}[\tau_1]\mathbf{P}^H\right)\circ\left(\mathbf{Y}[\tau_0]\mathbf{P}^H\right)^*\right] \\
=&E\left[\left(\bar{\mathbf{H}}\mathbf{F}[\tau_1]\right)\circ\bar{\mathbf{H}}^*\right] + E\left[\left(\bar{\mathbf{H}}\mathbf{F}[\tau_1]\right)\circ\boldsymbol{\mu}^*[\tau_0]\right] + E\left[\mathbf{N}[\tau_1]\circ\boldsymbol{\mu}^*[\tau_0]\right] + \\
&E\left[(\boldsymbol{\mu}[\tau_1]\mathbf{F}[\tau_1])\circ\mathbf{N}^*[\tau_0]\right] + E\left[(\boldsymbol{\mu}[\tau_1]\mathbf{F}[\tau_1])\circ\bar{\mathbf{H}}^*\right] + E\left[\mathbf{N}[\tau_1]\circ\bar{\mathbf{H}}^*\right] + \\
&E\left[(\boldsymbol{\mu}[\tau_1]\mathbf{F}[\tau_1])\circ\boldsymbol{\mu}^*[\tau_0]\right] + E\left[\mathbf{N}[\tau_1]\circ\mathbf{N}^*[\tau_0]\right] + E\left[\left(\bar{\mathbf{H}}\mathbf{F}[\tau_1]\right)\circ\mathbf{N}^*[\tau_0]\right] \\
=&||\bar{\mathbf{H}}||^2\mathbf{F}[\tau_1],
\end{aligned} \tag{2.19}$$

where $(\circ)$ represents the Hadamard product and $(*)$ represents conjugation. Here the perturbation factor can be extracted from the argument of the product of the individual scalar estimates. We will use these two detectors in the following fingerprint examples and demonstrate their performance.

2.5 Some Fingerprinting Scenarios

We now consider some simple fingerprinting functions as candidates for $\mathbf{F}[t]$. We will give examples for each fingerprinting function using the 2x2 Alamouti code [5] according to the polar representation of the complex valued intrinsic channel model (2.8), i.e.

$$\begin{aligned}
\mathbf{H}[t] &= \begin{bmatrix} \bar{H}_{0,0}+\mu_{0,0}[t] & \bar{H}_{0,1}+\mu_{0,1}[t] \\ \bar{H}_{1,0}+\mu_{1,0}[t] & \bar{H}_{1,1}+\mu_{1,1}[t] \end{bmatrix} + \begin{bmatrix} N_{0,0}[t] & N_{0,1}[t] \\ N_{1,0}[t] & N_{1,1}[t] \end{bmatrix} \\
&= \begin{bmatrix} \alpha_1 e^{j\theta_1} & \alpha_3 e^{j\theta_3} \\ \alpha_2 e^{j\theta_2} & \alpha_4 e^{j\theta_4} \end{bmatrix} + \begin{bmatrix} \mu_1[t] & \mu_3[t] \\ \mu_2[t] & \mu_4[t] \end{bmatrix} + \begin{bmatrix} N_1[t] & N_3[t] \\ N_2[t] & N_4[t] \end{bmatrix},
\end{aligned} \tag{2.20}$$

where the indices $\{i, j\}$ are serialized to $1, 2, \ldots, L_t M$ first column-wise and then row-wise, for simplicity of notation. Here $\bar{H}_{i,j}$ is represented in polar form, with amplitude $\alpha_x,\ x = 1, \ldots, ML_t$ and angle $\theta_x,\ x = 1, \ldots, ML_t$. In the case of the 2x2 code, $N = L_t = 2$.

2.5.1 Antenna Amplitude Modulation (AAM)

The first fingerprinting function we consider introduces a gain offset of ϵ between symbols to be transmitted by each antenna such that the overall transmission energy constraint is withheld. This function can also be thought of as a modulation of the gain of each antenna, and will be designated with the subscript AAM. The antenna gain fingerprinting function for the 2x2 code (i.e. $L_t = 2,\ M = 3,\ N = 2$) can be represented as

$$\mathbf{F}_{AAM}[t] = \gamma \begin{bmatrix} 1-\epsilon & 0 \\ 0 & 1+\epsilon \end{bmatrix},\ |\epsilon| < 1, \tag{2.21}$$

where γ is a normalization constant used to maintain the constant energy constraint as in (2.6). For the AAM fingerprinting function this normalization constant becomes

$$\gamma = \frac{1}{\sqrt{1+\epsilon^2}}. \tag{2.22}$$

Since the AAM fingerprinting function perturbs the amplitude of transmitted symbols, we apply the differential channel test statistic (2.18) to detect amplitude distortions between channel estimates. Using (2.18) and (2.20) applied to the AAM fingerprint func-

tion (2.21), test statistic for the 2x2 Alamouti code, denoted $\mathbf{Z}_{AAM}[\tau_1, \tau_0]$, becomes

$$\begin{aligned}\mathbf{Z}_{AAM}[\tau_1, \tau_0] &= E\left[\mathbf{Z}_{SUB}[\tau_1, \tau_0]\right] = \bar{\mathbf{H}}\mathbf{F}[\tau_1] - \bar{\mathbf{H}} \\ &= \begin{bmatrix} \alpha_1(1-\epsilon)e^{j\theta_1} & \alpha_3(1+\epsilon)e^{j\theta_3} \\ \alpha_2(1-\epsilon)e^{j\theta_2} & \alpha_4(1+\epsilon)e^{j\theta_4} \end{bmatrix} - \begin{bmatrix} \alpha_1 e^{j\theta_1} & \alpha_3 e^{j\theta_3} \\ \alpha_2 e^{j\theta_2} & \alpha_4 e^{j\theta_4} \end{bmatrix} = \begin{bmatrix} -\epsilon\alpha_1 e^{j\theta_1} & \epsilon\alpha_3 e^{j\theta_3} \\ -\epsilon\alpha_2 e^{j\theta_2} & \epsilon\alpha_4 e^{j\theta_4} \end{bmatrix}.\end{aligned} \tag{2.23}$$

The estimates received in each time slot $i = 0, \ldots, N-1$ for each antenna $j = 0, 1$ in (2.23), $Z_{AAM_{i,j}}$, can be combined by subtracting the amplitude of the estimates corresponding to the signals received by each antenna, i.e. the columns of (2.23). The ensemble estimate for ϵ becomes,

$$\begin{aligned}\hat{\epsilon} &= \sum_{i=0}^{N-1} Re\{Z_{AAM_{i,1}}[\tau_1, \tau_0]\} - \sum_{i=0}^{N-1} Re\{Z_{AAM_{i,0}}[\tau_1, \tau_0]\} \\ &= (\epsilon\alpha_3 + \epsilon\alpha_4) - (-\epsilon\alpha_1 - \epsilon\alpha_2) = \epsilon\lambda, \quad \lambda = \sum_{i=1}^{L_t N} \alpha_i,\end{aligned} \tag{2.24}$$

is the total channel gain measured during the sounding symbol transmitted at $t = \tau_0$.

From (2.24) we see that the performance of the test signal $\mathbf{Z}_{AAM}[\tau_1, \tau_0]$ depends on the aggregate signal gain of the channel λ and the value chosen for the perturbation amplitude ϵ. Therefore when using the AAM fingerprinting function we conclude that the symbol error rate (SER) for the authentication signal, and thus the detection performance of the fingerprint for the aware receiver, can be improved by increasing ϵ at the transmitter.

To analyze the performance of this fingerprinting function, we must also compute

the variance of the test statistic. This computation, similar to the proof in [64], becomes

$$\begin{aligned}
&\mathrm{Var}\left[\mathbf{Z}_{SUB}[\tau_1,\tau_0]\right] \\
&= \mathrm{Var}\left[\left(\mathbf{H}[\tau_1]\mathbf{F}[\tau_1]+\mathbf{N}[\tau_1]\mathbf{P}^H\right)-\left(\mathbf{H}[\tau_0]+\mathbf{N}[\tau_0]\mathbf{P}^H\right)\right] \\
&= \mathrm{Var}\left[\boldsymbol{\mu}[\tau_1]\mathbf{F}[\tau_1]\right]+\mathrm{Var}\left[\boldsymbol{\mu}[\tau_0]\right] \\
&- 2\mathrm{Cov}\left[\boldsymbol{\mu}[\tau_1]\mathbf{F}[\tau_1],\boldsymbol{\mu}[\tau_0]\right]+\mathrm{Var}\left[\mathbf{N}[\tau_1]\mathbf{P}^H\right]+\mathrm{Var}\left[\mathbf{N}[\tau_0]\mathbf{P}^H\right].
\end{aligned} \tag{2.25}$$

Due to the design of the AAM fingerprinting function, the gain of the j^{th} column of (2.23) is either increased or decreased by the perturbation factor ϵ, thus (2.25) becomes

$$\mathrm{Var}\left[\mathbf{Z}_{SUB}[\tau_1,\tau_0]\right]=\begin{cases}\sigma_T^2(1+(1-2a)(1-\epsilon)^2)+\sigma_N^2, & j=0, i=0,\ldots,N-1\\ \\ \sigma_T^2(1+(1-2a)(1+\epsilon)^2)+\sigma_N^2, & j=1, i=0,\ldots,N-1.\end{cases} \tag{2.26}$$

Therefore, the total variance of the estimate (2.24) for the 2x2 code becomes

$$\sigma_\epsilon^2=\mathrm{Var}\left[\hat{\epsilon}\right]=\frac{\sigma_T^2(2(1-a)+\epsilon^2(1-2a))+\sigma_N^2}{L_tN}. \tag{2.27}$$

If we select a typical antipodal binary signal constellation for the AAM fingerprint function $\mathbf{F}$ with parameter ϵ, i.e.

$$\mathbf{F}_{AAM}[t]\in\left\{\gamma\begin{bmatrix}1-\epsilon & 0\\ 0 & 1+\epsilon\end{bmatrix},\ \gamma\begin{bmatrix}1+\epsilon & 0\\ 0 & 1-\epsilon\end{bmatrix}\right\}, \tag{2.28}$$

it can be shown that the symbol error rate for the maximum-likelihood fingerprint detector detecting the transmitted fingerprint function $\mathbf{F}$ from the noisy estimate at the receiver $\hat{\mathbf{F}}$ is

$$P\left[\hat{\mathbf{F}}\neq\mathbf{F}\right]=Q\left(\sqrt{\frac{2\epsilon^2\lambda^2}{\sigma_\epsilon}}\right), \tag{2.29}$$

where $Q(\cdot)$ is the Gaussian tail function. We note that the variance (2.27) decreases linearly as the number of elements in the code increases, i.e. as L_t or N increase, however we also note that the variance also increases quadratically in ϵ.

2.5.2 AAM Fingerprint Distortion

We now consider the distortions experienced by the Maximum Ratio Combining (MRC) decoder operating on the 2x2 Alamouti code when equalizing the AAM-fingerprinted signal $\mathbf{Y}[\tau_1]$ according to an incorrect channel estimate that considers only the intrinsic channel estimate, i.e. if $\mathbf{H}[\tau_0]$ were used as the channel estimate for a symbol transmitted at $t = \tau_1 = 2k$ instead of $\mathbf{H}[\tau_1]\mathbf{F}[\tau_1]$. This important result delineates the worst-case degradation in performance the MRC receiver would experience due to channel model estimate mismatch, which generally destroys the orthogonality of the signals in the transmitted space-time code $\mathbf{D}[t]$. These distortions might be applicable to unaware receivers with non-adaptive equalization, and demonstrates how the perturbation parameter ϵ must be carefully chosen to limit maximum signal degradation when considering a heterogeneous system of receivers. For the 2x2 Alamouti code,

$$\mathbf{D}[t] = \begin{bmatrix} d_1 & -d_2^* \\ d_2 & d_1^* \end{bmatrix}, \tag{2.30}$$

the transmitted symbol $\mathbf{X}[t]$ with fingerprinting function (2.21) becomes

$$\mathbf{X}_{AAM}[t] = \begin{bmatrix} 1-\epsilon & 0 \\ 0 & 1+\epsilon \end{bmatrix} \begin{bmatrix} d_1 & -d_2^* \\ d_2 & d_1^* \end{bmatrix} = \begin{bmatrix} d_1(1-\epsilon) & -d_2^*(1-\epsilon) \\ d_2(1+\epsilon) & d_1^*(1+\epsilon) \end{bmatrix}. \tag{2.31}$$

The data signal estimate using MRC on the extracted data (2.15), using (2.20) becomes

$$\mathbf{Y}_{AAM}[t] = \begin{bmatrix} r_1 & r_3 \\ r_2 & r_4 \end{bmatrix} + \mathbf{N}[t]\mathbf{P}^H, \tag{2.32}$$

where

$$\begin{aligned} r_1 = \ & \alpha_1 d_1(1-\epsilon)e^{j\theta_1} + \alpha_3 d_2(1+\epsilon)e^{j\theta_3} \quad r_3 = \ -\alpha_1 d_2^*(1-\epsilon)e^{j\theta_1} + \alpha_3 d_1^*(1+\epsilon)e^{j\theta_3} \\ r_2 = \ & \alpha_2 d_1(1-\epsilon)e^{j\theta_2} + \alpha_4 d_2(1+\epsilon)e^{j\theta_4} \quad r_4 = \ -\alpha_2 d_2^*(1-\epsilon)e^{j\theta_2} + \alpha_4 d_1^*(1+\epsilon)e^{j\theta_4}. \end{aligned} \tag{2.33}$$

The estimates of the received signal using an MRC receiver with model mismatch distortion from the fingerprinting function present are given as

$$\begin{aligned} \tilde{d}_{1_{AAM}} &= \hat{\alpha}_1 e^{-j\hat{\theta}_1}\left(\alpha_1 d_1(1-\epsilon)e^{j\theta_1} + \alpha_3 d_2(1+\epsilon)e^{j\theta_3}\right) \\ &\quad + \hat{\alpha}_2 e^{-j\hat{\theta}_2}\left(\alpha_2 d_1(1-\epsilon)e^{j\theta_2} + \alpha_4 d_2(1+\epsilon)e^{j\theta_4}\right) \\ &\quad + \hat{\alpha}_3 e^{j\hat{\theta}_3}\left(-\alpha_1 d_2(1-\epsilon)e^{-j\theta_1} + \alpha_3 d_1(1+\epsilon)e^{-j\theta_3}\right) \\ &\quad + \hat{\alpha}_4 e^{j\hat{\theta}_4}\left(-\alpha_2 d_2(1-\epsilon)e^{-j\theta_2} + \alpha_4 d_1(1+\epsilon)e^{-j\theta_4}\right) + \eta_1, \\ \tilde{d}_{2_{AAM}} &= \hat{\alpha}_3 e^{-j\hat{\theta}_3}\left(\alpha_1 d_1(1-\epsilon)e^{j\theta_1} + \alpha_3 d_2(1+\epsilon)e^{j\theta_3}\right) \\ &\quad + \hat{\alpha}_4 e^{-j\hat{\theta}_4}\left(\alpha_2 d_1(1-\epsilon)e^{j\theta_2} + \alpha_4 d_2(1+\epsilon)e^{j\theta_4}\right) \\ &\quad - \hat{\alpha}_1 e^{j\hat{\theta}_1}\left(-\alpha_1 d_2(1-\epsilon)e^{-j\theta_1} + \alpha_3 d_1(1+\epsilon)e^{-j\theta_3}\right) \\ &\quad - \hat{\alpha}_2 e^{j\hat{\theta}_2}\left(-\alpha_2 d_2(1-\epsilon)e^{-j\theta_2} + \alpha_4 d_1(1+\epsilon)e^{-j\theta_4}\right) + \eta_2. \end{aligned} \tag{2.34}$$

where

$$\begin{aligned} \eta_1 &= \alpha_1 e^{-j\theta_1} N_1 + \alpha_2 e^{j\theta_2} N_2^* + \alpha_3 e^{-j\theta_3} N_3 + \alpha_4 e^{j\theta_1} N_4^*, \\ \eta_2 &= -\alpha_1 e^{j\theta_1} N_2^* + \alpha_2 e^{-j\theta_2} N_1 - \alpha_3 e^{j\theta_3} N_4^* + \alpha_4 e^{-j\theta_1} N_3, \end{aligned} \tag{2.35}$$

represent the the combined receiver noise in the estimates of d_1 and d_2, respectively. In (2.34), $\hat{\alpha}_1, \hat{\alpha}_2, \hat{\alpha}_3, \hat{\alpha}_4$ are the complex channel gain estimates for the intrinsic channel

given by (2.20), produced by the receiver during $\mathbf{H}[2Mk]$, $\hat{\theta}_1, \hat{\theta}_2, \hat{\theta}_3, \hat{\theta}_4$ are the channel phase estimates, and N_1, N_2, N_3, N_4 are the elements of $\mathbf{N}[t]\mathbf{P}^H$. We consider the case where the intrinsic channel component is perfectly coherent over the channel sounding symbol and the fingerprinted symbol, thus the time-variant component $\boldsymbol{\mu}[t]$ of (2.20) is omitted, i.e. $\boldsymbol{\mu}[\tau_1] = \boldsymbol{\mu}[\tau_0] = \mathbf{0}$, thus in the noiseless case, when $\mathbf{N}[\tau_1] = \mathbf{N}[\tau_0] = \mathbf{0}$, the channel estimates for $t = \tau_1$ and $t = \tau_0$ are equal

$$\hat{\mathbf{H}}[\tau_1] = \hat{\mathbf{H}}[\tau_0], \tag{2.36}$$

thus the estimates for channel amplitude and phase have perfectly determined the intrinsic channel response, or

$$\hat{\mathbf{H}}[t] = \mathbf{H}[t] = \begin{bmatrix} \hat{\alpha}_1 e^{j\hat{\theta}_1} & \hat{\alpha}_3 e^{j\hat{\theta}_3} \\ \alpha_2 e^{j\hat{\theta}_2} & \hat{\alpha}_4 e^{j\hat{\theta}_4} \end{bmatrix} = \begin{bmatrix} \alpha_1 e^{j\theta_1} & \alpha_3 e^{j\theta_3} \\ \alpha_2 e^{j\theta_2} & \alpha_4 e^{j\theta_4} \end{bmatrix}, \tag{2.37}$$

leaving only distortions due to the presence of the extrinsic fingerprint. Using (2.37), after some manipulation, (2.34) becomes

$$\begin{aligned} \tilde{d}_{1_{AAM}} &= (\lambda - \epsilon(\alpha_1 + \alpha_2 - \alpha_3 - \alpha_4))d_1 + 2\epsilon(\alpha_1\alpha_3 e^{j(\theta_3 - \theta_1)} + \alpha_2\alpha_4 e^{j(\theta_4 - \theta_2)})d_2 + \eta_1, \\ \tilde{d}_{2_{AAM}} &= (\lambda - \epsilon(\alpha_1 + \alpha_2 - \alpha_3 - \alpha_4))d_2 - 2\epsilon(\alpha_1\alpha_3 e^{j(\theta_1 - \theta_3)} + \alpha_2\alpha_4 e^{j(\theta_2 - \theta_4)})d_1 + \eta_2, \end{aligned} \tag{2.38}$$

From (2.38) we notice that an AAM-fingerprinted Alamouti code improperly equalized according to outdated CSI (i.e. CSI that does not reflect the distortions introduced by the fingerprinting function), is degraded in amplitude by an amount proportional to ϵ. Specifically, the estimate for d_1 is degraded in amplitude by $\epsilon(-\alpha_1 - \alpha_2)$ and a cross signal is introduced from the d_2 signal proportional to 2ϵ. Similar distortions are experienced for the d_2 symbol, which is also degraded by $\epsilon(-\alpha_1 - \alpha_2)$. These are the worse-case

distortions incurred due to channel model estimate miss-match when using (2.21) as a fingerprinting function, and demonstrates the importance of selecting ϵ when considering receivers with lower performance equalizers. For example, when the equalizer used by a receiver has a particularly slow learning curve, or if there is delay between when the channel is estimated and when this estimate can be used for equalization, the receiver can equalize the channel according to an outdated channel model and thus model mismatch distortions will occur.

We note that when the data symbol $\mathbf{D}[t]$ is equalized and decoded according to current CSI, (2.10) and (2.11), the intrinsic and extrinsic channel distortions will be corrected when decoding the symbols $\mathbf{D}[\tau_0]$ and $\mathbf{D}[\tau_1]$. Thus, when channel model mismatch error during the fingerprinted symbol (2.14) is omitted, $\mathbf{D}[t]$ will be recovered using the MMSE channel estimate according to (2.16) and (2.17). The primary signal estimates for the 2x2 Alamouti code when the antenna amplitude offset is properly corrected by equalization becomes

$$\tilde{d}_1 = \lambda^{(2)} d_1 + \eta_1, \quad \tilde{d}_2 = \lambda^{(2)} d_2 + \eta_2, \quad \lambda^{(2)} = \sum_{i=1}^{L_t N} \alpha_i^2, \tag{2.39}$$

which is the anticipated performance for the 2x2 MRC Alamouti decoder with the perfect channel estimation assumption.

2.5.3 Antenna Phase Modulation (APM)

We now consider a fingerprinting function that introduces a phase offset between the signals to be transmitted by each antenna, denoted with the subscript APM. The

fingerprinting function for the 2x2 code can be written

$$\mathbf{F}_{APM}[t] = \begin{bmatrix} e^{-j\epsilon} & 0 \\ 0 & e^{j\epsilon} \end{bmatrix}, \ 0 \leq \epsilon < 2\pi. \tag{2.40}$$

Since the APM fingerprinting function introduces a phase perturbation, we apply the Hadamard product detector (2.19). The APM fingerprinting function in (2.40) and equation (2.20) for the 2x2 code becomes

$$\begin{aligned} \mathbf{Z}_{APM}[\tau_1, \tau_0] &= E\left[\mathbf{Z}_{HAD}[\tau_1, \tau_0]\right] \\ &= E\left[\begin{bmatrix} \alpha_1 e^{j(\theta_1-\epsilon)} & \alpha_3 e^{j(\theta_3+\epsilon)} \\ \alpha_2 e^{j(\theta_2-\epsilon)} & \alpha_4 e^{j(\theta_4+\epsilon)} \end{bmatrix} \circ \begin{bmatrix} \alpha_1 e^{-j\theta_1} & \alpha_3 e^{-j\theta_3} \\ \alpha_2 e^{-j\theta_2} & \alpha_4 e^{-j\theta_4} \end{bmatrix}\right] = \begin{bmatrix} \alpha_1^2 e^{-j\epsilon} & \alpha_3^2 e^{j\epsilon} \\ \alpha_2^2 e^{-j\epsilon} & \alpha_4^2 e^{j\epsilon} \end{bmatrix}. \end{aligned} \tag{2.41}$$

Combining all scalar estimates from (2.41) by averaging the scalar estimates corresponding to the signals received by each antenna and taking the conjugate of the estimates from the second column, the ensemble estimate for ϵ becomes,

$$\begin{aligned} e^{-j\hat{\epsilon}} &= \sum_{j=0}^{N} Z_{APM_{1,j}}[\tau_1, \tau_0] + \sum_{j=0}^{N} Z^*_{APM_{0,j}}[\tau_1, \tau_0] \\ &= \lambda^{(2)} e^{-j\epsilon}, \end{aligned} \tag{2.42}$$

where the disturbance factor ϵ can be recovered by taking the argument of (2.42), and $\lambda^{(2)} = \sum_{x=1}^{L_t N} \alpha_x^2$ is the anticipated signal gain for the 2x2 MRC Alamouti decoder with the perfect channel estimation assumption.

From (2.42) we see that the performance of the test signal $\mathbf{Z}_{APM}[\tau_1, \tau_0]$ depends on the aggregate signal gain of the channel $\lambda^{(2)}$ and the magnitude of the perturbation factor, ϵ. Therefore when using the APM fingerprinting function we conclude that the authentication signal SER can be decreased by increasing ϵ at the transmitter.

The variance of the detection rule (2.42) can be written,

$$\mathrm{Var}\left[\mathbf{Z}_{HAD}[\tau_1,\tau_0]\right] = \left(\sigma_N^2+\sigma_T^2\right)^2\mathbf{1} + 2\left(\sigma_N^2+\sigma_T^2+a\sigma_T^2+a\sigma_T^4\right)\bar{\mathbf{H}}^{(2)}, \tag{2.43}$$

where $\mathbf{H}^{(2)} = \mathbf{H}\circ\mathbf{H}^*$ represents the element-wise square operation on the matrix $\mathbf{H}$ and it's conjugate. Therefore, the total variance of the estimate (2.42) for the case where all elements of $\bar{\mathbf{H}}^{(2)}$ are equal, becomes

$$\sigma_\epsilon^2 = \frac{\mathrm{Var}\left[\mathbf{Z}_{HAD}[\tau_1,\tau_0]\right]}{NL_t}. \tag{2.44}$$

If we select an antipodal signal constellation for (2.40) with phase perturbation parameter $\epsilon = \pi/2$, i.e.

$$\mathbf{F}[t] \in \left\{ \begin{bmatrix} e^{-j\pi/2} & 0 \\ 0 & e^{j\pi/2} \end{bmatrix}, \begin{bmatrix} e^{j\pi/2} & 0 \\ 0 & e^{-j\pi/2} \end{bmatrix} \right\}, \tag{2.45}$$

it can be shown that the symbol error rate for the maximum-likelihood fingerprint detector, detecting $\mathbf{F}$ from the received estimate $\hat{\mathbf{F}}$, is

$$P\left[\hat{\mathbf{F}} \neq \mathbf{F}\right] = Q\left(\lambda^{(2)}\sqrt{\frac{2}{\sigma_\epsilon}}\sin\left(\frac{\pi}{2}\right)\right), \tag{2.46}$$

where $Q(\cdot)$ is the Gaussian tail function. From (2.44) and (2.46) we observe that the authentication fingerprint signal SER decreases when N or L_t are increased, potentially allowing for fingerprint BERs lower than the primary signal BER in some channel stationarity conditions.

2.5.4 APM Fingerprint Distortion

We now consider worst case distortions present when equalizing the APM-fingerprinted signal according to incorrect channel information as was previously done for the AAM

fingerprinting function. The transmitted symbol with fingerprinting function present, (2.40), becomes

$$\mathbf{X}[t] = \begin{bmatrix} e^{-j\epsilon} & 0 \\ 0 & e^{j\epsilon} \end{bmatrix} \begin{bmatrix} d_1 & -d_2^* \\ d_2 & d_1^* \end{bmatrix} = \begin{bmatrix} d_1 e^{-j\epsilon} & -d_2^* e^{-j\epsilon} \\ d_2 e^{j\epsilon} & d_1^* e^{j\epsilon} \end{bmatrix}, \tag{2.47}$$

and the received ST signal becomes

$$\mathbf{Y}_{AMP}[t] = \begin{bmatrix} r_1 & r_3 \\ r_2 & r_4 \end{bmatrix} + \mathbf{N}[t]\mathbf{P}^H, \tag{2.48}$$

where

$$\begin{aligned} r_1 &= \alpha_1 d_1 e^{j(\theta_1-\epsilon)} + \alpha_3 d_2 e^{j(\theta_3+\epsilon)}, \quad r_3 = -\alpha_1 d_2^* e^{j(\theta_1-\epsilon)} + \alpha_3 d_1^* e^{j(\theta_3+\epsilon)}, \\ r_2 &= \alpha_2 d_1 e^{j(\theta_2-\epsilon)} + \alpha_4 d_2 e^{j(\theta_4+\epsilon)}, \quad r_4 = -\alpha_2 d_2^* e^{j(\theta_2-\epsilon)} + \alpha_4 d_1^* e^{j(\theta_4+\epsilon)}. \end{aligned} \tag{2.49}$$

Thus, the signal estimates for d_{P_1} and d_{P_2} using MRC without correcting for the phase perturbation, denoted $\tilde{d}_{P_1}$ and $\tilde{d}_{P_2}$, become

$$\begin{aligned} \tilde{d}_{1_{APM}} = & \\ & \hat{\alpha}_1 e^{-j\hat{\theta}_1}\left(\alpha_1 d_1 e^{j(\theta_1-\epsilon)} + \alpha_3 d_2 e^{j(\theta_3+\epsilon)}\right) + \hat{\alpha}_2 e^{-j\hat{\theta}_2}\left(\alpha_2 d_1 e^{j(\theta_2-\epsilon)} + \alpha_4 d_2 e^{j(\theta_4+\epsilon)}\right) \\ & + \hat{\alpha}_3 e^{j\hat{\theta}_3}\left(-\alpha_1 d_2 e^{-j(\theta_1-\epsilon)} + \alpha_3 d_1 e^{-j(\theta_3+\epsilon)}\right) \\ & + \hat{\alpha}_4 e^{j\hat{\theta}_4}\left(-\alpha_2 d_2 e^{-j(\theta_2-\epsilon)} + \alpha_4 d_1 e^{-j(\theta_4+\epsilon)}\right) + \eta_1, \\ \tilde{d}_{2_{APM}} = & \\ & - \hat{\alpha}_1 e^{j\hat{\theta}_1}\left(-\alpha_1 d_2 e^{-j(\theta_1-\epsilon)} + \alpha_3 d_1 e^{-j(\theta_3+\epsilon)}\right) - \hat{\alpha}_2 e^{j\hat{\theta}_2}\left(-\alpha_2 d_2 e^{-j(\theta_2-\epsilon)} + \alpha_4 d_1 e^{-j(\theta_4+\epsilon)}\right) \\ & + \hat{\alpha}_3 e^{-j\hat{\theta}_3}\left(\alpha_1 d_1 e^{j(\theta_1-\epsilon)} + \alpha_3 d_2 e^{j(\theta_3+\epsilon)}\right) \\ & + \hat{\alpha}_4 e^{-j\hat{\theta}_4}\left(\alpha_2 d_1 e^{j(\theta_2-\epsilon)} + \alpha_4 d_2 e^{j(\theta_4+\epsilon)}\right) + \eta_2. \end{aligned} \tag{2.50}$$

Using (2.37), after some manipulation, (2.50) becomes

$$\tilde{d}_{1_{APM}} = (\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4)d_1 e^{-j\epsilon} + \eta_1, \;\; \tilde{d}_{2_{APM}} = (\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4)d_2 e^{j\epsilon} + \eta_2, \tag{2.51}$$

with η_1 and η_2 given in (2.35).

From (2.51) we observe the worst case distortions from the extrinsic APM fingerprint function when equalizing according to outdated CSI for the 2x2 Allmouti code. Here, worst case model mismatch error introduces a phase rotation of $e^{-j\epsilon}$ in $\tilde{d}_{1_{APM}}$, and $e^{j\epsilon}$ in $\tilde{d}_{2_{APM}}$. Because the amount of distortion APM fingerprint is also proportional to ϵ, we note that like the AAM fingerprinting function, care must be taken when choosing ϵ when the performance of equalizers employed by unaware receivers must be considered.

2.6 Simulation Results

We now present simulation results for the AAM fingerprinting constellation (2.28), for different values ϵ and channel AR model parameter a in (2.9), using the MMSE channel estimator, the 2x2 Alamouti ST code with $M = 4$, and $N = L_t = L_r = 2$. A QPSK constellation was used for the primary signal. The authentication performance of arbitrary authentication messages can be computed directly from the fingerprint BER, therefore we will use BER in this section to demonstrate fingerprint signaling performance. The results for the AAM fingerprinting function for a fixed $\sigma_T = 0.01$ and values of a equal to 0.7 and 0.9 are presented in Figure 1. We observe that for both values of a, the authentication fingerprint is received with a BER advantage over the primary signal. We also see that the BER for the fingerprint signal is less when $a = 0.9$ than

when $a = 0.7$, suggesting that the fingerprint signal performance does indeed depend on correlation between channel estimates in time, determined by the AR-1 model parameter a.

A plot of the BER for both the primary signal and authentication signal is given in Figure 2 for a fixed $a = 0.7$ and values of ϵ equal to 0.45 and 0.47. As the value of ϵ increases, the signal strength for the authentication signal increases, resulting in an improved BER for the authentication signal at the expense of a slight increase in channel estimate MSE for the primary-user. We observe that the AAM fingerprinting function yields an authentication signal BER advantage over the primary signal for $\epsilon = 0.47$, over the range of SNR plotted.

In Figures 3 and 4 we present results for channel estimate MSE and the worst-case MSE for the simulations depicted in Figures 1 and 2, respectively. The worst-case MSE results represent the additional model error incurred if $\mathbf{Y}[\tau_1]$ were incorrectly equalized using $\mathbf{H}[\tau_0]$ as opposed to $\mathbf{H}[\tau_1]$, as suggested by (2.38).

We note that worst-case MSE is invariant of the AR-1 model parameter a, as the MSE for $a = 0.7$ and $a = 0.9$ are nearly indistinguishable. From Figure 4 we observe that increasing ϵ results in an increased channel model MSE as expected, and the worst-case error introduced by the fingerprinting function is apparent from the difference between MSE results when the fingerprint is present and when the fingerprint absent.

The results for the APM fingerprinting function for a fixed $\sigma_T = 0.3$ and values of a equal to 0.8 and 0.9, are presented in Figure 5. We observe that like the AAM function, the fingerprint is received with a greater BER advantage over primary signal for $a = 0.9$, suggesting that the APM fingerprint signal performance also depends on correlation be-

tween channel estimates in time, as determined by the AR-1 model parameter a. We also observe that in higher SNR, around 8dB, the slope of the BER curve for the authentication behaves differently for the case where $a = 0.8$, when compared to the authentication signal BER curve for $a = 0.9$. In particular, the authentication signal BER curve slope for $a = 0.8$ stops changing after 8dB. This can be explained as follows:

In our time-variant channel model (2.7), the channel matrix $H_{i,j}[t]$ is a summation of two independent noise processes, $N_{i,j}[t]$, which is a white Gaussian noise process with variance σ_N^2, and a colored Gaussian noise process $\mu_{i,j}[t]$, which is modeled as an AR-1 process driven by $u_{i,j}[t] \sim \mathcal{CN}(0, \sigma_T^2)$. For higher values of SNR, i.e. as σ_N^2 decreases, the dominating noise process when decoding the authentication signals becomes $\mu_{i,j}[t]$, and not $N_{i,j}[t]$. This effect becomes more pronounced as the bandwidth of the time-varying component $\mu_{i,j}[t]$ increases, which is inversely proportional to the AR-1 model parameter a. Thus, for large values of σ_T^2, the performance of the authentication signal degrades more rapidly under high SNR, as the value of a decreases. This is the scenario of rapidly varying channel.

A plot of the BER for both the primary signal and authentication signal is given in Figure 6 for a fixed $a = 0.9$ and values of σ_T equal to 0.1 and 0.3. As the value of σ_T^2 increases, the power of the time-varying channel component increases resulting in a greater channel estimate MSE for both the primary and authentication signal, and a decreased system BER for both signals. We note from Figure 6 that the APM fingerprint signal BER is lower than the primary for the range of SNR simulated.

In Figures 7 and 8 we plot the worst-case mean-squared error of the channel estimate using the APM fingerprinting function, suggested by (2.51), as was done for the

AAM fingerprinting function. We observe the MSE, and worst-case MSE, experienced by the MMSE receiver, as suggested by (2.51). We note that worst-case MSE is relatively invariant of the AR-1 model parameter a, as the MSE for $a = 0.8$ and $a = 0.9$ are completely overlapping and indistinguishable. From Figure 8 we observe that increasing σ_T^2 results in an increased channel model MSE for worst-case distortions as expected, and the worst-case error introduced by the fingerprinting function is apparent from the difference between MSE results when the fingerprint is present and when the fingerprint absent.

We conclude from these results that the APM fingerprinting function generally has better performance over the AAM fingerprinting function, for larger values of σ_T^2, for given parameters.

Some more discussion on the threat model used in this work may be in order, to describe the kinds of security attacks our system may come under. We note that a rigorous security analysis is not given in this work to show how well the proposed scheme works when attackers are present in the system. For this topic we present the following discussion:

Because the fingerprint is embedded at the PHY-layer and recovered at the PHY-layer, an attack model must consider a few PHY-layer attacks. We propose that the fingerprint function $\mathbf{F}[\tau_1]$, which conveys one symbol of the authentication message every two STC blocks is applied to the entire fingerprinted block (i.e. embedded into M primary-signal symbols), including the pilot signals. If an attacker wanted to, for example, replace the identity of a transmission with a forgery via a PHY-layer attack, she would be required to not only forge $\mathbf{F}[\tau_1]$ but the entire transmission $\mathbf{U}[\tau_1]$ as a whole including both the deterministic pilot signals $\mathbf{P}$ as well as the the primary-data transmission $\mathbf{D}[\tau_1]$. If one were

to attack only the pilot signal $\mathbf{P}$ with a forged $\mathbf{F}[\tau_1]$, then $\mathbf{P}$ would be out-of-reference with $\mathbf{D}[\tau_1]$ (these match in a normal, healthy transmission), a phenomena that will most certainly lead to severe distortion when decoding the primary-signal $\mathbf{D}[k]$ thus making the attack ineffective (i.e. The attacker is now merely a jammer). The proposed embedding scheme does not exacerbate the efficiency of jamming attacks, nor does it attempt mitigate these attacks. If the authentication signal were designed using a modulation and FEC that is more robust than the primary-signal $\mathbf{D}[k]$, then the attacker would require more jamming energy to jam the authentication signal than is required to jam $\mathbf{D}[k]$, and thus $\mathbf{D}[k]$ has already been destroyed at this point. The best that the attacker can do in conducting a PHY-layer attack against the original transmission would be to jam the signal entirely, as $\mathbf{F}[\tau_1]$, $\mathbf{P}$, and $\mathbf{D}[\tau_1]$ become a gestalt and are decoded as a whole.

We may also want to discuss an attacker, where the attacker targets only transmissions where the fingerprint is present. We offer the following explanation: The function $\mathbf{F}[\tau_1]$, which conveys one symbol of the authentication message every two STC blocks is applied to the entire fingerprinted block (i.e. embedded into M primary-signal symbols), including the pilot signals. If an attacker wanted to, for example, replace the identity of a transmission with a forgery via a PHY-layer attack, she would be required to not only forge a new $\mathbf{F}[\tau_1]$ but the entire transmission $\mathbf{U}[\tau_1]$ as a whole including both the deterministic pilot signals $\mathbf{P}$ as well as the the primary-data transmission $\mathbf{D}[\tau_1]$. If the attacker were to attack only the pilot signal $\mathbf{P}$ with a forged $\mathbf{F}[\tau_1]$, then $\mathbf{P}$ would be out-of-reference with $\mathbf{D}[\tau_1]$ (these match in a normal, healthy transmission), a phenomena that will most certainly lead to severe distortion when decoding the primary-signal $\mathbf{D}[k]$ thus making the attack ineffective (i.e. Eve is now merely a jammer). The proposed

embedding scheme does not exacerbate the efficiency of this type of attack, nor does it mitigate these attacks. If the authentication signal were designed using a modulation and FEC that is more robust than the primary-signal $\mathbf{D}[k]$, then the attacker would require more jamming energy to jam the authentication signal than is required to jam $\mathbf{D}[k]$, and thus $\mathbf{D}[k]$ has already been destroyed at this point. The best that the attacker can do in conducting a PHY-layer attack against the original transmission would be to jam the signal entirely, as $\mathbf{F}[\tau_1]$, $\mathbf{P}$, and $\mathbf{D}[\tau_1]$ become a gestalt, and are decoded as a whole.

Additionally, in this work we explicitly avoid detailed discussion of the actual message being authenticated. This is because we assume that best practices are observed when designing the message. These are important details for the implementation of a system, however the focus of this paper is the PHY-layer embedding of the fingerprint signal using the differential precoding method described. This is why we use BER for the authentication signal as a measure of performance, since we are primarily concerned with the performance of our embedding and signaling scheme. The design of a protocol and accompanying message structure that uses our signaling method to address a particular security problem problem is interesting, however, such designs have been discussed heavily in distributed systems literature and are outside of the scope of this work.

One viewpoint is that we suggest replacing the strong authentication abilities provided by traditional cryptographic schemes with a potentially weaker PHY-layer mechanism. This viewpoint, however, is a misconception. We do not propose the replacement of traditional schemes, but instead want leverage traditional cryptographic schemes in the design of the embedded authentication message. As stated in the introduction section of this chapter, our work is on *embedding* a cryptographically secure message at PHY-layer.

This implies that traditional cryptographic schemes can be applied to the physical layer without changing higher-layer protocols.

In one embodiment a full authentication system may send independent authentication messages via both PHY-layer and higher-layer methods, and use both messages to jointly authenticate the transmission. While we suggest that best practices are observed in designing both of these messages, to cryptographically secure them from attack, this paper focuses on the physical-layer details of conveying the message to the receiver as an embedded signal.

One might also question if the proposed scheme only works when the receiver knows the identity of the sender, and if we need to know the sender's fingerprint signal *a priori*. This is of concern since because in most cases authentication is needed when a new node joins the system, and may not know the exact message the joining node will send, and this problem is typically called the *bootstrapping problem*.

While our work focuses on the PHY-layer details of embedding an authentication message via a novel precoding approach, we do suggest that a full system should have a trusted authority that can issue unforgeable tokens (specifically, capabilities [18]) prescribing the time and frequency a signal is allowed to transmit. However, this "bootstrapping" problem is outside of the scope of our work and could be considered as a future research direction. We suggest that best practices are observed when designing the mechanism by which new nodes join the system, leveraging prior work conducted in the fields of distributed computing and cryptography.

We may also want to provide additional discussion about the comprehensiveness of our comparison of this work and existing fingerprint techniques. The PHY-layer em-

bedding of fingerprint signals into digital wireless transmissions is a relatively new field of research. In fact, to the best our knowledge this is the first work considering the fingerprinting of MIMO transmissions (i.e. where the fingerprint message can be embedded via both time an spacial perturbations. Four previous approaches were briefly discussed, however these approaches are fundamentally different from the method described in this work. These methods rely on blind superposition and do not exploit receiver preprocessing algorithms in their embedding. Further, none of the approaches consider MIMO transmissions. Of the previous work considered, [64] is the closest to our approach, is it investigates the evolution of time-varying channel state implicate malicious transmissions. This work is discussed in the most detail.

One may question why we argue that a PHY-layer authentication can be decoupled from higher-layer authentication mechanisms. This is in part true for intrinsic fingerprinting (channel based fingerprinting). However, for the extrinsic scheme we propose, we rely on a certificate to obtain a public key/private key which is used for signing an authenticator message. This approach does, therefore, rely on some notion of higher-layer authentication. We may also want to discuss how the benefits of PHY-layer authentication are tightly connected to the context of this work.

In this chapter we use the term "higher-layer" refers to the embedding of the authentication message in the upper layers (i.e. as in the OSI model, for example) of the protocol stack of the primary-signal; specifically, in the bit-level protocol of the payload of the primary-signal.

We suggest leveraging traditional, cryptographically secure, best practices when designing the digital authentication message. We do not propose the replacement of tra-

ditional schemes, we want leverage traditional cryptographic schemes in the design of the embedded authentication message. As stated in our introduction in the last paragraph, our work is on *embedding* a cryptographically secure message at PHY-layer. So the traditional cryptographic schemes can be applied when designing the physical layer authentication signal, without changing the higher-layer protocol.

We would also like to discuss the odd-even delivery scheme in greater detail, an give another explanation of its purpose. The odd-even delivery scheme describes a *differential modulation scheme*, where the bits of each authentication signal symbol are decoded based on the difference between channel estimates obtained during the channel-sounding symbol and the adjacent block. The intrinsic channel will add its own distortions to the signal, and this differential modulation scheme helps mitigate the effects of these distortions by using the channel-sounding block as a reference when decoding the authentication symbol. The odd-even transmission scheme does not suggest that only half of the transmission is fingerprinted.

This scheme transmits one authentication symbol (i.e. through application of $\mathbf{F}[k]$) every $2M$ space-time symbols. The channel estimate derived from the first M symbols of the channel-sounding block is used for reference when detecting $\mathbf{F}[k]$, which has been applied to the next M symbols. Changing $\mathbf{F}[k]$ more frequently than M symbols will cause implicit distortion since the block-fading assumption leveraged by the channel estimation mechanism no longer holds (i.e. the manipulations of $\mathbf{F}[k]$ are happening faster than the channel estimation mechanism can mitigate, and thus the transmitter will send irreparably distorted blocks. Preforming channel sounding less frequently will degrade the detection of $\mathbf{F}[k]$.

Perhaps the most interesting theoretical aspects of this work are related to the fact that $\mathbf{F}[k]$ and $\mathbf{H}[k]$ are both time-varying, and thus the product of $\mathbf{H}[k]\mathbf{F}[k]$ has less coherence time than $\mathbf{H}[k]$ alone. Some discussion on more general channel sounding practices is in order. The most efficient channel sounding scheme would dynamically change the frequency at which $\mathbf{F}[k]$ is transmitted, in relation to the actual coherence time of $\mathbf{H}[k]$. For very slowly time-varying channels the physical channel does not require as frequent sounding, and thus $\mathbf{F}[k]$ can be transmitted more frequently leading to an increased baud rate for the authentication message. The trade-off here is that more frequent channel sounding provides better resistance against quickly changing channels where less frequent channel sounding results in less overhead for the transmission of the authentication signal.

In this work we refer only to the fingerprinting of the modulated signal, or waveform, and not the manipulation of the bits of the payload message. In fact, we specifically seek designs that minimize bit errors in the primary-signal and use the bit error rate of the primary-signal as a measure of perceptual distortion to the primary-signal due to the embedding of the authentication message. Steganography is the practice of hiding messages in such a way that no one, apart from the sender and intended recipient, suspects the existence of the message. The practice of steganography is not limited to embedding messages in digital data, and in striking contrast to aforementioned definition we explicitly want everyone, friend or foe, to know that the message exists and have the ability to decode the message to revel the identity and authenticity of the message's source.

When referring to the embedding of the authentication message into the primary-signal, we use the term 'signal' in it's most general form, and that is the analog form of the

signal, or the discretized version, or digital signal. Because of this generality, we use the terms watermarking and fingerprinting interchangeably in the introduction when referring to previous work using both of these terms. However, in both cases we are fingerprinting the modulated signal and not the bits of the payload.

To bolster our assumption that $\mathbf{N}[k]\mathbf{P}^H$ will be white noise, as opposed to colored noise, we give the following discussion: For the case where $\mathbf{P}$ is a sparse matrix with pilot signals uniformly distributed throughout, $\mathbf{N}[k]\mathbf{P}^H$ will be Gaussian and white, as the projection simply decimates the Gaussian process $N_{i,j}[k]$, for a given row of $\mathbf{N}[k]$, and combines the Gaussian elements linearly. For example, consider the case where $\mathbf{P}^H$ is a uniform, sparse matrix, with 1's every 4th element (both row-wise and column-wise), and zeros in the remaining elements. In this case $\mathbf{N}[k]\mathbf{P}^H$ simply decimates each row of $\mathbf{N}[k]$ by 4 by selecting every 1-in-4 elements and creates a linear combination of Gaussian random variables $N_{i,j}[k]$. In most practical pilot embedding systems, the pilot signal matrix $\mathbf{P}^H$ is comprised of zeros and constant energy signals, otherwise the SNR of elements in the channel estimate $\mathbf{H}[k]$ would be unequal necessitating a non-ML decoder.

As the linear combination of Gaussian random variables is also Gaussian, and for practical pilot matrices with zeros and constant energy elements the projection of $\mathbf{N}[k]\mathbf{P}^H$ should be Gaussian and white if $\mathbf{N}[k]$ is Gaussian and white.

We note that the additive work in [66] also claims to use very little power for the perturbation used to create the primary-signal, resulting in almost no impact on the receiver BER. To delineate the main differences between our work and this work, with respect to primary-signal degradation, we give the following discussion:

In [66] the impact to primary-signal BER due to interference from the authentica-

tion signal can be manipulated to achieve arbitrarily high performance through the design of an arbitrarily long code. The length of the resulting code, however, adversely affects the rate that the authentication signal can send information. For example, one could design a code that is spread over 10 million primary-signal symbols, perturbing each symbol an immeasurable amount, and the matched-filter for this code would produce a high confidence result for one bit of information (i.e the signal is present or not). The drawback is, of course, that this design transmits only a single bit of authentication for every 10 million symbols of primary data. Another drawback would be the computation required to run a 10 million tap matched-filter at the full baud rate of the primary-signal.

The blind superposition approaches, like [66], implicitly achieve authentication signal capacity by reducing primary-signal capacity, since the capacity of the original band-limited channel must remain the same under a TX energy constraint. In the example above, one ten-millionth of a bit of capacity is borrowed from every symbol of primary-signal capacity and is transfered to the authentication signal.

Our system takes a fundamentally different approach. In our scheme the authentication signal, as modulated by the fingerprint function $\mathbf{F}[k]$, is detected by analyzing channel estimate data, not primary-signal data. As channel estimates are derived from the pilot signals $\mathbf{P}$, the capacity for our authentication signal is recovered from the pilot signals ('recovered' because pilot signals traditionally carry zero information at the time of transmission). Essentially $\mathbf{H}[k]\mathbf{F}[k]$ conveys two pieces of information to the receiver: 1) Changes in intrinsic channel state, and 2) Changes in $\mathbf{F}[k]$. For example, in completely stationary channel conditions the channel estimates $\mathbf{H}[k]$ are correlated, and thus convey zero new information to the receiver. This wasted capacity can be recovered by

extrinsically modulating the evolution of channel estimates at the transmitter, i.e. through $\mathbf{H}[k]\mathbf{F}[k]$. In the perfect channel stationarity case, the system capacity reserved for $\mathbf{P}$ can be recovered by extrinsically increasing uncertainty in consecutive channel estimates via $\mathbf{F}[k]$, without changing the density of the pilot signals (which would decrease the efficacy of the channel estimation mechanism).

A casual reader familiar with previous blind-superposition fingerprinting methods may assert that this work attempts to achieve equivalent perturbations in amplitude and phase shift in the constellation points of the primary-signal, using the additive model like [66]. This is, however, a misconception which we will offer the following clarification:

The spirit of our approach is not to change the primary-signal constellation points. Our approach processes the evolution of channel-estimate data, exclusively, and ignores primary-signal data. This assertion is corroborated by (2.3), since $\mathbf{A}$ and $\mathbf{P}$ are designed such that the pilot signals and the data signals are orthogonal. Since channel estimates are derived from pilot signals when using pilot-aided channel estimation, the underlying constellation points used by $\mathbf{D}[k]$ are explicitly orthogonal to the statistics used by our fingerprint detector, since the precoding function $\mathbf{F}[k]$ does not destroy the orthogonality of (2.3).

Two example designs for $\mathbf{F}[k]$ are presented in this chapter, the AAM design and the APM design, and these designs modulate the fundamental parameters of signal amplitude and signal phase respectively. In this chapter only the receiver perceptual model is considered. Designs that also consider anticipated channel distortions in the design of $\mathbf{F}[k]$, are presented in Chapter 3. In this chapter analysis is presented for the 2x2 Alamouti, however we would like to address the question of what can be said about larger

codes. We offer the following discussion on this topic:

The performance of a particular code will, in general, depend greatly on the channel model used for $\mathbf{H}[k]$ and the design of $\mathbf{F}[k]$. The underlying STC used for $\mathbf{D}[k]$ is inconsequential, as we use only channel estimate data and not primary-signal data. The statistical analysis in this chapter is for the fingerprint design and detection in Rayleigh fading, which is not specific to the the dimensionality of a particular STC. It is instead parameterized by the dimensionality of $\mathbf{P}$. In the design of the fingerprint, all elements of $\mathbf{Q}[k]$ are combined to produce a single statistic, $\hat{\epsilon}$, which may not be the case in all designs of $\mathbf{F}[k]$. Our embedding scheme can be easily extended to any orthogonal space-time code just with more parameters, such as a 4×4 code, by simply perturbing the additional elements (i.e. in the spirit of AAM or APM). For example, the APM could trivially be extended to something like:

$$\mathbf{F}_{APM}[t] = \begin{bmatrix} e^{-j\epsilon} & 0 & 0 & 0 \\ 0 & e^{j\epsilon} & 0 & 0 \\ 0 & 0 & e^{-j\epsilon} & 0 \\ 0 & 0 & 0 & e^{j\epsilon} \end{bmatrix}, \ 0 \leq \epsilon < 2\pi. \tag{2.52}$$

One may question why simulation results were not presented for the stationary channel case. We offer the following explanation on why these results were were not presented:

The stationary channel case is obtained by setting $\sigma_T^2 = 0$, since σ_T^2 is the predominant noise term from $u_{i,j}[t]$ and contributes the most uncertainty between $\hat{\mathbf{H}}[\tau_0]$ and $\hat{\mathbf{H}}[\tau_1]$. This term represents the power of intrinsic channel distortions and is primarily

responsible for decreasing fingerprint detection performance. As depicted in Figure 6, the two values simulated for σ_T^2 are currently quite large ($\sigma_T^2 = .1$ and $\sigma_T^2 = .3$). Setting $\sigma_T^2 = 0$ would remove most of the uncertainty between channel estimates resulting in extremely good authentication signal BER. Simulating this case, while possible, would require an an exceptional number of simulations due to the infrequency of authentication bit errors under these conditions.

The conclusions from the simulations may seem obvious: The authentication SER goes down if you increase the amplitude or phase changes. The simulation results, however, provide a quantitative evaluation of the performance of the authentication scheme that may useful for future system design. The primary focus of the simulation results is to corroborate the analysis, and the veracity of the presentation of these results.

2.7 Conclusion

In this chapter we presented a framework for fingerprinting MIMO transmissions with a digital PHY-layer message for the purpose of transmitter authentication. We demonstrated that the fingerprint signal can be added without modifying the decoding process of unaware, or traditional MIMO receivers. Further, the distortions introduced by the fingerprint can be partially removed by the receiver's equalizer to reduce the degradation in performance of the primary transmission. It was demonstrated that the fingerprint signal can be designed with a BER lower than the primary signal, and that the probability of symbol error for the proposed method improves as the correlation between time-varying channel estimates increases. Our proposed scheme provides the foundation of

fingerprint signaling which can be used to embed authentication messages of arbitrary lengths for secure wireless transmissions.

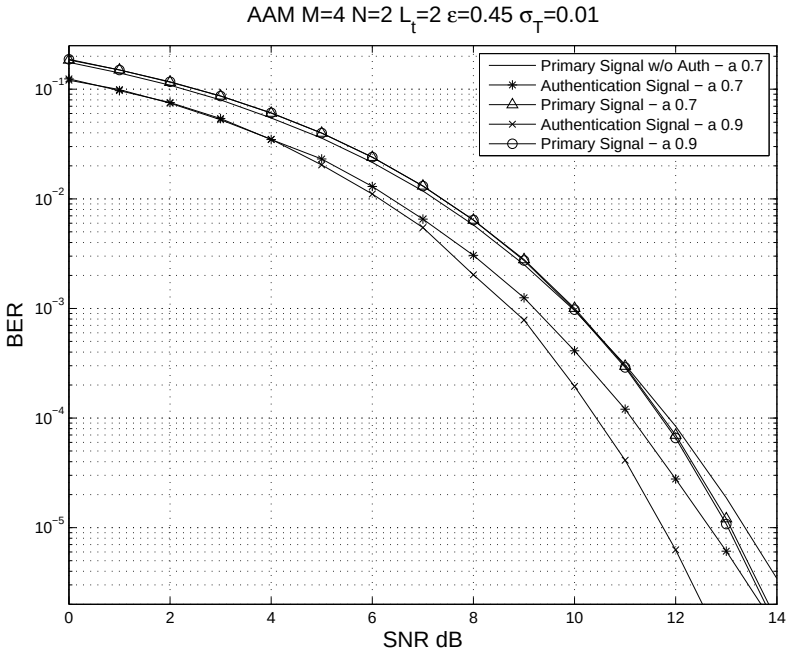


Figure 1: BER for primary and AAM fingerprint signal for various a

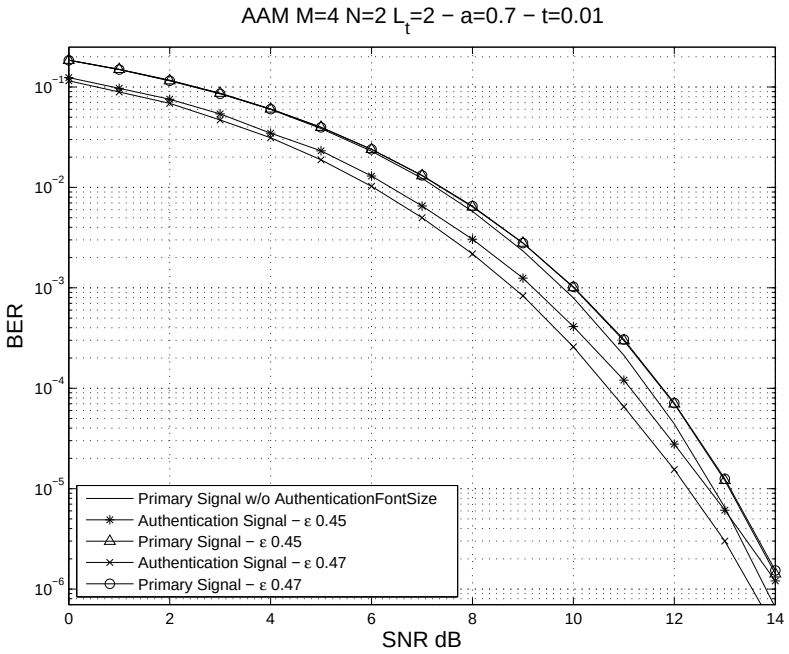


Figure 2: BER for primary and AAM fingerprint signal for various ϵ

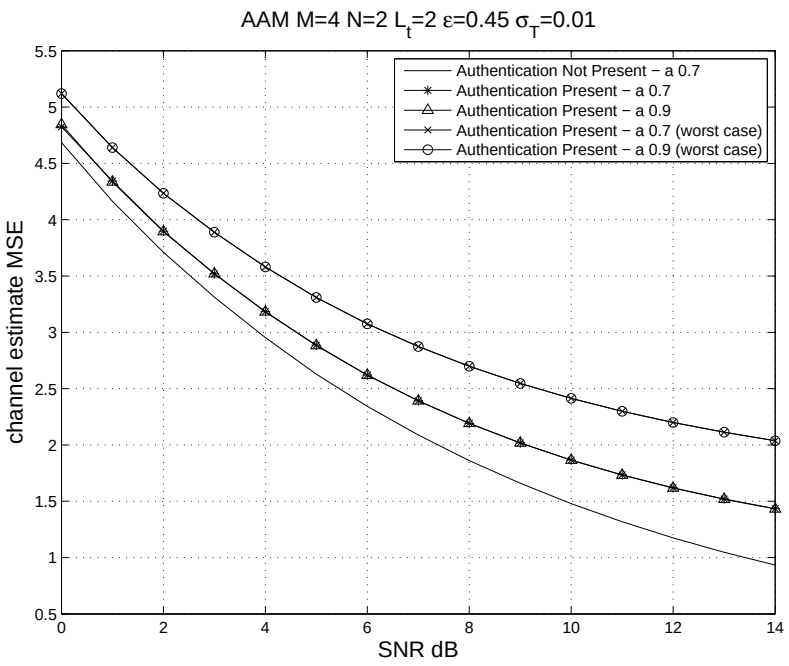


Figure 3: MSE of the channel estimate with and without AAM fingerprint signal for various a

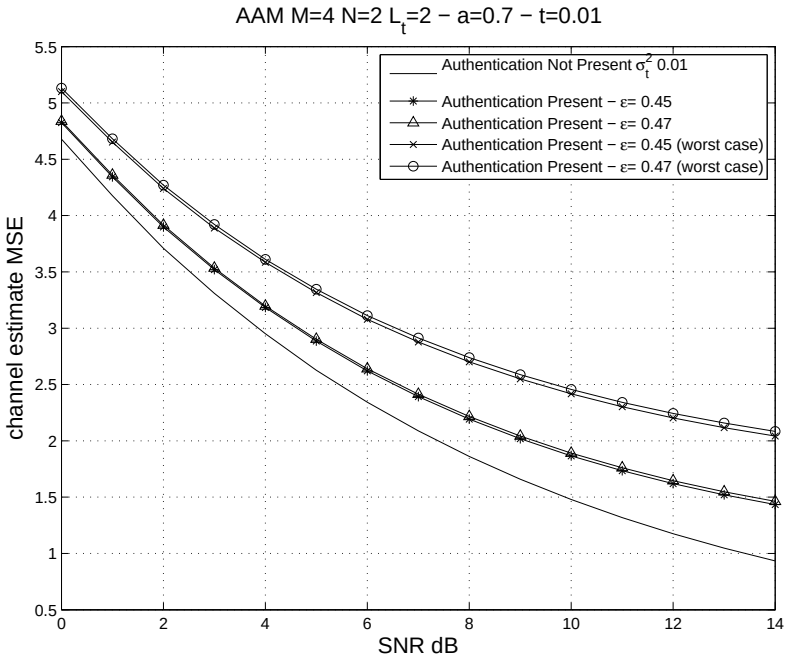


Figure 4: MSE of the channel estimate with and without AAM fingerprint signal for various ϵ

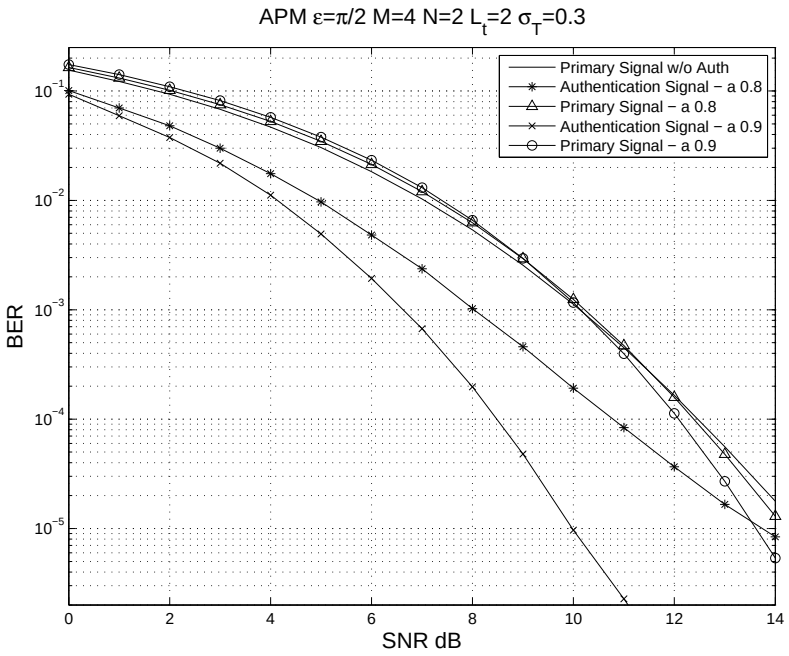


Figure 5: BER for primary and APM fingerprint signal for various a

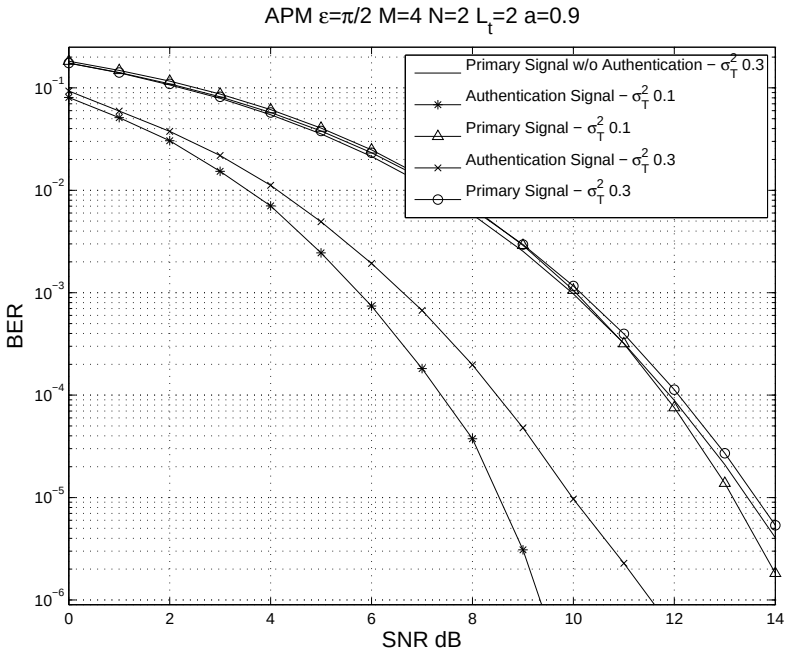


Figure 6: BER for primary and APM fingerprint signal for various σ_T

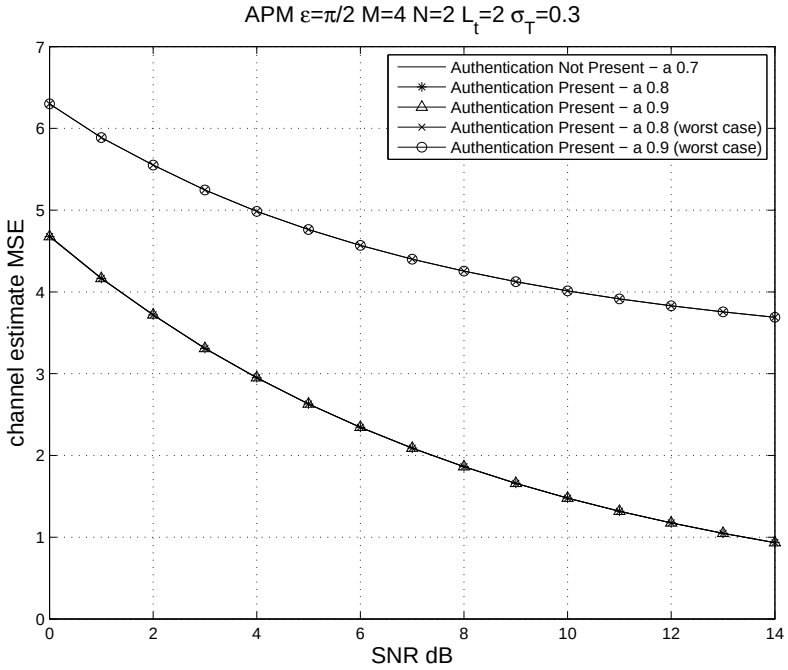


Figure 7: MSE of the channel estimate with and without APM fingerprint signal for various a

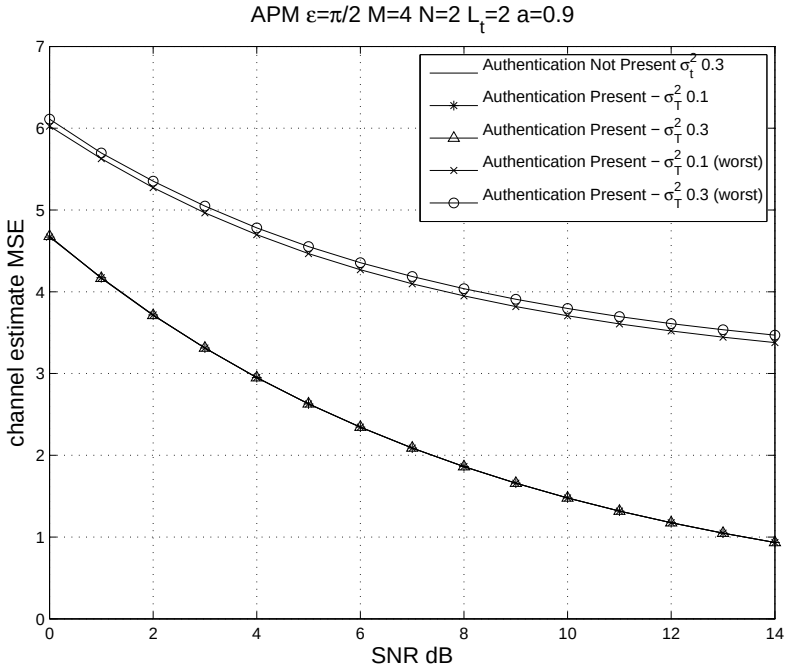


Figure 8: MSE of the channel estimate with and without APM fingerprint signal for various σ_T

Chapter 3

Extrinsic Channel-Like Fingerprinting Overlays Using Subspace Embedding

3.1 Overview

We present a physical-layer fingerprint-embedding scheme for orthogonal frequency division multiplexing (OFDM) transmissions, where the fingerprint signal conveys a low capacity communication suitable for authenticating the transmission and further facilitating secure communications. Our system strives to embed the fingerprint message into the noise subspace of the channel estimates obtained by the receiver, using a number of signal spreading techniques. When side information of channel state is known and leveraged by the transmitter, the performance of the fingerprint embedding can be improved. When channel state information is not known, blind spreading techniques are applied. The fingerprint message is only visible to aware receivers who explicitly preform detection of the signal, but is invisible to receivers employing typical channel equalization. A taxonomy of overlay designs is discussed and these designs are explored through experiment using time-varying channel-state information (CSI) recorded from IEEE802.16e Mobile WiMax base stations. The performance of the fingerprint signal as received by a WiMax subscriber is demonstrated using CSI measurements derived from the downlink signal. Detection performance for the digital fingerprint message in time-varying channel

conditions is also presented via simulation.

3.2 Introduction

In this chapter we present a physical layer (PHY) fingerprinting method for orthogonal frequency division multiplexed (OFDM) transmissions, where side information about anticipated channel conditions is incorporated into the fingerprint design. A number of PHY-layer fingerprinting approaches for wireless communications have been investigated using basic signal superposition methods [61, 47, 63, 66]. The main disadvantage of blind superposition is that the fingerprint signal appears as interference to the original signal and is fully present when the signal is decoded, resulting in decreased SNR for the original signal. Instead we investigate fingerprint designs that consider how the signal will be perceived by the receiver, and side information of the channel distortions that the signal will experience. By leveraging channel side information and considering the receiver's perception of the signal, improved fingerprint designs [17] are possible as the undesirable effects of the fingerprint signal associated with blind superposition approaches [47] and [63] are partially removed by the receiver when preprocessing the signal.

In [24] an extrinsic channel-like fingerprint for narrow-band single-input single-output (SISO) digital television signals was considered, where the fingerprint message is applied at the transmitter by emulating nominal multipath channel responses. In [26] an extrinsic channel-like fingerprint signal is considered for multiple-input-multiple-output (MIMO) systems using space-time block codes (STBC). In this work we extend these previous approaches to Orthogonal Frequency Division Multiplexing (OFDM) signals,

and we incorporate previous channel state knowledge as side information into the design of the fingerprint signal.

We propose a number of techniques to spread the fingerprint message in both frequency and time domain using subspace decomposition. When the full Channel State Information (CSI) is known and leveraged by the transmitter, the fingerprint can be embedded into the noise subspace of the receiver's channel estimates to ensure that the fingerprint signal incurs minimal distortion during transmission. When channel state information is not known by the transmitter, blind spreading designs using orthogonal codes, such as Walsh codes, can be applied. We present a number fingerprint spreading designs that incorporate various amounts of previous CSI into the design of the spread fingerprint signaling bases. Our designs demonstrate that additional CSI knowledge can be leveraged by the transmitter to improve the performance of the fingerprint embedding.

This chapter is organized as follows: Section 3.3 describes the OFDM system and presents a framework for introducing the channel-like fingerprint. A time-varying channel model is given. In Section 3.5 an analysis of the extrinsic fingerprint overlay design is given, and the embedding and recovery of the fingerprint message is demonstrated. A taxonomy of overlay designs is also presented. In Section 3.6 we present experimental results, where one fingerprint overlay design is evaluated using CSI extracted from an IEEE802.16e WiMax transmission. In Section 3.7 we present simulation results for two of the fingerprint overlay designs. In Section 3.8 we present our conclusions. The following is a list of the most frequently used notation in Sections II, III, and VI.

- $\boldsymbol{\xi}[k]$: Extrinsic fingerprint signal vector

- $\mathbf{f}[k]$: Extrinsic fingerprint signaling vector to be applied the TFB transmission at time k
- $\hat{\mathbb{H}}[k]$: Block of frequency domain intrinsic channel estimate vectors
- $\mathbb{F}[k]$: Block of frequency domain extrinsic fingerprint vectors
- $\hat{\mathbb{Q}}[k]$: Block of frequency domain aggregate channel estimate vectors
- $\hat{\mathbb{L}}[k]$: Block of frequency domain intrinsic channel estimate vectors
- $\mathbb{U}_2[k]$: Time-spreading fingerprint signaling basis
- $\mathbb{V}_2[k]$: Frequency-spreading fingerprint signaling basis
- $\mathbb{K}[k]$: Extrinsic fingerprint overlay block
- $\mathbb{E}[k]$: Intrinsic channel model mismatch error block

3.3 System Model

We consider an OFDM system where the transmission is subjected to a linear time-domain channel response $g(t)$, given as

$$g(t) = \sum_c A_c(t)\delta(t - \Delta\tau_c), \tag{3.1}$$

where Δ is the sampling interval, τ_c are the delays for each channel component, and $A_c(t)$ are the complex valued delay-spread gains at time t for multipath component c. The OFDM system is modulated using an N-point discrete-time inverse Fourier transform

($IDFT_N$), and then subsequently demodulated using an N-point discrete-time Fourier transform (DFT_N). The matrix representation of the OFDM system is given as

$$\mathbf{y} = \mathbf{gWX} + \mathbf{n}, \tag{3.2}$$

where $\mathbf{y} = [y_0\ y_1\ \cdots\ y_{N-1}]$ is the received band-limited signal vector after DFT_N, where elements $y_n \in \mathcal{C}$, the matrix $\mathbf{X} = \text{diag}([x_0\ x_1\ \cdots\ x_{N-1}])$ is of size $\mathcal{C}^{(N\times N)}$, $\mathbf{x} = [x_0\ x_1\ \cdots\ x_{N-1}]$ are the data symbols to be fingerprinted and transmitted, $\mathbf{W}$ is a $\mathcal{C}^{(N\times N)}$ DFT-matrix with elements $W_N^{nk} = \frac{1}{\sqrt{N}} e^{-j2\pi\frac{nk}{N}}$, using row index $n = 0, \ldots, N-1$ and column index $k = 0, \ldots, N-1$, the vector $\mathbf{g} = [g_0\ g_1\ \cdots\ g_{N-1}]$ is the sampled channel impulse response, where each element $g_i \in \mathbf{g}$ is defined as $g_i = \sum_c A_c e^{j\frac{\pi}{N}(i+(N-1)\tau_c)} \frac{sin(\pi\tau_c)}{sin(\frac{\pi}{N}(\tau_c - i))}$, and $\mathbf{n} = [n_0\ n_1\ \cdots\ n_{N-1}]$ is the frequency-domain representation of complex Gaussian noise.

To recover the data transmission the receiver must estimate the channel response $\mathbf{g}$ or its frequency-domain equivalent $\mathbf{h} = \mathbf{gW}$. A number of channel estimation techniques have been considered for OFDM systems, including the minimum mean-squared error (MMSE) estimator, and the least-squares (LS) estimator. These estimators, with some improvements, are discussed in [59]. A discussion of particular channel estimation techniques is beyond the scope of this chapter, therefore without loss of generality, we use the least-square (LS) channel estimator [59] in Section 3.6 and Section 3.7 due its widespread adoption in OFDM systems and low computational complexity.

We now augment the OFDM transmission system with our extrinsic fingerprint function $\mathbf{f} = [f_0\ f_1\ \cdots\ f_{N-1}]^T$, and its matrix equivalent $\mathbf{F}^{(N\times N)} = \text{diag}\left([f_0\ f_1\ \cdots\ f_{N-1}]\right)$. The transmitted OFDM symbol of (3.2) with fingerprinting function applied after the

modulating IDFT matrix becomes

$$\mathbf{y} = \mathbf{gWFX} + \mathbf{n} \;=\; \mathbf{u}_{\mathbf{x},\mathbf{f},\mathbf{g}} + \mathbf{n}, \tag{3.3}$$

where $\mathbf{u}_{\mathbf{x},\mathbf{f},\mathbf{g}}$ is the received noiseless OFDM transmission.

To facilitate channel estimation, we use pilot-aided channel estimation, where pilot signals are periodically embedded into the transmitted signal. To represent periodic preamble and pilot signals, we expand the OFDM symbol transmission system above (3.3) to a block-based system consisting of M consecutive OFDM symbol vectors in time. The resulting Time-Frequency Block (TFB) received by the receiver at time index $t = kM$ is represented by the matrix $\mathbf{U}[k] \in \mathcal{C}^{(N\times M)}$, where each column of $\mathbf{U}[k]$ is an OFDM symbol vector $\mathbf{u}^m_{\mathbf{x},\mathbf{f},\mathbf{g}}$ received at time $(k-1)M + m$, $m = 0, 1, \ldots, M-1$. The TFB received at time index k becomes $\mathbf{Y}[k]$, i.e.

$$\mathbf{Y}[k] = \mathbf{U}[k] + \mathbf{N}[k], \tag{3.4}$$

with

$$\mathbf{U}[k] = \left[\left\{\mathbf{u}^0_{\mathbf{x},\mathbf{f},\mathbf{g}}\right\}^T \;\; \left\{\mathbf{u}^1_{\mathbf{x},\mathbf{f},\mathbf{g}}\right\}^T \;\; \ldots \;\; \left\{\mathbf{u}^{M-1}_{\mathbf{x},\mathbf{f},\mathbf{g}}\right\}^T \right], \tag{3.5}$$

and $\mathbf{N}[k] = \left[\mathbf{n}_0^T \; \mathbf{n}_1^T \; \ldots \; \mathbf{n}_{M-1}^T\right]_k$ are the noise vectors from (3.2).

If we assume the elements of $\mathbf{g}$ to be independent Rayleigh block-stationary and quasi-static, then $\mathbf{g}$ in (3.3) remains constant over the entire TFB for a total of M symbols. Similarly, if the fingerprinting function is also designed to be block-stationary, then (3.4) can be written as

$$\mathbf{Y}[k] = \mathbf{H}[k]\mathbf{F}[k]\mathbf{X}[k] + \mathbf{N}[k], \tag{3.6}$$

where $\mathbf{H}[k] = \mathrm{diag}\,(\mathbf{h}[k])$, $\mathbf{F}[k] = \mathrm{diag}\,(\mathbf{f}[k])$, $\mathbf{f}[k]$ is the fingerprinting function applied

to the entire TFB, and $\mathbf{g}[k]$ is the block-stationary channel response experienced by the received TFB, $\mathbf{Y}[k]$.

We construct $\mathbf{X}[k]$ as a composite signal composed of two components: the user-data signal and the embedded preamble and pilot signals used for channel estimation and equalization. Such a scheme enables the channel-like fingerprints to be detected using the known pilot signals. The frame preamble occupying $M - L$ time slots is followed by a section containing user-data symbols mixed with pilot signals occupying the remaining L time slots [50]. The TFB to be transmitted, augmented with embedded pilots signals becomes

$$\mathbf{X}[k] = \mathbf{D}[k]\mathbf{A} + \mathbf{P}, \tag{3.7}$$

where $\mathbf{D}[k] \in \mathcal{C}^{(N\times L)}$ is the TFB data matrix, $\mathbf{A} \in \mathcal{R}^{(L\times M)}$ is the data-projection matrix, and $\mathbf{P} \in \mathcal{R}^{(N\times M)}$ is the pilot signal matrix. The data-projection and pilot matrix satisfy the following properties:

$$\mathbf{A}\mathbf{P}^T = \mathbf{0} \in \mathcal{R}^{(L\times N)}, \quad \mathbf{A}\mathbf{A}^T = \mathbf{I} \in \mathcal{R}^{(L\times L)}, \quad \mathbf{P}\mathbf{P}^T = \mathbf{I} \in \mathcal{R}^{(N\times N)}. \tag{3.8}$$

The properties (3.8) of the data-projection matrix $\mathbf{A}$ and pilot matrix $\mathbf{P}$ essentially allow $\mathbf{A}$ to project the data component $\mathbf{D}[k]$ onto the orthogonal subspace of the pilot matrix $\mathbf{P}$, allowing for signal demodulation by means of a Maximum Likelihood (ML) receiver. In the simulation Section, 3.7, we will use a Time-Multiplexed (TM) single symbol preamble structure [50], which is given as

$$\mathbf{A} = \left[\mathbf{0}^{(L\times 1)}; \mathbf{I}^{(L\times M-1)}\right] \qquad \mathbf{P} = \left[\mathbf{I}^{(N\times 1)}; \mathbf{0}^{(N\times M-1)}\right], \tag{3.9}$$

where $\mathbf{I}^{(\cdot)}$ and $\mathbf{0}^{(\cdot)}$ are the identity matrix and the zero matrix, respectively, with dimensionality denoted in the superscript $(\cdot)$.

The expanded form of the TFB signal at the receiver (3.4), using (3.3), (3.6), and (3.7), becomes

$$\mathbf{Y}[k] = \mathbf{Q}[k](\mathbf{D}[k]\mathbf{A} + \mathbf{P}) + \mathbf{N}[k], \tag{3.10}$$

where $\mathbf{Q}[k] = \mathrm{diag}\,(\mathbf{H}[k]\mathbf{F}[k]) = \mathrm{diag}\,(\mathbf{q}[k])$ is the aggregate channel-like distortion.

In [26] delineation of the intrinsic channel response and the extrinsic fingerprint signal was accomplished through an even-odd time-division delivery (TDD) scheme that implemented a differential modulation scheme for transmission of the fingerprint signal. In the even-odd transmission scheme the fingerprint transmission $\mathbf{F}[k]$ is omitted during even block transmissions, i.e. $k = [0, 2, 4, \dots]$, and is present during odd block transmissions, i.e. $k = [1, 3, 5, \dots]$. To sound the physical channel, the fingerprint function is omitted by simply replacing $\mathbf{F}[k]$ with the identity matrix. This scheme allows for periodic sampling of intrinsic channel distortions, denoted $\mathbf{g}[k]$ in this work, when a generalized pilot embedding scheme, (3.7), (3.8), and (3.10), is employed. In Section 3.5 we will consider a similar fingerprint transmission scheme that enables periodic channel-sounding.

In this paper, we aim to design channel-like fingerprint signaling schemes that result in minimal degradation to the primary data signal, $\mathbf{D}[k]$, and the detection thereof, when the fingerprint is present.

3.4 Signal Recovery

In this section we briefly describe how the primary signal is recovered, using traditional unmodified channel estimation algorithms, when the fingerprint signal is present.

Upon receiving the TFB, the first step for both aware and unaware receivers is channel estimation. The channel estimation problem is to extract and estimate channel distortions in the received signal (3.10) for performing channel equalization and further recovering $\mathbf{D}[k]$. By post-multiplying both sides of (3.10) by $\mathbf{P}^H\mathbf{W}^H$ and using the properties of (3.8), an estimate of the aggregate channel response $\mathbf{Q}[k] = (\mathbf{H}[k]\mathbf{F}[k])$ of (3.26) can be obtained from $\mathbf{Y}[k]$, i.e.

$$\begin{aligned}\hat{\mathbf{Q}}[k] &= \mathbf{Y}[k]\mathbf{P}^H\mathbf{W}^H \\ &= \mathbf{H}[k]\mathbf{F}[k]\mathbf{W}(\mathbf{D}[k]\mathbf{A}+\mathbf{P})\mathbf{P}^H\mathbf{W}^H + \mathbf{N}[k]\mathbf{P}^H\mathbf{W}^H \\ &= \mathbf{Q}[k] + \mathbf{N}[k]\mathbf{P}^H\mathbf{W}^H.\end{aligned} \tag{3.11}$$

If the fingerprinting function is replaced by the ones vector, a channel-sounding process that will be described in Section 3.5, $\mathbf{F}[k]$ from (3.11) becomes the identity matrix and the intrinsic channel response estimate $\hat{\mathbf{H}}[k]$ can trivially be recovered from $\mathbf{Y}[k]$, i.e.

$$\begin{aligned}\hat{\mathbf{H}}[k] &= \mathbf{Y}[k]\mathbf{P}^H\mathbf{W}^H \\ &= \mathbf{H}[k]\mathbf{W}(\mathbf{D}[k]\mathbf{A}+\mathbf{P})\mathbf{P}^H\mathbf{W}^H + \mathbf{N}[k]\mathbf{P}^H\mathbf{W}^H \\ &= \mathbf{H}[k] + \mathbf{N}[k]\mathbf{P}^H\mathbf{W}^H.\end{aligned} \tag{3.12}$$

In (3.11) and (3.12), $\mathbf{N}[k]\mathbf{P}^H\mathbf{W}^H$ is the projection of the noise vector $\mathbf{n}$ from (3.2) onto the pilot signals embedded in the TFB, and represents noise in the channel estimate.

The next step preformed by the receiver is the recovery of the transmitted primary-signal data $\mathbf{D}[k]$. By post-multiplying both sides of (3.10) by $\mathbf{A}^H$ and using the properties (3.8), the data signal $\mathbf{D}[k]$ can be extracted from the received signal $\mathbf{Y}[k]$, i.e.

$$\begin{aligned}\mathbf{Y}[k]\mathbf{A}^H &= \mathbf{H}[k]\mathbf{F}[k]\mathbf{W}\left(\mathbf{D}[k]\mathbf{A}+\mathbf{P}\right)\mathbf{A}^H + \mathbf{N}[k]\mathbf{A}^H \\ &= \left(\mathbf{H}[k]\mathbf{F}[k]\right)\mathbf{W}\mathbf{D}[k] + \mathbf{N}[k]\mathbf{A}^H.\end{aligned} \tag{3.13}$$

An estimate for the aggregate channel response $\mathbf{Q}[l] = (\mathbf{H}[l]\mathbf{F}[l])$ is demonstrated in (3.11), and thus the data signal may be recovered by pre-multiplying (3.13) by the inverse of the aggregate channel estimate and taking the FFT, i.e. $\mathbf{W}^H\mathbf{Q}^{-1}[k]$. For the block-stationary channel case, both $\mathbf{H}[l]$ and $\mathbf{F}[l]$ are square, diagonal, non-negative matrices of size $\mathcal{C}^{(N\times N)}$ and thus the inverse of $\mathbf{Q}[l]$ exists and is simply the element-wise inverse of its diagonal elements.

While equations (3.11) and (3.12) demonstrate that the channel estimate can be recovered using generalized pilot embedding schemes, in Sections 3.6 and 3.7 we will specifically consider the LS channel estimator in experiment and simulation due to its widespread use. When using LS channel estimation, the columns of the embedded pilot signals of $\mathbf{P}$ in (3.7) become the training data for the estimator.

3.5 Subspace Extrinsic Channel-Like Fingerprinting

In this section, we will first derive the capacity of the extrinsic fingerprint under time-varying channels, and then demonstrate the maximal-capacity fingerprint embedding scheme using subspace decomposition of the channel information. We will also introduce the fingerprint recovery process at the receiver's side, as well as example fingerprint design. In this section we consider the subspace decomposition of channel state information and describe how a fingerprint message can be embedded in the noise subspace of these channel estimates. While a perfectly embedded fingerprint will occupy only the noise subspace of the received fingerprint message, we will use the framework

presented in this section to create a number of sub-optimal fingerprint designs that leverage various amounts of CSI.

3.5.1 Time-Varying Channel Model

We now briefly consider the capacity of an extrinsic message, signaled via the fingerprinting function $\mathbf{f}[k]$ and decoded through consideration of historical CSI obtained via pilot-aided channel estimation in time-varying channels. In this motivating example, we leverage the fact that the time-variant noise process modeling the intrinsic channel response, (3.15) and (3.17), is a colored Gaussian noise process. By a *water-filling* argument we surmise that an extrinsic fingerprint process, applied by way of the fingerprinting function $\mathbf{f}[k]$, may be designed as a Gaussian process with a power spectrum that is large at frequencies where the intrinsic channel noise spectrum is small.

We now briefly describe the channel model of our intrinsic time-varying channel response $\mathbf{g}[k]$, or equivalently, its frequency-domain representation $\mathbf{h}[k] = \mathrm{diag}\left(\mathbf{G}[k]\mathbf{W}\right)$ by the time-variant OFDM channel model used in [64] where the block-stationary channel response vector $\mathbf{h}[k]$ has scalar complex gain elements $h_j[k],\ j = 0, \ldots, N-1$ corresponding to the N independent OFDM subcarriers. Each element $h_j[k]$ is the summation of three model components:

- A fixed time-invariant channel gain denoted $\bar{h}_j = E\left[h_j[k]\right],\ j = 0, \ldots, N-1$
- A zero-mean time-variant channel gain component denoted $\mu_j[k],\ j = 0, \ldots, N-1$
- A zero-mean receiver noise component denoted $n_j[k],\ j = 0, \ldots, N-1$,

where $\bar{h}_j$ is the mean of the random variable $h_j[k]$, and $\mu_j[k]$ is a linear system. Thus, $h_j[k]$ becomes

$$h_j[k] = \bar{h}_j + \mu_j[k] + n_j[k], \;\; j = 0, \ldots, M-1. \tag{3.14}$$

While in general the channel gain means $\bar{h}_j$ will vary in time, for the sake of exposition we will assume that this component will remain stationary between the interval when the intrinsic channel is sounded and the fingerprint is detected. According to the block-stationary channel model, $h_j[k]$ remains constant over $i = 0, \ldots, M-1$ symbols, which allows $\mathbf{H}[k]$ to be defined as $\mathbf{H}[k] = \text{diag}\left([h_0[k], h_1[k], \ldots, h_{N-1}[k]]\right)$. To model the time-varying channel gain elements, $\mu_j[k]$, we adopt the *wide-sense stationary uncorrelated scattering* (WSSUS) temporal fading model [8], where the complex magnitudes in (3.1) are zero-mean, i.e. $E\left[A_c(t)\right] = 0$, and $\Delta\tau = 1/W$ is the carrier spacing of the DFT. The frequency response of the variable channel gains is the Fourier transform of $g(t, \tau)$, i.e.

$$\mu_j[r] = \mathcal{F}\left\{g(t,\tau)\right\}|_{t=rT, f=f_o-W/2+j\Delta f} = \sum_{c=0}^{\infty} A_c[r] e^{-j2\pi(f_o - W/2 + j\Delta f)c\Delta}, \tag{3.15}$$

where $A_c[r] = A_c(rT)$ is the discrete sample of the multipath item at time rT. In [64], a one-sided exponential distribution is considered for modeling the power delay spectrum of $A_c[r]$, i.e.,

$$P_\tau[c] = \text{Var}\left[A_c[r]\right] = \sigma_T^2 (1 - e^{\gamma\Delta\tau}) e^{-\gamma\Delta\tau c}, \tag{3.16}$$

where $\gamma = 2\pi B_c$, B_c is the coherence bandwidth of the time-variant noise, and σ_T^2 is the average power of $A_l[r]$ over all taps. Each tap gain is modeled as an order 1 autoregressive model (AR-1) characterizing the process $A_c[r]$ as

$$A_c[r] = aA_c[r-1] + \sqrt{(1-a^2)P_\tau[c]}u_c[r], \tag{3.17}$$

where the AR coefficient a controls the temporal correlation between two gains, A_c, separated in time by T for a given delay index c, and the independent noise process of (3.17) is represented by $u_c[r] \sim \mathcal{CN}(0,1),\ u_c[r] \perp A_c[r-1]$. Therefore σ_T is the variance of the driving source of the AR-1 process that is used to model each channel gain. When σ_T increases, the amplitude of the fluctuation of each channel gain in the model increases, and changing channel conditions become more violent. We note that (3.17) models the gains of each multipath component as a *colored Gaussian noise process*, an important property that will be considered in the next section for fingerprint analysis.

In essence, an optimal extrinsic fingerprint signaling scheme, embodied by $\mathbf{f}[k]$, would adapt to the intrinsic channel response $\mathbf{g}[k]$, and by means of spectral water-filling these two processes work in consort to produce the received band-limited signal, $\mathbf{Y}[k]$, given in (3.10). Traditional transmission precoding by way of water-filling methods typically strive to increase the capacity of the primary-signal, represented by $\mathbf{D}[k]$, using partial CSI at the transmitter. Instead, in this work we consider the case where any additional capacity obtained via water-filling is provisioned to the fingerprint signal which operates independently of the primary-signal.

In [64] it was demonstrated that the differential test statistic for the intrinsic channel noise process has the relationship

$$\mathbf{g}[k] - \mathbf{g}[k-1] \sim \mathcal{CN}(\mathbf{0}, \mathbf{R}), \tag{3.18}$$

with covariance matrix

$$\mathbf{R} = \mathrm{Cov}\left[\mathbf{g}[k] - \mathbf{g}[k-1]\right] = [r(m-n)]_{m\times n},\ m,n = 0,\ldots,N-1, \tag{3.19}$$

and has an autocorrelation function $r(\tau)$ given as

$$r(\tau) = \begin{cases} 2(1-a)\sigma_T^2 + 2\sigma_N^2 & \tau = 0 \\ \frac{2\sigma_T^2(1-a)\left(1-e^{-2\pi B_c/W}\right)}{1-e^{-2\pi B_c/W - j2\tau/N}} & \tau = -N+1, \ldots, N-1. \end{cases} \tag{3.20}$$

From (3.20) we can derive the noise spectrum of the intrinsic channel gains $\mathbf{g}[k]$ in time, which is given as the Fourier transform of the autocorrelation function, i.e.

$$\begin{aligned} S(f) &= \int_{-\infty}^{\infty} r(\tau) e^{-j2\pi f\tau} d\tau \\ &= \int_{-\infty}^{\infty} \frac{2\sigma_T^2(1-a)\left(1-e^{-2\pi B_c/W}\right) e^{-j2\pi f\tau}}{1-e^{-2\pi B_c/W - j2\tau/N}} d\tau \\ &= \gamma e^{-\pi f}, \end{aligned} \tag{3.21}$$

with constants ξ, γ, ρ

$$\xi = e^{-2\pi B_c/W}\ ,\ \gamma = 2\sigma_T^2(1-a)\,(1-\xi)\int_{-\infty}^{\infty} \frac{1}{e^{j2\tau}-\rho} d\tau\ ,\ \rho = \xi e^{-1/N}. \tag{3.22}$$

We see that the power spectrum of this channel model is a decaying complex exponential with constants π and γ. It can be shown that the water-filling capacity C of an additive Gaussian noise channel with power spectrum $S(f)$ is simply [16]

$$C = \int_{-\pi}^{\pi} \frac{1}{2} \log \left[1 + \frac{(\nu - S(f))^+}{S(f)}\right] df \tag{3.23}$$

where the spectral *water-level* ν is selected such that $\int (\nu - S(f))^+ df = P$.

3.5.2 Subspace Decomposition of Channel Information

We consider a sequence of P previous block-stationary channel estimate vectors $\hat{\mathbf{h}}[k]$, obtained via channel estimation and arranged as column vectors in a matrix $\hat{\mathbb{H}}[k] \in \mathcal{C}^{(N\times P)},\ N \geq P$, i.e.

$$\hat{\mathbb{H}}[k] = \left[\hat{\mathbf{h}}^T[k-P+1]\ \hat{\mathbf{h}}^T[k-P+2]\ \ldots\ \hat{\mathbf{h}}^T[k-1]\right]. \tag{3.24}$$

We note from this point forward that $\hat{\mathbb{H}}[k]$ and any derivations are based on the estimate of the true channel gain $\mathbb{H}[k] = \left[\mathbf{h}^T[k-P+1] \ \ldots \ \mathbf{h}^T[k-1]\right]$ since both the receiver and transmitter only have the information of the estimated channel gain. We consider the case where the fingerprinting function is designed using a block-based scheme such that $\mathbf{f}[k]$ in (3.10) is applied by the transmitter to a sequence of TFBs as a matrix denoted $\mathbb{F}[k]$. We select the length of the fingerprint block to be P, so that manipulations of one fingerprint block over P consecutive TFBs will coincide with the evolution of CSI at the receiver, i.e. $\hat{\mathbb{H}}[k]$ which is also of length P TFBs. Let the block-based fingerprinting function, $\mathbb{F}[k] \in \mathcal{C}^{(N\times P)}$, be the matrix-representation of P fingerprinting functions $\mathbf{f}^T[k]$ applied by the transmitter, such that

$$\mathbb{F}[k] = \left[\mathbf{f}^T[k-P+1]\, \mathbf{f}^T[k-P+2] \ \ldots \ \mathbf{f}^T[k-1]\right]. \tag{3.25}$$

For the fingerprinting transmission scheme described by (3.10), we arrange the aggregate block-stationary channel estimate vectors $\hat{\mathbf{q}}[k]$ that are observed by the receiver as columns in $\hat{\mathbb{Q}}[k]$, which are related to $\hat{\mathbb{H}}[k]$ and $\mathbb{F}[k]$ via the Hadamard product, i.e.

$$\hat{\mathbb{Q}}[k] = \hat{\mathbb{H}}[k] \circ \mathbb{F}[k], \tag{3.26}$$

where $(\circ)$ represents the Hadamard product. The intrinsic time-varying channel measurement $\hat{\mathbb{H}}[k]$ is easily obtained by omitting the fingerprint component and replacing the fingerprinting function $\mathbf{f}[k]$ with the ones vector $\mathbf{1}^{(1\times N)}$. Thus $\mathbb{F}[k]$ in (3.26) becomes the identity matrix for the Hadamard product, which is the ones matrix $\mathbf{1}^{(N\times P)}$. We denote this process as the *channel-sounding phase* with accompanying channel-sounding fingerprint function $\mathbb{F}_{snd}[k]$. The aggregate distortion perceived by the receiver during the

channel-sounding phase, denoted $\hat{\mathbb{Q}}_{snd}[k]$, is simply

$$\hat{\mathbb{Q}}_{snd}[k] = \hat{\mathbb{H}}[k] \circ \mathbb{F}_{snd}[k] = \hat{\mathbb{H}}[k] \circ \mathbf{1}^{(N \times P)} = \hat{\mathbb{H}}[k]. \tag{3.27}$$

CSI used by the transmitter must first be estimated by the receiver and then conveyed to the transmitter as feedback, resulting in a delay. Thus, we denote CSI obtained by the transmitter during the channel-sounding phase as $\hat{\mathbb{Q}}_{snd}[l-\epsilon] = \hat{\mathbb{H}}[l-\epsilon]$, where ϵ is the number of OFDM symbols of delay experienced by the channel-sounding CSI, as received by the transmitter, and the current fingerprint transmission at time $l = \lfloor \frac{k}{P} \rfloor$. The diagonalization of $\hat{\mathbb{Q}}_{snd}[l-\epsilon]$ in (3.27) via singular-value decomposition (SVD) for the case $N \geq P$ yields

$$\hat{\mathbb{H}}^T[l-\epsilon] = \hat{\mathbb{Q}}^T_{snd}[l-\epsilon] = \mathbb{U}[l-\epsilon]\mathbb{S}[l-\epsilon]\mathbb{V}^H[l-\epsilon], \tag{3.28}$$

where $\mathbb{U}[l-\epsilon] \in \mathcal{C}^{(P \times P)}$ is the *left* unitary matrix of the decomposition with orthonormal columns representing the *left* singular-vectors of $\hat{\mathbb{H}}[l-\epsilon]$, and $\mathbb{V}[l-\epsilon] \in \mathcal{C}^{(N \times N)}$ contains the *right* orthonormal columns of the singular-vectors of $\hat{\mathbb{H}}[l-\epsilon]$, and $\mathbb{S}[l-\epsilon] \in \mathcal{R}^{(P \times N)}$ contains the diagonalized eigenvalues of the decomposition. The column vectors of $\mathbb{U}[l-\epsilon]$ are the projections of the eigenvectors of $\hat{\mathbb{H}}[l-\epsilon]$ in the time dimension for the previous P TFBs, while $\mathbb{V}[l-\epsilon]$ are the projections of the eigenvectors of $\hat{\mathbb{H}}[l-\epsilon]$ in the frequency dimension.

According to our time-varying channel model (3.14), we surmise that linearly correlated time-varying channel components $\bar{h}$ and $\mu_j[k]$ can be separated from the uncorrelated noise components $N_j[k]$ of (3.14) via subspace decomposition of (3.28), if P is large enough [19]. Hence, $\mathbb{U}[l-\epsilon]$ and $\mathbb{V}[l-\epsilon]$ form a basis describing the linear dependencies in $\hat{\mathbb{H}}[l-\epsilon]$ in both time and frequency. We consider the case where P is selected

to be large enough so that the decomposition of $\hat{\mathbb{H}}[l-\epsilon]$ is over-determined. Via subspace decomposition we bifurcate each matrix of (3.28) into two subspaces:

1. A correlated time-varying *signal subspace* characterizing the linear dependencies in $\hat{\mathbb{H}}[l-\epsilon]$ spanning both time and frequency, thus forming a temporal-spectral model for the linearly correlated components present in previous CSI estimates.

2. An uncorrelated *noise subspace*, also spanning $\hat{\mathbb{H}}[l-\epsilon]$ in both time and frequency, and thus characterizing the noise component present in previous CSI.

The correlated signal subspace will also be referred to as the *intrinsic* distortion subspace, as this subspace characterizes the principle linear relationships between consecutive channel estimates intrinsic to the linear time-varying fading channel. The bifurcation of $\mathbb{U}[l-\epsilon]$, $\mathbb{S}[l-\epsilon]$, and $\mathbb{V}[l-\epsilon]$ into principal components and noise, according to a model parameter p, becomes

$$\begin{aligned}\hat{\mathbb{H}}^T[l-\epsilon] &= \left\{\mathbb{U}\mathbb{S}\mathbb{V}^H\right\}_{l-\epsilon} \\ &= \left[\mathbb{U}_1^{(P\times p)} \quad \mathbb{U}_2^{(P\times(P-p))},\right]_{l-\epsilon} \left[\mathbf{\Lambda}^{(P\times P)} \quad \mathbf{0}^{(P\times(N-P))},\right]_{l-\epsilon} \left[\mathbb{V}_1^{(N\times p)} \quad \mathbb{V}_2^{(N\times(N-p))},\right]_{l-\epsilon}^{H},\end{aligned} \tag{3.29}$$

where the columns of $\mathbb{U}_1$ and $\mathbb{V}_1$ represent the singular-vectors of the left and right unitary matrices $\mathbb{U}$ and $\mathbb{V}$, respectively, spanning the correlated time-varying signal subspace, $\mathbb{U}_2$ and $\mathbb{V}_2$ represent the singular-vectors of $\mathbb{U}$ and $\mathbb{V}$, respectively, spanning the noise subspace, and $\mathbf{\Lambda} \in \mathcal{R}^{(P\times P)} = \mathrm{diag}\left(\lambda_0,\ \lambda_1,\ \ldots,\ \lambda_{P-1}\right)$, are the eigenvalues $\lambda_i,\ i = 0, 1, \ldots, P-1$ of $\mathbb{S}$. The dimensionality of these subspaces, determined by p, should be chosen to be equal to the effective rank of $\hat{\mathbb{H}}[l-\epsilon]$, for optimal bifurcation of the

subspaces. That is, if we sort the eigenvectors of $\mathbb{S}[l-\epsilon]$ according to magnitude in descending order, i.e. $\lambda_0 \geq \lambda_i \geq \lambda_{P-1}, i = 1, \ldots, P-2$, the effective rank of $\mathbb{H}[l-\epsilon]$, and thus the optimal value for p, is equal to the number of eigenvalues of $\mathrm{diag}\,(\mathbf{\Lambda})$ that are not equal to σ_H^2, i.e.

$$\lambda_0 \geq \lambda_i \geq \lambda_j = \sigma_H^2,\ i = 1,\ \ldots, p-1,\ \ j = p,\ \ldots, P-1, \tag{3.30}$$

where σ_H^2 is the noise power of $\mathbf{N}[l-\epsilon]$ projected onto the pilot matrix $\mathbf{P}$. In practice, the eigenvalues of the noise subspace may not all be equal, making the estimation of the effective rank of $\hat{\mathbb{H}}[l-\epsilon]$ difficult. However, these values will be very close to σ_H^2 [40]. Additionally, the spectrum of the signal spanned by the noise subspace is orthogonal to the intrinsic channel disturbance spectrum. We will exploit this orthogonality property later in our fingerprinting designs to be discussed in Section 3.5.3. Various criteria have been proposed to estimate p in these cases [40], however this discussion is beyond the scope of this chapter. For the sake of exposition we will assume that p is perfectly selected to be the effective rank of $\mathbb{H}[l-\epsilon]$, $\lambda_j = \sigma_H^2,\ j = p,\ \ldots, P-1$, and note that improper estimation of p will result in degraded performance as the orthogonality between the two subspaces is degraded in this case.

For sufficiently stationary and non-trivial time-varying channels, the channel estimates $\hat{\mathbf{h}}[k-i],\ i = 1, \ldots, P+1$ obtained using embedded pilot signals are correlated in both frequency and in time. Thus, when P is properly selected, the resulting decomposition of $\mathbb{H}[l-\epsilon]$ in (3.28) will be over-determined and therefore $p < P$ and the size of both subspaces will be non-zero. For the case $p \geq P$, the system is under-determined and an

accurate delineation of both subspaces in (3.29) is not possible. Therefore, for the sake of exposition we consider here the case

$$1 \leq p < P,\ P \geq 2, \tag{3.31}$$

which is the most interesting case for our extrinsic fingerprinting method, as both the intrinsic time-varying channel distortion subspace and the noise subspace are of non-zero size.

3.5.3 Fingerprint Design using Subspace Modeling

We now describe how the fingerprinting function $\mathbf{F}[k]$ is embedded into the transmission to produce a time-frequency fingerprint that leverages knowledge of the channel, $\mathbf{H}[k]$.

According to a spectral water-filling observation that was presented in (3.23), the capacity of a sequence of channel estimate data, i.e. the amount of information conveyed by $\hat{\mathbb{Q}}[k]$ to the receiver using the embedded pilot signals $\mathbf{P}$ of (3.7) to drive $\hat{\mathbb{Q}}[k]$, can be maximized by introducing a Gaussian process with energy where the spectrum of $\mathbb{H}[k]$ is lowest. The transmitter can first estimate the noise subspace of $\mathbb{H}[k]$ by decomposing $\hat{\mathbb{H}}[k]$ as in (3.30), in which $\mathrm{diag}\,(\mathbf{\Lambda})$ denotes spectral magnitudes and the smallest eigenvalues represent the power spectrum of the noise subspace of $\hat{\mathbb{H}}[k]$. We introduce our fingerprinting signal using a *fingerprint overlay* onto these parts of the spectrum.

Hence, the fingerprint function should be in the noise subspace, and the transmitter has to ensure that the receiver can detect the fingerprint from the aggregate channel estimate $\hat{\mathbb{Q}}[k]$. To aid in analysis, we split the subspace decomposition of the channel

state information in possession by the transmitter, (3.29), into a summation of significant channel gains and noise, i.e.

$$\hat{\mathbb{H}}^T[l-\epsilon] = \hat{\mathbb{L}}[l-\epsilon] + \hat{\mathbb{N}}[l-\epsilon], \tag{3.32}$$

where $\hat{\mathbb{L}}[l-\epsilon]$ is an estimate of the significant intrinsic channel gains $\mathbb{L}[l-\epsilon]$ on the signal subspace, and $\hat{\mathbb{N}}[l-\epsilon]$ is an estimate of noise on the noise subspace $\mathbb{N}[l-\epsilon]$, of $\mathbb{H}[l-\epsilon]$. From this definition, the principle linear components of the intrinsic time-varying channel fading patterns are parameterized by the basis $\mathbb{L}[l]$. The definition of $\mathbb{L}[l-\epsilon]$ and $\mathbb{N}[l-\epsilon]$ according to (3.29) and (3.30) is

$$\hat{\mathbb{L}}[l-\epsilon] = \left[\mathbb{U}_1 \quad \mathbf{0}^{(P\times(P-p))}\right]_{l-\epsilon} \left[\mathrm{diag}\left(\boldsymbol{\lambda}^{(p)}, \mathbf{0}^{(P-p)}\right) \quad \mathbf{0}^{((N-p)\times P)}\right]_{l-\epsilon} \left[\begin{array}{c} \mathbb{V}_1^H \\ \mathbf{0}^{((N-p)\times N)} \end{array}\right]_{l-\epsilon}, \tag{3.33}$$

where $\boldsymbol{\lambda}[l-\epsilon]$ is defined as

$$\boldsymbol{\lambda}[l-\epsilon] = \left[\lambda_0 \; \lambda_1 \; \ldots \; \lambda_{p-1}\right]_{l-\epsilon}, \tag{3.34}$$

and

$$\hat{\mathbb{N}}[l-\epsilon] = \left[\mathbf{0}^{(P\times p)} \quad \mathbb{U}_2,\right]_{l-\epsilon} \left[\mathrm{diag}\left(\mathbf{0}^{(p)}, \boldsymbol{\sigma}^{(P-p)}\right) \quad \mathbf{0}^{(P\times(N-P))}\right]_{l-\epsilon} \left[\begin{array}{c} \mathbf{0}^{(p\times N)} \\ \mathbb{V}_2^H \end{array}\right]_{l-\epsilon}, \tag{3.35}$$

where $\boldsymbol{\sigma}[l-\epsilon]$ is defined as

$$\boldsymbol{\sigma}[l-\epsilon] = \left[\lambda_p \; \lambda_{p+1} \; \ldots \; \lambda_{P-1}\right]_{l-\epsilon}, \quad \lambda_p = \lambda_{p+1} = \cdots = \lambda_{P-1} = \sigma_H^2. \tag{3.36}$$

This low-rank modeling of intrinsic channel conditions by $\mathbb{L}[l]$ will help reduce feedback overhead when conveying CSI to the transmitter, as this feedback decreases system efficiency.

3.5.3.1 Subspace Fingerprint Design

The transmitter designs $\mathbb{F}[l]$, which will be recovered by the receiver using the estimate of the aggregate channel, $\hat{\mathbb{Q}}[l]$. Since $\hat{\mathbb{Q}}[l] = \mathbb{F}[l] \circ \hat{\mathbb{H}}[l]$, for the transmitter, designing $\mathbb{F}[l]$ is the same as designing $\hat{\mathbb{Q}}[l]$ if the transmitter has the current channel estimate, $\hat{\mathbb{H}}[l]$. However, the transmitter possesses only a delayed version of the channel estimate data, $\hat{\mathbb{H}}[l - \epsilon]$. If prediction of future channel state is not employed, the transmitter must approximate $\hat{\mathbb{H}}[l]$ using $\hat{\mathbb{H}}[l - \epsilon]$. Therefore, the transmitter will approximate the aggregate channel estimate $\hat{\mathbb{Q}}[l]$ at the receiver's side by $\hat{\mathbb{Q}}_{des}[l]$, where $\hat{\mathbb{Q}}_{des}[l] = \mathbb{F}[l] \circ \hat{\mathbb{H}}[l - \epsilon]$.

Similarly, we can decompose $\hat{\mathbb{Q}}_{des}[l]$ using a summation model

$$\hat{\mathbb{Q}}_{des}[l] = \mathbb{P}[l] + \mathbb{K}[l], \tag{3.37}$$

where $\mathbb{P}[l]$ is the the projection of $\hat{\mathbb{Q}}_{des}[l]$ onto the intrinsic channel subspace corresponding to $\hat{\mathbb{L}}[l]$ in (3.32), and $\mathbb{K}[l]$ is the extrinsic fingerprinting overlay matrix that we will design to overlay the noise component $\hat{\mathbb{N}}[l]$ in (3.32).

Consequently, the transmitter will design $\mathbb{F}[l]$ according to $\hat{\mathbb{Q}}_{des}[l]$, such that

$$\mathbb{F}[l] = (\mathbb{P}[l] + \mathbb{K}[l]) \circ \left(\hat{\mathbb{L}}[l - \epsilon] + \hat{\mathbb{N}}[l - \epsilon]\right)^{(-1)}, \tag{3.38}$$

where $(\cdot)^{(-1)}$ is the Hadamard inverse operation.

3.5.3.2 Fingerprint Extraction

The aggregate channel estimate $\hat{\mathbb{Q}}[l]$ that the receiver will obtain can be formulated by substituting (3.38) into (3.26), i.e. $\hat{\mathbb{Q}}[l]$, becomes

$$\begin{aligned}\hat{\mathbb{Q}}[l] &= \hat{\mathbb{H}}[l] \circ (\mathbb{P}[l] + \mathbb{K}[l]) \circ \left(\hat{\mathbb{L}}[l-\epsilon] + \hat{\mathbb{N}}[l-\epsilon]\right)^{(-1)} \\ &= \left(\hat{\mathbb{L}}[l] \circ \mathbb{P}[l] + \hat{\mathbb{L}}[l] \circ \mathbb{K}[l] + \hat{\mathbb{N}}[l] \circ \mathbb{P}[l] + \hat{\mathbb{N}}[l] \circ \mathbb{K}[l]\right) \circ \left(\hat{\mathbb{L}}[l-\epsilon] + \hat{\mathbb{N}}[l-\epsilon]\right)^{(-1)}.\end{aligned} \tag{3.39}$$

We will now show that $\hat{\mathbb{Q}}[l]$ is an unbiased estimate of $\hat{\mathbb{Q}}_{des}[l] = \hat{\mathbb{P}}[l] + \hat{\mathbb{K}}[l]$ enabling recovery of the fingerprint $\hat{\mathbb{F}}[l]$ by the receiver, without bias. Equivalently, we want to show that $E\left[\hat{\mathbb{Q}}[l]\right] = \mathbb{P}[l] + \mathbb{K}[l]$.

Assuming that $\hat{\mathbb{L}}[l]$ and $\hat{\mathbb{L}}[l-\epsilon]$ are unbiased estimates of their respective channel gain components, i.e.

$$E\left[\hat{\mathbb{L}}[l]\right] = \mathbb{L}[l] \text{ and } E\left[\hat{\mathbb{L}}[l-\epsilon]\right] = \mathbb{L}[l-\epsilon], \tag{3.40}$$

then the expectation of (3.39) yields

$$\begin{aligned}E\left[\hat{\mathbb{Q}}[l]\right] = \Big(\mathbb{L}[l] \circ \mathbb{P}[l] + \mathbb{L}[l] \circ \mathbb{K}[l] + E\left[\hat{\mathbb{N}}[l]\right] \circ \mathbb{P}[l] + E\left[\hat{\mathbb{N}}[l]\right] \circ \mathbb{K}[l]\Big) \circ \\ \left(\hat{\mathbb{L}}[l-\epsilon] + E\left[\hat{\mathbb{N}}[l-\epsilon]\right]\right)^{(-1)}.\end{aligned} \tag{3.41}$$

In the derivation of (3.41), we recall that $(\circ)$ is the Hadamard product, therefore regular matrix multiplication and inversion is not used in this result.

We recall that $\hat{\mathbb{N}}[l]$ is the projection of $\mathbf{N}[k]$ on the noise subspace of channel estimates $\hat{\mathbb{H}}[l]$. As $\mathbf{N}[k]$ is a matrix of zero-mean Gaussian random variables and the basis of the noise subspace is formed from Gaussian random variables $\mathbf{N}[k]$ projecting on the pilot signals $\mathbf{P}$, then $\hat{\mathbb{N}}[l]$ is also Gaussian with each element having zero mean, i.e.

$E\left[\hat{\mathbb{N}}[l]\right] = \mathbf{0}$. The elements of $\mathbf{N}[k]$ are also i.i.d. Gaussian, thus $E\left[\hat{\mathbb{N}}[l-\epsilon]\right]$ is also zero-mean Gaussian and $\hat{\mathbb{N}}[l-\epsilon]$ is uncorrelated with $\hat{\mathbb{N}}[l]$, therefore by a similar argument $E\left[\hat{\mathbb{N}}[l-\epsilon]\right] = \mathbf{0}$, and (3.41) becomes

$$E\left[\hat{\mathbb{Q}}[l]\right] = (\mathbb{P}[l] + \mathbb{K}[l]) \circ \mathbb{L}[l] \circ \mathbb{L}[l-\epsilon]^{(-1)}. \tag{3.42}$$

The above equation demonstrates that if the intrinsic channel is stationary over ϵ blocks, i.e., $\mathbb{L}[l] = \mathbb{L}[l-\epsilon]$, then the aggregate channel estimate at the receiver's side $\hat{\mathbb{Q}}[l]$ is an unbiased estimate of the information that the transmitter conveyed, $\hat{\mathbb{Q}}_{des}[l]$.

Moreover, to obtain an estimate for only the extrinsic fingerprinting overlay $\mathbb{K}[l]$ from $\hat{\mathbb{Q}}[l]$, in (3.42) we immediately see that our estimate $\hat{\mathbb{Q}}[l]$ will be biased by $\mathbb{P}[l]$. By a water-filling argument which was discussed in Section 3.5.1, the extrinsic fingerprinting signal should contribute energy to the noise subspace of $\mathbb{H}[l]$ to maximize the information conveyed by $\mathbb{H}[l]$ to the receiver, as this basis represents the spectral elements of $\mathbb{H}[l]$ with the lowest energy. Thus, we design $\mathbb{F}[l]$ to contribute only to the noise subspace while leaving the intrinsic channel subspace unperturbed by setting $\mathbb{P}[l]$ to be the identity matrix for the Hadamard product, i.e.

$$\mathbb{P}[l] = \mathbf{1}. \tag{3.43}$$

With (3.43), (3.42) becomes

$$\begin{aligned} E\left[\hat{\mathbb{Q}}[l]\right] &= \mathbb{L}[l] \circ \mathbb{L}^{(-1)}[l-\epsilon] + \mathbb{L}[l] \circ \mathbb{K}[l] \circ \mathbb{L}^{(-1)}[l-\epsilon] \\ &= (\mathbf{1} + \mathbb{K}[l]) \circ \mathbb{M}[l], \end{aligned} \tag{3.44}$$

where we introduce the definition

$$\mathbb{M}[l] = \mathbb{L}[l] \circ \mathbb{L}^{(-1)}[l-\epsilon]. \tag{3.45}$$

In (3.45), $\mathbb{M}[l]$ is the intrinsic-channel model mismatch error in estimating $\mathbb{K}[l]$, and is the Hadamard product between $\mathbb{L}[l]$ and the Hadamard inverse of the previous intrinsic channel estimate, $\mathbb{L}^{(-1)}[l-\epsilon]$. From this definition, the model error mismatch matrix, $\mathbb{E}[l]$, is simply $\mathbb{E}[l] = \mathbf{1}^{(N\times P)} - \mathbb{M}[l]$.

From (3.45) we readily see that when the low-rank subspace approximation of $\mathbb{H}[l]$, $\mathbb{L}[l-\epsilon]$, is a perfect match of the low-rank approximation of the current channel conditions $\mathbb{L}[l]$, then $\mathbb{M}[l] = \mathbf{1}$, and by removing the bias introduced by (3.43), (3.44) becomes

$$E\left[\hat{\mathbb{Q}}[l]\right] - \mathbf{1} = \mathbb{K}[l], \tag{3.46}$$

thus an unbiased estimate for $\mathbb{K}[l]$ can be obtained from $\hat{\mathbb{Q}}[l]$. When the error matrix $\mathbb{E}[l]$ has non-zero elements, additional model mismatch error will result in degraded performance when detecting the fingerprint signal.

Using (3.43), the design of the fingerprinting function $\mathbb{F}[l]$ from (3.38) becomes simply

$$\mathbb{F}[l] = \mathbb{K}[l] \circ \mathbb{L}^{(-1)}[l-\epsilon]. \tag{3.47}$$

We note from (3.47) that either the transmitter or the receiver may apply $\mathbb{L}^{(-1)}[l-\epsilon]$ before recovering $\mathbb{K}[l]$ as the Hadamard product is commutative. The case where $\mathbb{L}^{(-1)}[l-\epsilon]$ is applied by the transmitter is analogous to linear OFDM block precoding [62], for the purpose of channel fade mitigation. If we assume that the receiver has memory and can store $\mathbb{L}^{(-1)}[l-\epsilon]$ for future computation, this would allow the receiver to preform this computation, eliminating the need to transmit $\mathbb{L}^{(-1)}[l-\epsilon]$ to the transmitter, thus decreasing the amount of CSI feedback required and reducing overhead.

3.5.4 Subspace Fingerprinting Overlays

We now discuss a methodology for designing the extrinsic fingerprint overlay $\mathbb{K}[l]$ that will allow the authentication signal to overlay the noise subspace $\mathbb{N}[l]$ of $\mathbb{H}[l]$. Additionally, we demonstrate how $\mathbb{K}[l]$ can be used to modulate the extrinsic fingerprinting signal.

Similar to the definition of $\hat{\mathbb{L}}[l]$ in (3.33) and $\hat{\mathbb{N}}[l]$ in (3.35), according to (3.29) and (3.30) we define the extrinsic fingerprinting overlay matrix $\mathbb{K}[l]$ as

$$\begin{aligned}\mathbb{K}^T[l] =& \mathbb{U}_{l-\epsilon}\mathbb{S}_l\mathbb{V}^H_{l-\epsilon} \\ =& \begin{bmatrix}\mathbf{0}^{(P\times p)} & \mathbb{U}_2,\end{bmatrix}_{l-\epsilon} \begin{bmatrix}\text{diag}\left(\mathbf{0}^{(p)}, \boldsymbol{\xi}^{(P-p)}\right) & \mathbf{0}^{(P\times(N-P))}\end{bmatrix}_l \begin{bmatrix}\mathbf{0}^{(N\times p)} \\ \\ \mathbb{V}_2^H\end{bmatrix}_{l-\epsilon},\end{aligned} \tag{3.48}$$

where $\boldsymbol{\xi}[l] \in \mathcal{R}^{((P-p)\times(P-p))}$ is defined as

$$\boldsymbol{\xi}[l] = [\xi_p\ \xi_{p+1}\ \ldots\ \xi_{P-1}]_l\,,\ \ \xi_p,\ \xi_{p+1},\ \ \ldots,\ \xi_{P-1} > 0, \tag{3.49}$$

and $\mathbb{U}_2$ and $\mathbb{V}_2$ are the left and right singular bases, respectively, that are constructed using CSI obtained during the channel-sounding phase at time $l-\epsilon$, and will be used as an orthonormal basis to signal our extrinsic fingerprinting function. Using this signaling basis, the vector $[\xi_p\ \xi_{p+1}\ \ldots\ \xi_{P-1}]_l$ will convey the extrinsic fingerprint message to the receiver.

In general, the channel stationary conditions will not hold, thus some model mismatch error between $\mathbb{L}[l-\epsilon]$ and $\mathbb{L}[l]$ will occur. This model mismatch error manifests itself as $\mathbb{M}[l]$, defined in (3.45). We will first consider system design and performance using the assumption of quasi-stationary behavior between $\mathbb{L}[l-\epsilon]$ and $\mathbb{L}[l]$ without attempting to predict $\mathbb{L}[l]$.

3.5.5 Fingerprint Recovery and Modulation

We now describe how the fingerprint signal vector $[\xi_p\ \xi_{p+1}\ \ldots\ \xi_{P-1}]_l$ may be recovered from $\mathbb{K}[l]$, of which $\hat{\mathbb{Q}}[l] - \mathbf{1}$ is an unbiased estimate. Substituting (3.48) into (3.44) we obtain

$$
\begin{aligned}
&E\left[\mathbb{Q}^T[l]\right] - \mathbf{1} = \left(\mathbf{1} + \mathbb{K}^T[l]\right) \circ \left(\mathbf{1} - \mathbb{E}^T[l]\right) - \mathbf{1} \\
&= \begin{bmatrix} \mathbf{0}^{(P\times p)} & \mathbb{U}_2 \end{bmatrix}_{l-\epsilon} \begin{bmatrix} \mathrm{diag}\left(\mathbf{0}^{(p)}, \boldsymbol{\xi}^{(P-p)}\right) & \mathbf{0}^{(P\times(N-P))} \end{bmatrix}_l \begin{bmatrix} \mathbf{0}^{(p\times N)} \\ \mathbb{V}_2^H \end{bmatrix}_{l-\epsilon} - \mathbb{E}^T[l] - \mathbb{K}^T[l] \circ \mathbb{E}^T[l],
\end{aligned}
\tag{3.50}
$$

where the terms $\mathbb{E}[l]$ and $(\mathbb{K}[l] \circ \mathbb{E}[l])$ represent model mismatch error in (3.50). These model mismatch error terms degrade the performance of the fingerprint detector. To recover the fingerprint signal $[\xi_p\ \xi_{p+1}\ \ldots\ \xi_{P-1}]_l$, we must first estimate $\mathbb{S}[l]$ from (3.48). To produce the estimate $\hat{\mathbb{S}}[l]$, the receiver pre-multiplies (3.50) by $\begin{bmatrix} \mathbf{0} & \mathbb{U}_2 \end{bmatrix}_{l-\epsilon}^H$ and post-multiplies by $\begin{bmatrix} \mathbf{0} & \mathbb{V}_2 \end{bmatrix}_{l-\epsilon}$, and (3.50) becomes

$$
\begin{aligned}
\hat{\mathbb{S}}[l] &= \begin{bmatrix} \mathbf{0} \\ \mathbb{U}_2^H \end{bmatrix}_{l-\epsilon} \left(E\left[\mathbb{Q}^T[l]\right] - \mathbf{1}\right) \begin{bmatrix} \mathbf{0} & \mathbb{V}_2 \end{bmatrix}_{l-\epsilon} \\
&= \mathbb{R}_u[l-\epsilon] \begin{bmatrix} \mathrm{diag}\left(\mathbf{0}^{(p)}, \boldsymbol{\xi}^{(P-p)}\right) & \mathbf{0} \end{bmatrix}_l \mathbb{R}_v[l-\epsilon] - \mathbb{B}[l],
\end{aligned}
\tag{3.51}
$$

where we define the fingerprint model mismatch error component $\mathbb{B}[l]$,

$$
\mathbb{B}[l] = \begin{bmatrix} \mathbf{0} & \mathbb{U}_2, \end{bmatrix}_{l-\epsilon}^H \left(\mathbb{K}[l] \circ \mathbb{E}[l] - \mathbb{E}[l]\right)^T \begin{bmatrix} \mathbf{0} & \mathbb{V}_2 \end{bmatrix}_{l-\epsilon}, \tag{3.52}
$$

and

$$
\mathbb{R}_u[l-\epsilon] = \begin{bmatrix} \mathbf{0} \\ \mathbb{U}_2^H \end{bmatrix}_{l-\epsilon} \begin{bmatrix} \mathbf{0} & \mathbb{U}_2, \end{bmatrix}_{l-\epsilon}, \tag{3.53}
$$

and

$$\mathbb{R}_v[l-\epsilon] = \begin{bmatrix} \mathbf{0} \\ \mathbb{V}_2^H \end{bmatrix}_{l-\epsilon} \begin{bmatrix} \mathbf{0} & \mathbb{V}_2 \end{bmatrix}_{l-\epsilon}. \tag{3.54}$$

The left and right correlation matrices, $\mathbb{R}_u[l-\epsilon]$ and $\mathbb{R}_v[l-\epsilon]$ respectively, can be used as a measure of the *closeness* of the left and right extrinsic signaling basis $\mathbb{U}_2$ and $\mathbb{V}_2$, respectively, to a true unitary basis for their respective subspaces. An optimal selection of extrinsic fingerprinting bases, yielding the correlation matrices $\mathbb{R}_u^*[l-\epsilon]$ and $\mathbb{R}_v^*[l-\epsilon]$, respectively, would preserve the orthogonality of the intrinsic and extrinsic subspaces. Thus, an optimal selection of bases would yield $\mathbb{R}_u^*[l-\epsilon] = \begin{bmatrix} \mathbf{0}^{(P\times p)} & \mathbf{I}^{(P\times P-p)} \end{bmatrix}$ and $\mathbb{R}_v^*[l-\epsilon] = \begin{bmatrix} \mathbf{0}^{(N\times p)} & \mathbf{I}^{(N\times N-p)} \end{bmatrix}$, respectively. We define deviation from a true orthonormal signaling basis for the left and right bases, e_{R_u} and e_{R_v} respectively, as the Frobenius norm of the difference between $\mathbb{R}_v[l-\epsilon]$ and $\begin{bmatrix} \mathbf{0}^{(P\times p)} & \mathbf{I}^{(P\times P-p)} \end{bmatrix}$, and $\mathbb{R}_u[l-\epsilon]$ and $\begin{bmatrix} \mathbf{0}^{(N\times p)} & \mathbf{I}^{(N\times N-p)} \end{bmatrix}$, respectively, i.e.,

$$e_{R_u} = \| \mathbb{R}_u - \begin{bmatrix} \mathbf{0}^{(P\times p)} & \mathbf{I}^{(P\times P-p)} \end{bmatrix} \|_F, \tag{3.55}$$

and

$$e_{R_v} = \| \mathbb{R}_v - \begin{bmatrix} \mathbf{0}^{(N\times p)} & \mathbf{I}^{(N\times N-p)} \end{bmatrix} \|_F, \tag{3.56}$$

which are both non-negative values.

When both $\mathbb{R}_v[l-\epsilon]$ and $\mathbb{R}_u[l-\epsilon]$ are perfectly unitary, a condition which we will

denote with the subscripts (ru) and (rv) respectively, (3.51) becomes simply

$$\begin{aligned}\hat{\mathbb{S}}_{ru,rv}[l] &= \begin{bmatrix}\mathbf{0}\\ \mathbb{U}_2^H\end{bmatrix}_{l-\epsilon} \begin{bmatrix}\mathbf{0} & \mathbb{U}_2\end{bmatrix}_{l-\epsilon} \begin{bmatrix}\mathrm{diag}\left(\boldsymbol{0}^{(p)}, \boldsymbol{\xi}^{(P-p)}\right) & \mathbf{0}\end{bmatrix}_{l} \begin{bmatrix}\mathbf{0}\\ \mathbb{V}_2\end{bmatrix}_{l-\epsilon} \begin{bmatrix}\mathbf{0} & \mathbb{V}_2^*\end{bmatrix}_{l-\epsilon} - \mathbb{B}[l] \\ &= \begin{bmatrix}\mathrm{diag}\left(\boldsymbol{0}^{(p)}, \boldsymbol{\xi}^{(P-p)}\right) & \mathbf{0}\end{bmatrix}_{l} - \mathbb{B}[l].\end{aligned} \tag{3.57}$$

From (3.57) we note that the channel model mismatch error term $\mathbb{B}[l]$ is still present. However, under perfect channel estimation conditions when $\mathbb{L}[l] = \mathbb{L}[l-\epsilon]$, a condition which we will denote with the subscript (m), the channel model-mismatch term becomes the zero matrix and no model mismatch error is present. Thus (3.57) is simply

$$\hat{\mathbb{S}}_{ru,rv,m}[l] = \mathbb{S}[l] = \begin{bmatrix}\mathrm{diag}\left(\boldsymbol{0}^{(p)}, \boldsymbol{\xi}^{(P-p)}\right) & \mathbf{0}^{(P\times N-P)}\end{bmatrix}_{l}, \tag{3.58}$$

and the extrinsic fingerprint signal of a fingerprinted block transmitted at time l may be recovered from $\hat{\mathbb{S}}[l]$ by simply extracting the elements $[\xi_p\ \xi_{p+1}\ \ldots\ \xi_{P-1}]$ using (3.58) and (3.49). From (3.57) we observe that any model mismatch will degrade the fingerprint statistics $[\xi_p\ \xi_{p+1}\ \ldots\ \xi_{P-1}]$ as interference.

3.5.6 Extrinsic Fingerprint Overlay Design

In this section we evaluate various methodologies for incorporating previous CSI into the design of the left and right signaling bases $\mathbb{U}_2$ and $\mathbb{V}_2$, respectively, and discuss the performance trade-offs of these designs. The design taxonomy we present will be ordered according to the amount of CSI required, in descending order. Therefore, we will lead our discussion with designs that require the greatest amount of CSI, and end our discussion with designs that do not require CSI at the transmitter at all.

3.5.6.1 Direct Fingerprint Overlay Using Full CSI

We first consider the possibility of directly using the left and right singular vectors $\mathbb{U}_2$ and $\mathbb{V}_2$ from (3.48) to implement $\mathbb{U}_2$ and $\mathbb{V}_2$, respectively, and denote this design $\mathbb{K}_{direct}$, i.e.

$$\mathbb{K}_{direct}[l] = \left[\mathbf{0}^{(P\times p)} \quad \mathbb{U}_2^H,\right]_{l-\epsilon} \left[\text{diag}\left(\mathbf{0}^{(p)}, \boldsymbol{\xi}^{(P-p)}\right) \quad \mathbf{0}^{(P\times N-P)}\right]_l \begin{bmatrix} \mathbf{0}^{(p\times N)} \\ \\ \mathbb{V}_2^H \end{bmatrix}_{l-\epsilon} . \quad (3.59)$$

While $\mathbb{K}_{direct}$ demonstrates that $\mathbb{U}_2$ and $\mathbb{V}_2$ can be used directly to implement an orthonormal basis for signaling $[\xi_p\ \xi_{p+1}\ \ldots\ \xi_{P-1}]$, there are a number of downfalls to this approach. To recover $\mathbb{K}_{direct}$ via (3.51), both $\mathbb{U}_2$ and $\mathbb{V}_2$ must be communicated to the transmitter from the receiver, requiring a total of $P(P-p)+N(N-p)$ units of feedback information. Also, if $\mathbb{V}_2[l-\epsilon]$ accurately models the noise subspace in the frequency dimension under particular channel conditions while the singular vectors of $\mathbb{U}_2[l-\epsilon]$ inaccurately model the noise subspace in the time dimension, $\mathbb{U}_2[l-\epsilon]$ will predominately contribute to model mismatch error component, $\mathbb{E}[l]$. This would be the case when $\mathbb{U}_2[l-\epsilon]$ captures transient fading or other irrelevant temporal information. Vice-versa, a similar argument may be made for $\mathbb{U}_2[l-\epsilon]$ under some channel conditions, where in this case $\mathbb{U}_2[l-\epsilon]$ is an accurate model of $\mathbb{U}_2[l]$ but $\mathbb{V}_2[l-\epsilon]$ has captured some inaccurate information and thus predominately contributes to model-mismatch error. To ameliorate these effects we will now consider the possibility of using a standard, uniform orthonormal basis for $\mathbb{U}_2[l-\epsilon]$ and/or $\mathbb{V}_2[l-\epsilon]$ when designing the overlay.

3.5.6.2 Uniform Fingerprint Overlays Using Partial CSI

In general, the columns of $\mathbb{U}_2$ forming the left signaling basis for our fingerprint message in the time dimension can be selected from any unitary matrix that is at least size $\mathcal{C}^{(P\times(P-p))}$. Similarly, the columns of $\mathbb{V}_2$ forming the right signaling basis for our fingerprint message in the frequency dimension can be selected from any unitary matrix that is at least size $\mathcal{C}^{(N\times N-p)}$, and further, this basis may be selected independently from $\mathbb{U}_2$. We note that deviation from $\mathbb{U}_2$ and/or $\mathbb{V}_2$ as used in $\mathbb{K}_{direct}$ necessarily degrades the orthogonality between the intrinsic bases $\mathbb{U}_1[l-\epsilon]$ and $\mathbb{V}_1[l-\epsilon]$ and the extrinsic fingerprinting overlay formed by $\mathbb{U}_2$ and $\mathbb{V}_2$. When the orthogonality between these subspaces is degraded $\mathbb{U}_1$ will partially project on $\mathbb{U}_2$ as interference, $\mathbb{V}_1$ will partially project onto $\mathbb{V}_2$ as interference, and vice-versa.

A number of matrices with the unitary property exist in the literature that would suffice for selecting $\mathbb{V}_2$ and/or $\mathbb{U}_2$, however three typical unitary matrices will be considered here: the Identity matrix, the Discrete Fourier Transform matrix (DFT) matrix, and the Walsh-Hadamard matrix. Since $\mathbb{Q}[k]$ in (3.26) is the Hadamard product between intrinsic channel estimate matrix $\mathbb{H}[k]$ and the fingerprinting matrix $\mathbb{F}[k]$, another desirable property of $\mathbb{F}[k]$ is that it does not unduly bias particular elements of $\mathbb{Q}[k]$ in either the time or frequency dimensions when conveying $[\xi_p\ \xi_{p+1}\ \ldots\ \xi_{P-1}]$ to the receiver. Both the Walsh-Hadamard matrix, denoted $\mathcal{H}$ and the DFT matrix, denoted $\mathcal{W}$, are felicitous choices for selecting $\mathbb{U}_2$ and/or $\mathbb{V}_2$, since the majority of elements in these matrices are non-zero. This property effectively allows these bases to *spread* the extrinsic fingerprint signal in the respective dimension, i.e. *frequency spreading* for $\mathbb{V}_2$ and *time spreading* for

$\mathbb{U}_2$. We define the following design criteria when selecting $\mathbb{U}_2[k]$ and $\mathbb{V}_2[k]$:

- If $\mathbb{U}_2[k]$ is to be designed from a uniform signaling basis, select the $P-p$ columns of $\mathbb{U}_2[k]$ from a column subset of a unitary matrix of size $\mathcal{C}^{(P\times P)}$.
- Similarly, if $\mathbb{U}_2[k]$ is to be designed from a uniform signaling basis, select the $N-p$ columns of $\mathbb{V}_2[k]$ from a column subset of a unitary matrix of size $\in \mathcal{C}^{(N\times N)}$.

Using the criteria above, $\mathbb{U}_2[k]$ can be designed using either a Walsh-Hadamard matrix or a DFT matrix as a basis, and $\mathbb{V}_2[k]$ can use either a Walsh-Hadamard matrix or a DFT matrix as a basis, and the the basis selections for $\mathbb{U}_2[k]$ and/or $\mathbb{V}_2[k]$ may be made independently.

With this design criteria, we define an extrinsic fingerprint overlay design for $\mathbb{K}[l]$ where $\mathbb{U}_2$ is drawn from a Walsh-Hadamard basis such that

$$\mathbb{K}_{hu}^{T}[l] = \left[\mathbf{0}^{(P\times p)} \quad \mathcal{H}^{(P\times P-p)}\right]_{l-\epsilon} \left[\mathrm{diag}\left(\mathbf{0}^{(p)}, \boldsymbol{\xi}^{(P-p)}\right) \quad \mathbf{0}^{(P\times N-P)}\right]_{l} \begin{bmatrix} \mathbf{0}^{(N\times p)} \\ \\ \mathbb{V}_2 \end{bmatrix}_{l-\epsilon}^{H}, \tag{3.60}$$

where the subscript (hu) on $\mathbb{K}_{hu}$ denotes that the columns of $\mathbb{U}_2$ are selected from a subset of columns of a Walsh-Hadamard matrix of size $\mathcal{R}^{(P\times P)}$, while $\mathbb{V}_2$ are the original columns of the noise subspace projected in the frequency dimension derived from channel-sounding information obtained from $\mathbb{H}[l-\epsilon]$. While $\mathbb{K}_{hu}$ can be an improvement over $\mathbb{K}_{direct}$ for some channel conditions, it still requires transmission of $\mathbb{V}_2$ to the transmitter using $N(N-p)$ resources of feedback.

To improve on the feedback requirement of (3.60), we also consider the case where

$\mathbb{V}_2$ is selected from a standard, uniform basis, i.e.

$$\mathbb{K}_{hv}[l] = \left[\mathbf{0}^{(P\times p)} \quad \mathbb{U}_2\right]_{l-\epsilon} \left[\mathrm{diag}\left(\mathbf{0}^{(p)}, \boldsymbol{\xi}^{(P-p)}\right) \quad \mathbf{0}^{(P\times N-P)}\right]_{l} \left[\begin{array}{c} \mathbf{0}^{(N\times p)} \\ \mathcal{H}^{(N\times N-p)} \end{array}\right]^{H}_{l-\epsilon}, \quad (3.61)$$

where the subscript hv on $\mathbb{K}_{hv}$ denotes that $\mathbb{V}_2[k]$ is selected from a subset of columns of a Walsh-Hadamard matrix of size $\mathcal{R}^{(N\times N)}$, while $\mathbb{U}_2$ are the original columns of the noise subspace projected in the time dimension and derived from channel-sounding information obtained from $\mathbb{H}[l-\epsilon]$. While $\mathbb{K}_{hv}$ improves on the feedback required by $\mathbb{K}_{wu}$, requiring transmission of only $\mathbb{U}_2$ to the transmitter using $P(P-p)$ resources for feedback, $\mathbb{K}_{hv}$ leverages much less CSI in the design of $\mathbb{K}[l]$. Additionally, if the information captured by $\mathbb{U}_2$ represents transient temporal information while $\mathbb{V}_2$ captures accurate frequency-selective fading behavior, $\mathbb{K}_{hu}$ may yield greater model mismatch error than $\mathbb{K}_{hv}$ because the CSI used in the design of $\mathbb{K}_{hv}$ may be an inaccurate representation of channel state during $\mathbb{H}[l]$.

3.5.6.3 Fingerprint Overlays Requiring Zero CSI

For comparison, we also consider the *blind* orthonormal signaling basis overlay, where both $\mathbb{U}_2$ and $\mathbb{V}_2$ are selected from standard, uniform signaling bases and previous CSI is not needed or used by the transmitter. When both $\mathbb{U}_2$ and $\mathbb{V}_2$ are replaced with columns of the Walsh-Hadamard matrix, we denote the result $\mathbb{K}_{hu,hv}$

$$\mathbb{K}_{hu,hv}[l] = \left[\mathbf{0}^{(P\times p)} \quad \mathcal{H}^{(P\times P-p)}\right]_{l-\epsilon} \left[\mathrm{diag}\left(\mathbf{0}^{(p)}, \boldsymbol{\xi}^{(P-p)}\right) \quad \mathbf{0}^{(P\times N-P)}\right]_{l} \left[\begin{array}{c} \mathbf{0}^{(N\times p)} \\ \mathcal{H}^{(N\times N-p)} \end{array}\right]^{H}_{l-\epsilon}. \quad (3.62)$$

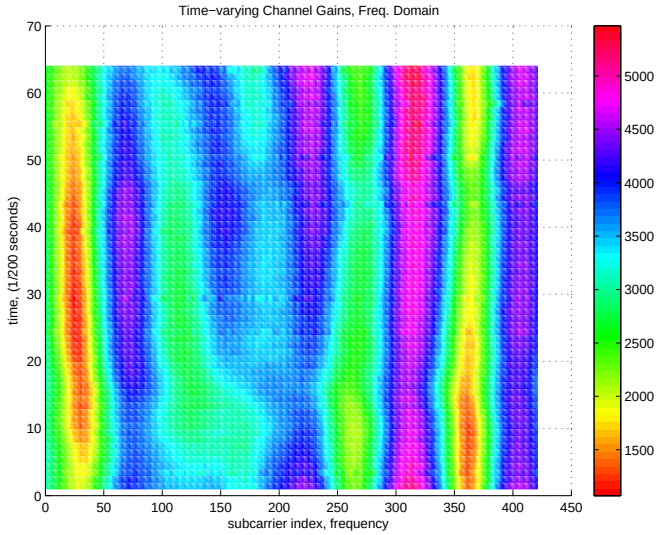
One advantage of using $\mathbb{K}_{hu,hv}[l]$ is that no previous CSI is required in the design of the extrinsic fingerprint overlay, as receiver feedback is not required. The primary disadvantage of $\mathbb{K}_{hu,hv}[l]$ is that orthogonality between the intrinsic channel distortions and the extrinsic fingerprinting subspace is not present, thus distortions indicative of the intrinsic time-varying channel will act as interference in the detection of the extrinsic fingerprint signal. Since the $\mathbb{K}_{hu,hv}[l]$ design does not need or use CSI, in (3.62) p is not the effective rank but merely determines the dimensionality of the fingerprint. We will consider this design for comparison in our experiments in Section 3.6 and simulations in Section 3.7.

Similarly, (3.60), (3.61), and (3.62) may select columns from the DFT matrix for $\mathbb{U}_2$ and/or $\mathbb{V}_2$, yielding $\mathbb{K}_{wu}[l]$, $\mathbb{K}_{wv}[l]$, and $\mathbb{K}_{wu,wv}[l]$, respectively, however the delineation of these designs will be omitted as they are similarly defined. The bases for $\mathbb{U}_2$ and $\mathbb{V}_2$ may be selected independently, yielding the following possible designs:

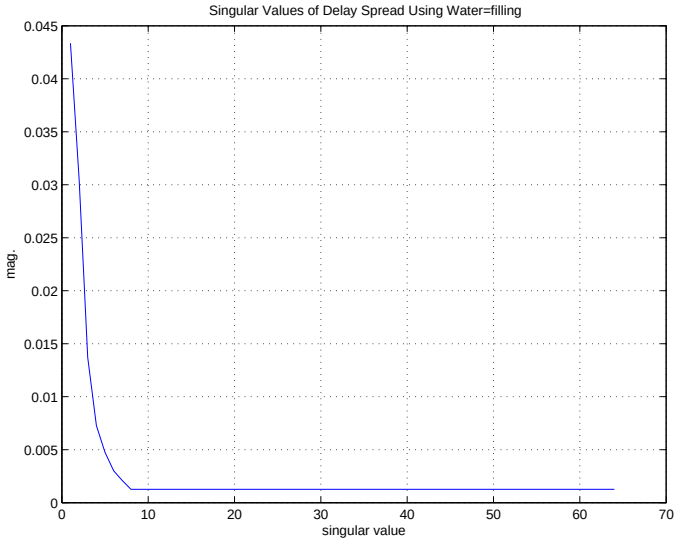
$$\mathbb{K}_{x1,x2}[l], \;\; x1 \in \{-, hu, wu\}\,, \; x2 \in \{-, hv, wv\}\,. \tag{3.63}$$

3.6 Experimental Results

We now present experimental results for the fingerprinting method described, using time-varying channel estimate data collected from the IEEE 802.16e WiMax waveform. Raw signals were collected from a 5 MHz WiMax base-station configured to use a $N = 512$ subcarrier FFT, and a sequence of channel estimates were obtained using the training data present in each 5ms frame preamble. This experiment data represents the channel conditions of a hand-held mobile unit, where other than subtle hand movement the mobile user is stationary.



(a) Top view



(b) Sorted Singular Values

Figure 1: Time-Varying Channel Gains $\mathbb{H}[l - \epsilon]$

The WiMax preamble uses a known data sequence that is duplicated three consecutive times in the time domain, therefore a $3\times$ sinc interpolating filter was used to complete the channel estimate. A sequence of $P = 64$ frames was selected to form $\mathbb{H}[l-\epsilon]$, and the magnitudes of the equalizer-tap gains associated with each of the N subcarriers are presented in top view in Figure 1(a). From Figure 1(a) we readily observe the slow frequency-selective fading behavior of this channel where areas of deepest fade are lightest, while the areas with the least fading are darker. We observe that the locations of frequency-selective fades are highly correlated in time, and that the environment is slightly changing since the locations of frequency-selective fades drift slightly.

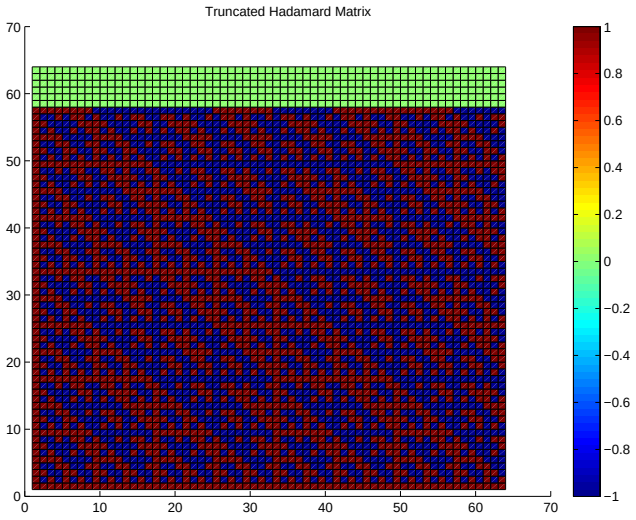
After diagonalizing the time-varying CSI of $\mathbb{H}[l-\epsilon]$ presented in Figure 1(a) via SVD, a plot of the sorted singular values of $\mathrm{diag}\left(\mathbf{\Lambda}[l-\epsilon]\right)$ is given in Figure 1(b). From Figure 1(b) we estimate the effective rank of $\mathbb{H}[l-\epsilon]$ to be approximately 8, and thus we select $p = 8$ accordingly, yielding a fingerprinting subspace of size $|\boldsymbol{\xi}[l]| = P - p = 56$, i.e. 56 eigenvectors and accompanying eigenvalues may be used for embedding the fingerprint message. The water-level used in this experiment was selected to be equal to $\xi_p = 0.0013$. In a full fingerprinting system, the elements $\xi_i,\ i = p+1,\ldots,P$ will be selected from a uniform PAM-like constellation to signal the digital fingerprint message.

In this experiment the $\mathbb{K}_{hu}[l]$ fingerprint overlay design was used, where the left singular-vectors of $\mathbb{U}_2[l-\epsilon]$ spanning $\mathbb{K}[l]$ in time are replaced with columns from a Hadamard matrix of size $\mathcal{H} \in \mathcal{R}^{(P\times P)}$ yielding the augmented version of this basis denoted $\mathbb{U}_2[l-\epsilon]$. According to the $\mathbb{K}_{hu}[l]$ design, the right singular-vectors are used directly, i.e. $\mathbb{V}_2[l-\epsilon] = \mathbb{V}_2[l-\epsilon]$. This design effectively *spreads* the fingerprinting signal in the time dimension using a time-uniform basis consisting of Walsh codes, in a way similar

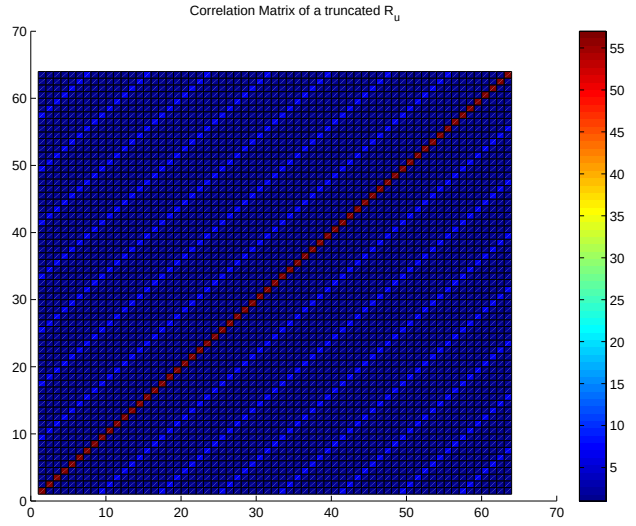
to frequency spreading via Walsh codes in CDMA systems. The modified singular basis, $\mathbb{U}_2[l-\epsilon]$, used in the design of $\mathbb{K}_{hu}[l]$ is presented in Figure 2(a). In Figure 2(a) the "$+1$" values of the Walsh code column vectors are represented by white, the "-1" values are represented by black, and the zero values represented by gray.

When a non-unitary basis is used in the construction of $\mathbb{U}[l-\epsilon]$, the innate orthogonality of the columns of the original matrix is degraded. For example, this degradation is apparent in $\mathbb{R}_u[l-\epsilon]$ from equation (3.53), depicted in Figure 2(b) for a truncated Hadamard matrix. In this example the top eight rows of the matrix are zero, thus truncating the Walsh codes that form the columns of $\mathbb{U}[l-\epsilon]$. From Figure 2(b), we notice that the diagonal elements of $\mathbb{U}^H[l-\epsilon]\mathbb{U}[l-\epsilon]$ represented by the white elements, which were exactly 1 for the $P \times P$ sized Hadamard matrix, now take on values slightly less than 1 while the off-diagonal elements which were previously equal to zero now have non-zero elements represented by dark gray. The gray elements denote off-diagonal correlation, or *cross-projection* of the fingerprint signaling vectors which leads to inner-signal interference during fingerprint recovery.

We first consider the case of no model-mismatch error, i.e. when $\mathbb{L}[l] = \mathbb{L}[l-\epsilon]$, by applying the fingerprint to the same block of CSI used in the construction of $\mathbb{K}_{hu}[l]$. The magnitude of the fingerprinted time-varying CSI, $\mathbb{Q}[l]$, using the same intrinsic channel distortions of Figure 1(a) and precoding using the $\mathbb{K}_{hu}[l]$ fingerprinting design, is depicted in Figure 3(a). We note that the fingerprinted CSI of Figure 3(a) and the original CSI of Figure 1(a) are very similar, and that the fingerprinted version is visually a *noisier* version of the original intrinsic time-varying channel distortions. By preforming the fingerprint recovery steps of (3.51), $\hat{\mathbb{S}}[l-\epsilon]$ can be recovered and the fingerprint signal elements



(a) $\mathbb{U}$ derived from a Hadamard matrix



(b) The left correlation matrix $\mathbb{R}_u$

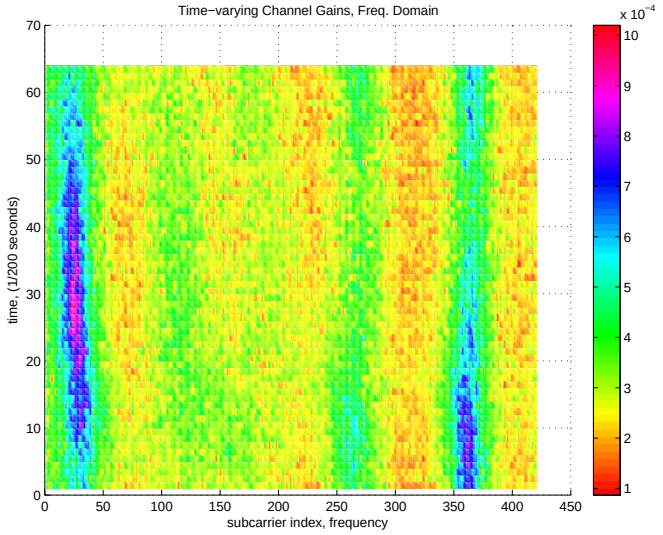
Figure 2: Left eigenspace $\mathbb{U}$

$[\xi_{p+1} \ \ldots \ \xi_{P-1}]$ can then be recovered from the diagonal elements of $\hat{\mathbb{S}}[l-\epsilon]$, as depicted in Figure 3(b), while the elements $[\xi_i \ \xi_0 \ \ldots \ \xi_P]$ are omitted.

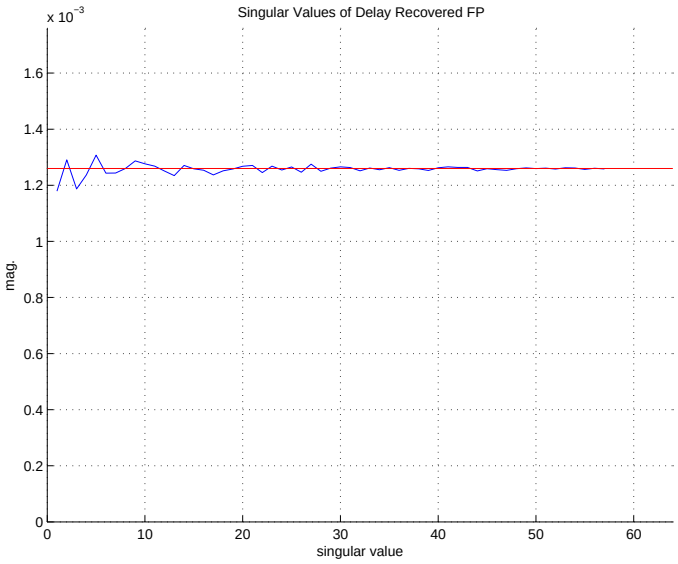
From Figure 3(b) we readily observe the effect of receiver noise on the received fingerprint signal, as the values $[\xi_{p+1} \ \ldots \ \xi_{P-1}]$ should all be identically equal to $\xi_i = \xi_p = 0.0013, \ \forall i = p+1, \ldots, P-1$. We note that even in the absence of model-mismatch error these values are distorted by noise. It is clear, however, that in this example a ML receiver can recover the digital fingerprint, all 1's in this case, without bit errors using 2-level PAM signaling and a symbol decision region that is half way between 0 and 0.0013.

While Figures 3(a) and 3(b) demonstrate fingerprint application in the absence of model mismatch error, we now consider the performance of a the fingerprint overlay when $\mathbb{K}[k]$ is applied to a future block of data transmissions. For this result we select the next $P = 64$ channel estimates from the same signal used to create Figure 1(a). The resulting fingerprinted time-varying CSI, $\mathbb{Q}[l]$, for the more general case when $\mathbb{L}[l] \neq \mathbb{L}[l-\epsilon]$ is presented in Figure 4(a), using the same $\mathbb{K}_{hu}[k]$ design delineated in Figure 3(a). We note that the fading behavior depicted in Figure 4(a) is highly correlated with the fading behavior shown in Figure 3(a), as this channel estimate data was obtained from the same WiMax signal and these blocks of CSI are exactly $P = 64$ OFDM frames, or $320ms$, apart. This adjacent block of CSI also demonstrates frequency-selective fade locations that are highly correlated in time. We also present the the extracted fingerprinting signal elements $[\xi_{p+1} \ \ldots \ \xi_{P-1}]$ in Figure 4(b).

Comparing the results of Figure 4(b) to Figure 3(b), the recovered fingerprinting signal elements $[\xi_{p+1} \ \ldots \ \xi_{P-1}]$ are even more distorted due to the additional model miss-match error.

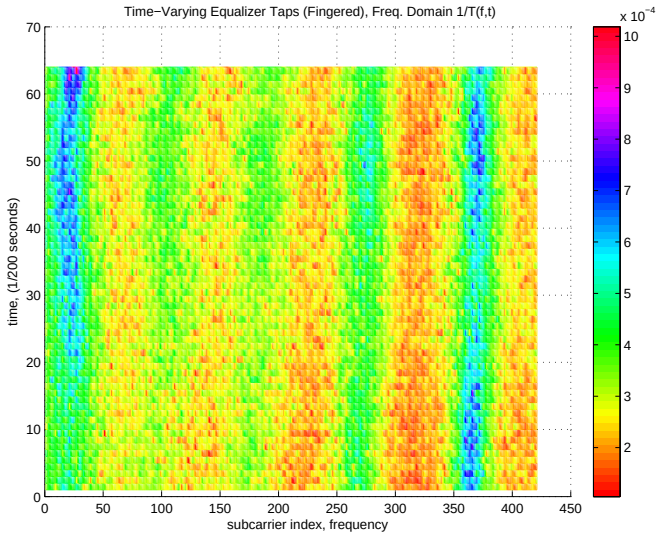


(a) Fingerprinted Time-Varying CSI

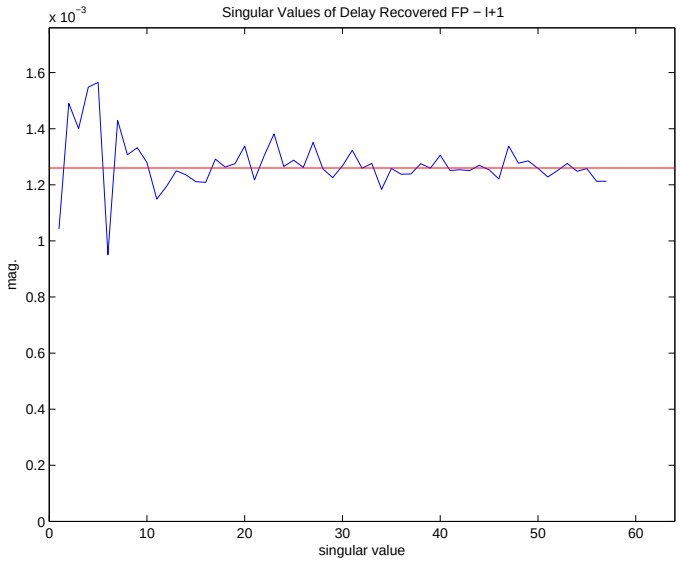


(b) Singular values

Figure 3: Results for $\mathbb{L}[l] = \mathbb{L}[l - \epsilon]$ using the $\mathbb{K}_{hu}$ overlay



(a) Fingerprinted CSI $\mathbb{Q}$



(b) Singular values of the fingerprinted symbol $\hat{\mathbb{S}}$

Figure 4: $\mathbb{K}_{hu}$ overlay with model mismatch error

3.7 Simulation Results

In this section we present simulation results for one of the fingerprint overlay designs in (3.63), using the intrinsic time-varying channel model described by (3.14), (3.15), (3.16), and (3.17) in [64]. In [64], each channel gain of the delay profile is modeled using an Auto Regressive process of order 1 (AR-1), and the driving process for each AR-1 model is a Gaussian process with variance σ_T^2. As σ_T increases, the magnitude of the fluctuation of each channel gain in the model increases, and channel conditions change more violently.

Through simulation, we can quantitatively compare the performance of overlay designs and measure any degradation experienced by the primary signal from the embedding of the fingerprint. To measure primary signal degradation, we compare the BERs of the primary signal with and without the fingerprint present. Additionally, we use BER to compare the efficacy of each overlay design.

To simulate the embedding of various fingerprint designs, a full OFDM system and accompanying channel simulator were created in Matlab. The OFDM signal generated uses a 512-point FFT with 430 occupied subcarriers and 41 left and right guard subcarriers. A BPSK-modulated preamble occupies the first symbol of each frame using a $2\times$ time-domain repetition. The preamble is represented in (3.9) by $\mathbf{P}$, while the primary signal payload is represented by $\mathbf{D}[k]$. In the simulated signal, every odd frame is fingerprinted while every even frame is used for channel sounding.

The OFDM frame is then subjected to a simulated time-varying channel by applying $g(t)$ in the time domain using a transversal filter, according to (3.1). Timing jitter was also

added to test the fingerprint's affect on typical frame synchronization algorithms. For the channel delay-spread gains, $\bar{h}_j$, the simplified *Typical Urban* delay spread profile of section B.1 in the 3GPP specification [2] was used, while the values for the AR-1 model coefficients, α_j were chosen empirically.

To decode the signal, the Schmidl and Cox algorithm [51] was first applied for course timing synchronization. The LS channel estimator [59] was applied to estimate the aggregate channel distortion using the frame preamble as training data, and the resulting estimate was then up-sampled using $2\times$ sinc interpolation. The channel sounding symbol and fingerprinted symbol were then equalized independently, and $\hat{\mathbf{q}}[l-\epsilon]$ and $\hat{\mathbf{h}}[l]$ are recorded for each frame for later use during the fingerprint recovery phase. Both the BPSK preamble data and QPSK payload data are demodulated into bits, and bit errors for the fingerprinted frames and non-fingerprinted frames are recorded.

In Figures 5(a) through 5(d), the $\mathbb{K}_{hu,wu}$ overlay design was used using $p = 0$ and the *Typical Urban* channel delay spread profile. A plot of the simulated time-varying CSI, $\mathbb{H}[l]$, is given in Figure 5(a), while a plot of the fingerprinted CSI, i.e. $\mathbb{Q}[l]$, is given in Figure 5(b) By selecting $\mathbb{K}_{hu,wu}$ and $p = 0$, CSI is not used in the design of $\mathbb{K}_{hu,wu}$. We will use this design for comparison when considering designs that do leverage knowledge of previous CSI.

In Figure 5(b) we observe minimal distortion in the fingerprinted CSI due to the embedding of the fingerprint, while the actual embedded fingerprint signal, $\mathbb{F}[l]$, is depicted in Figure 5(c) for the $\mathbb{K}_{hu,wu}$ overlay design using $p = 0$. The recovered fingerprint signal, $\hat{\mathbb{F}}[l]$, is depicted in Figure 5(d), where we see the effects of the intrinsic time-varying channel on the blind superposition design in the additional distortions in Figure

5(d) which are not present in Figure 5(c). This is caused by the lack of orthogonality between the intrinsic and extrinsic subspaces discussed in Section 3.5, thus demonstrating the susceptibility of the blind fingerprint approach to intrinsic channel distortions. For comparison, the same plots are presented for the $\mathbb{K}_{hu}$ overlay design using $p = 8$ in Figure 6(a) and 6(b). By comparing Figure 5(c) to Figure 6(a), we see that the fingerprint in Figure 6(a) is more noise-like since its basis incorporates CSI derived from the noise subspace.

The BER results for the $\mathbb{K}_{hu,wu},\ p = 0$ design are given in Figure 7(a) with $\sigma_T^2 = 0.01$ and the values $\lambda_{fp} = 0.7$ and $\lambda_{fp} = 1.0$, where λ_{fp} is the signal magnitude of ξ_i representing transmission of a 1 while transmission of a 0 is represented by zero, when using two-level signaling. Figure 7(b) shows results for the $\mathbb{K}_{hu,wu},\ p = 0$ design, with $\lambda_{fp} = 1.0$ and the values $\sigma_T^2 = 0.01$ and $\sigma_T^2 = 0.015$. We observe that the $\mathbb{K}_{hu,wu}$ fingerprint design does yield a BER improvement of 10 to 20 dB over the primary signal for both values of λ_{fp}, for SNR greater than 7 dB. This design also operates with a BER advantage for the low SNR regions of Figure 7(b), however, for $\sigma_T^2 = 0.015$ the authentication signal BER flattens out around 10^{-5} as σ_T^2 becomes the predominate noise term in (3.14) and the time-varying distortions represented by $\mu_j[k]$ further degrade the fingerprint signal.

For comparison, BER results for the $\mathbb{K}_{hu},\ p = 8$ overlay design are given in Figures 7(c) and 7(d). We observe that the $\mathbb{K}_{hu}$ fingerprint design also achieves a BER improvement over the primary signal for SNR greater than 7 dB. The 'flattening out' phenomenon of the authentication signal BER for $\sigma_T = .015$ is also apparent, as the fingerprint yields BER slightly lower than 10^{-5} in higher SNR. In Figures 7(a), 7(b), 7(c), and 7(d), we see

zero impact to primary signal BER do to the presence of the fingerprint message, since the primary signal series with and without the fingerprint present completely overlap for a given value of σ_T.

To observe the benefits of incorporating previous CSI into the design of the fingerprinting overlay, we display the authentication signal BERs for the $\mathbb{K}_{hu,wu},\ p = 0$ and $\mathbb{K}_{hu},\ p = 8$ designs together, for the values $\sigma_T = [0.02,\ 0.03]$ in Figure 8(a). The primary signal BER for these simulations is depicted in Figure 8(b), and from this figure we observe nearly zero impact to primary signal BER for both values of σ_T^2. From 8(a) we observe that the fingerprint overlay design incorporating CSI, i.e. $\mathbb{K}_{hu}$, outperforms the design that does not incorporate previous CSI, i.e. $\mathbb{K}_{hu,wu}$. This advantage is demonstrated by the lower BER of the $\mathbb{K}_{hu}$ design, for all values of σ_T. From 8(a) we also observe that the BER advantage of the $\mathbb{K}_{hu}$ design increases as σ_T increases. This is because an increase in σ_T corresponds to an increase in model mismatch error, which manifests itself as $\mathbb{B}[l]$ in (3.51). The incorporation of CSI into the $\mathbb{K}_{hu}$ design helps mitigate the distortions caused by model-mismatch error.

One observation that can be made is that the proposed extrinsic fingerprinting overlays are very similar to pre-coding or pre-filtering in OFDM systems. However, some of the fingerprint designs discussed in this chapter require much more detailed channel information than common precoding schemes typically require. One may question the practicality of the required channel information used in some designs, in pragmatic wireless systems. Additional comment on the amount of feedback required on the transmitter side may also be in order. On these topics we present the following discussion.

Our work presents a design taxonomy that incorporates various amounts of CSI into

the fingerprint design and one very important design that we discuss, the $\mathbb{K}_{hu,hv}[l]$ design, *does not require* CSI at the transmitter, because it simply spreads the fingerprint in both time and frequency using both the Hadamard and DFT spreading bases. Therefore, we lead our discussion with designs that require the greatest amount of CSI, and end our discussion with designs that do not require CSI at the transmitter at all.

Because of the practicality of the $\mathbb{K}_{hu,hv}[l]$ design, and similar designs that require zero CSI feedback, we compare every simulation result is compared to this extremely important design. Our goal is to discuss, analyze, and critique the performance of various subspace fingerprint embedding designs, using different amounts of CSI, via a design taxonomy. The $\mathbb{K}_{hu,wu}$ design is practical for sure, and depending on the application, other designs using partial CSI can be leveraged in scenarios when feedback channels are available.

One may question why we consider the Least Squares (LS) channel estimator in the simulation section of this paper, instead of more accurate methods such as MMSE, when a great deal of channel information is assumed to be known. On this topic, we note that any other channel estimator can work for our embedding scheme, however the selection of a particular channel estimator does not necessarily influence the design of the fingerprint. Hence we use LS channel estimation in the simulation and experiment sections due to its widespread adaptation in OFDM systems and low computational requirements which make the LS channel estimator very practical to implement.

One may also question why proposed scheme is based on the assumption that the attackers cannot forge the fingerprints $\mathbf{F}[k]$, while the transmitter could potentially derive the fingerprint according to the channel information feedback from the receiver side. To

this question, we offer the following discussion.

The elements of the fingerprint signal vector $\boldsymbol{\xi}[l]$ which determines the fingerprinting function $\mathbf{F}[k]$ are used as symbols to modulate a multi-bit digital authentication message. Therefore, to forge $\mathbf{F}[k]$ an attacker would also need to fabricate the fingerprint signal vector $\boldsymbol{\xi}[l]$, and thus the bits of the digital authentication message, which is protected from forgery using a cryptographic signature and secret key as stated in paragraph 1 of page 2 of the manuscript. While our paper focuses on the embedding and signaling of a digital fingerprint message for OFDM systems, and the design of specific digital authentication messages is outside of the scope of the paper. We note that the basic authentication message example in Section III-F is only an example message, and is not the focus of our work.

We would like to offer some additional discussion on compromised keys. The security of the authentication signature systems rests on the secrecy of the private keys used to design the secure signatures within the authentication message that is conveyed by $\boldsymbol{\xi}[l]$. In the event that an attacker knows all information about the signaling bases used by $\mathbf{F}[k]$, the signal constellation and symbol-to-bit-map used by the elements in $\boldsymbol{\xi}[l]$, all channel state information used to design $\mathbf{F}[k]$, and the authorized user's secret keys are compromised, an attacker would be able to forge an authentic fingerprint signal $\mathbf{F}[k]$ and masquerade as an authorized user. Therefore secret keys should be used to sign the authentication message are protected.

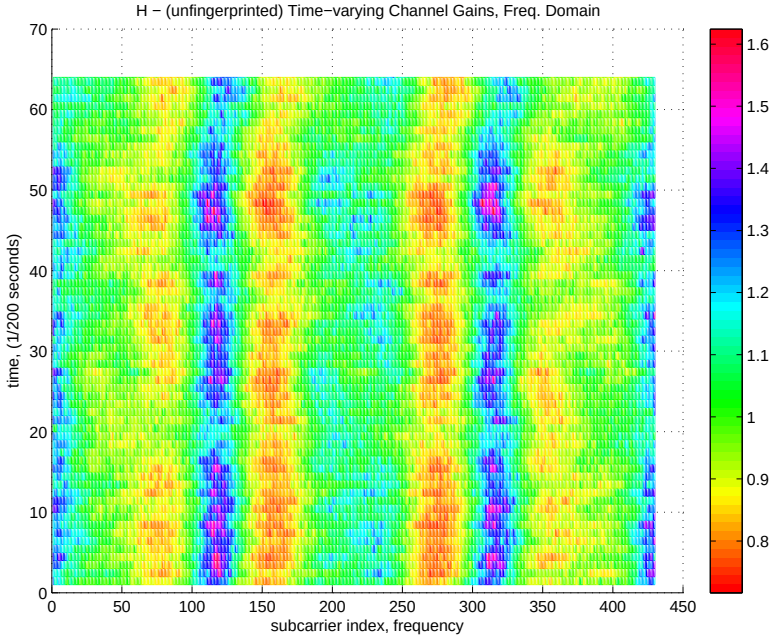
Figures 4-7 depict some useful information on the spreading behavior of our fingerprint designs, by demonstrating how the received channel estimates are perturbed by the presence of the fingerprint overlay. The primary purpose of these figures is to show

how the fingerprint is uniformly spread throughout fingerprinted block, and to demonstrate the relative magnitude of the perturbed channel estimates when compared to the original channel estimates. Error rates for a particular digital authentication message of any given length can be formulated using the BER results presented in Figures 8 and 9, and by considering any additional coding that may be employed.

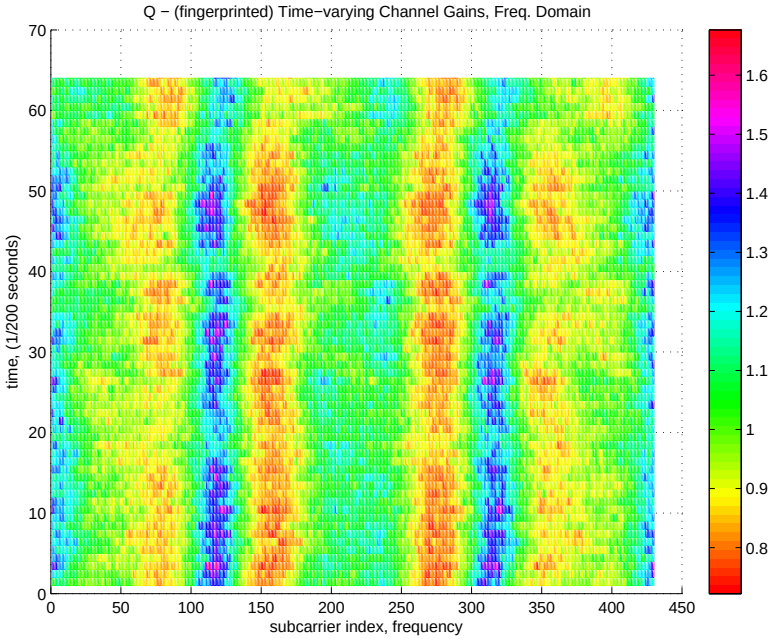
In this chapter we assume that eigenvalues for components $p+1$ through $P-1$ are all equal to σ_H^2. Clearly, as shown by experiment, the eigenvalues are decreasing, and many will be close enough to zero to be useful as 'extrinsic' eigenvectors. However, in this chapter we do not provide a general proof that eigenvalues for components $p+1$ through $P-1$ are all equal to σ_H^2, and some additional discussion on this topic may be in order. We do acknowledge that, in practice, the eigenvalues will not be exactly σ_H^2, however, a detailed explanation of effective rank and its relationship to real and modeled noisy signals is given in [18]. We note our fingerprinting system does not require that these diminishing singular values $p+1$ through $P-1$ be equal for correct operation. Because our fingerprint overlay is designed to occupy the noise subspace represented by the singular values $p+1$ through $P-1$, these values represent the power spectrum of the noise subspace. When $p+1$ through $P-1$ are not identically equal, the noise spectrum of the recovered fingerprint overlay will be nonuniform, but the fingerprint can still be recovered with nonuniform SNR.

3.8 Conclusion

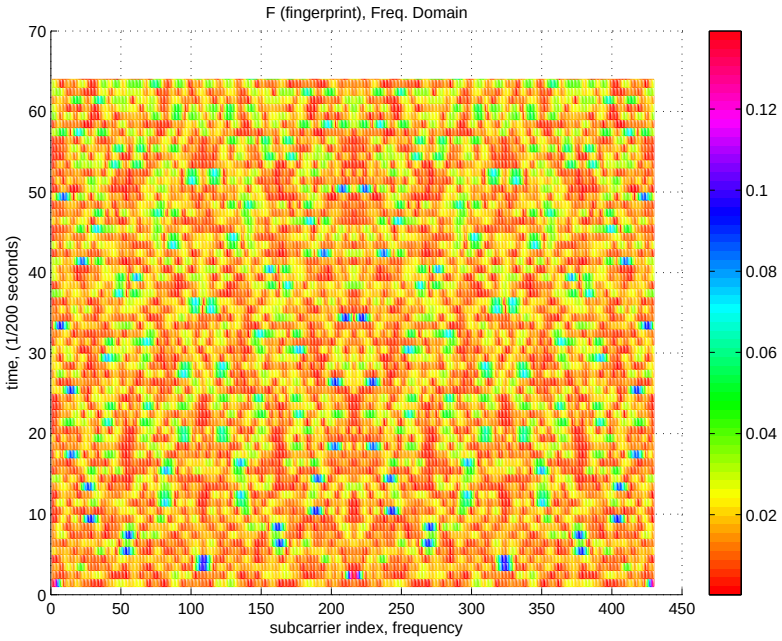
In this chapter we presented a new OFDM physical-layer fingerprint embedding scheme that incorporates previous channel-state information into the design of a overlay signaling basis. The transmitter embeds the fingerprint only onto the noise subspace of the wireless channel in a water-filling manner that maximizes the fingerprint capacity. We have demonstrated the embedding scheme through experimentation using real channel data collected from WiMax base stations, and the presented simulation results demonstrate that the BER of the primary signal is not influenced by the presence of the fingerprint. Also, the BER of the fingerprint signal outperforms the primary transmission by 20 dB, or more, in the channel conditions tested. Additionally, the proposed embedding scheme has demonstrated robustness to time-varying block-stationary fading.



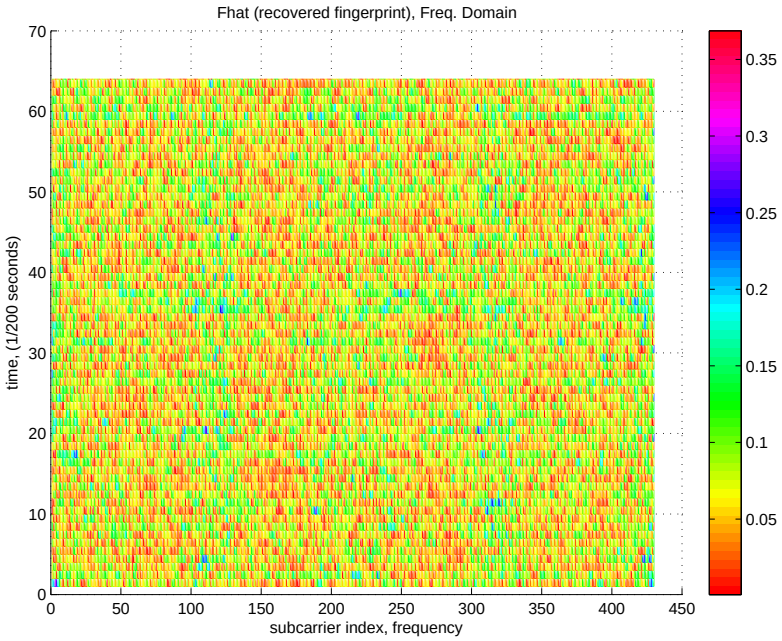
(a) Simulated Sounding Data



(b) Simulated Fingerprinted Data

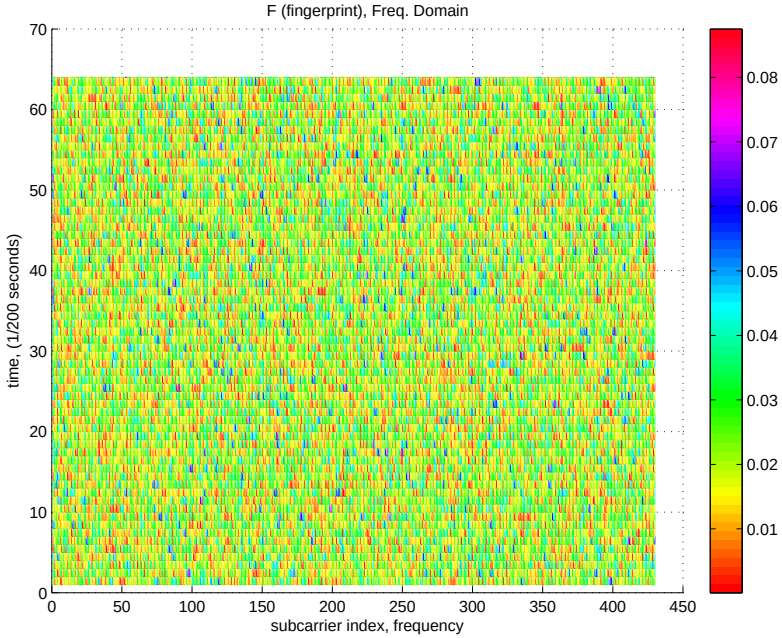


(c) Original Fingerprint

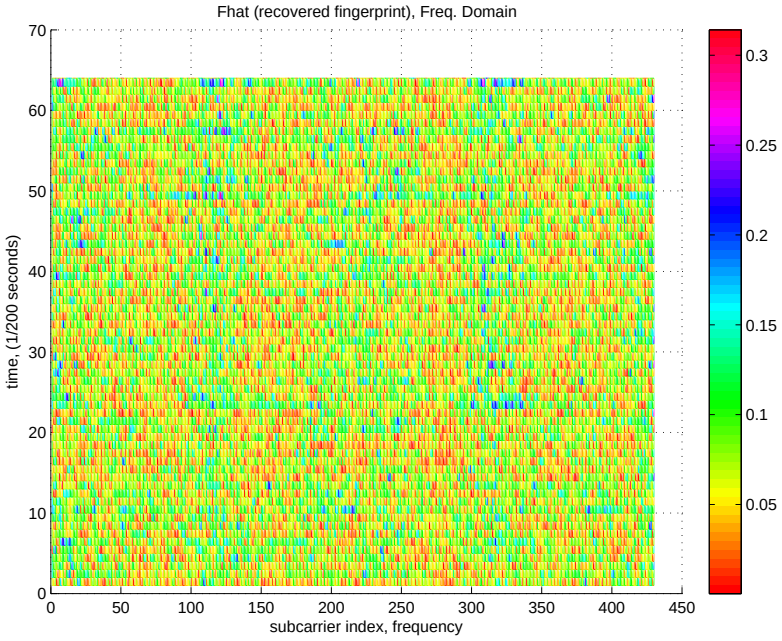


(d) Recovered Fingerprint

Figure 5: Simulated fingerprint embedding using the $\mathbb{K}_{hu,wu}$ overlay design with p=0

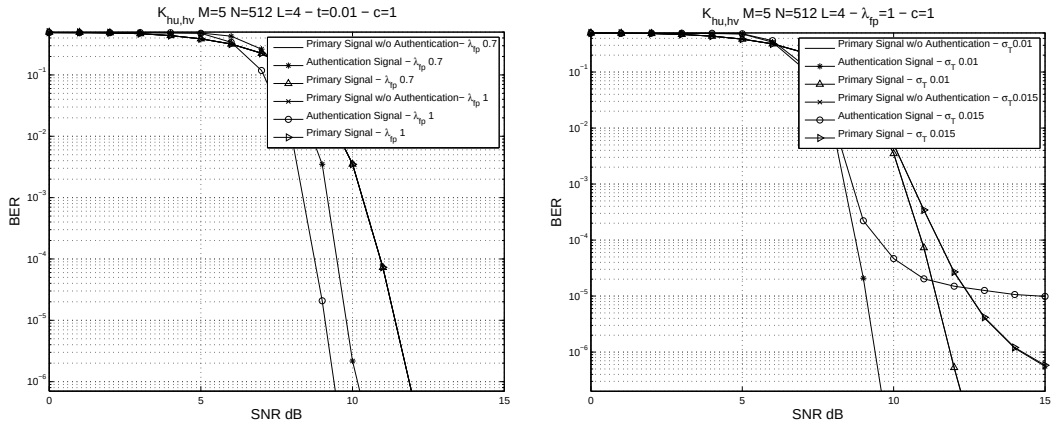


(a) Original Fingerprint



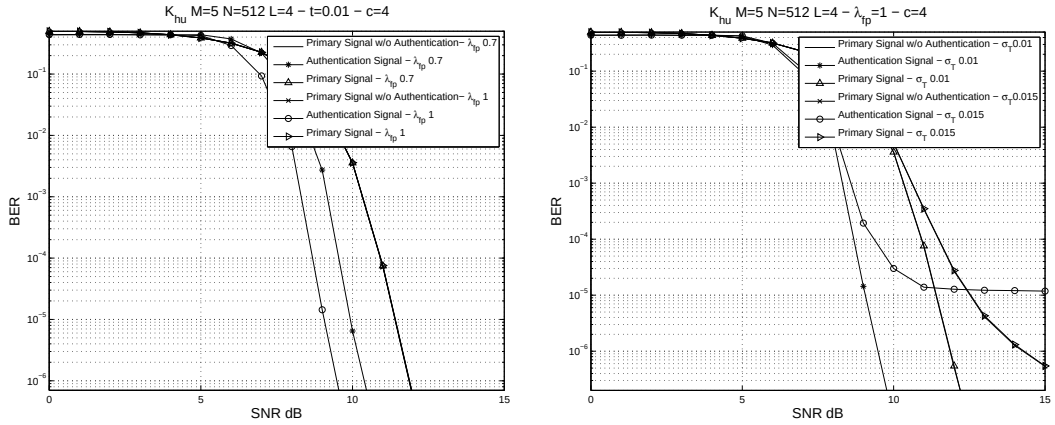
(b) Recovered Fingerprint

Figure 6: Simulated fingerprint embedding using the $\mathbb{K}_{hu}$ overlay design with p=8



(a) $\mathbb{K}_{hu,wu}$ design vs λ_{fp}

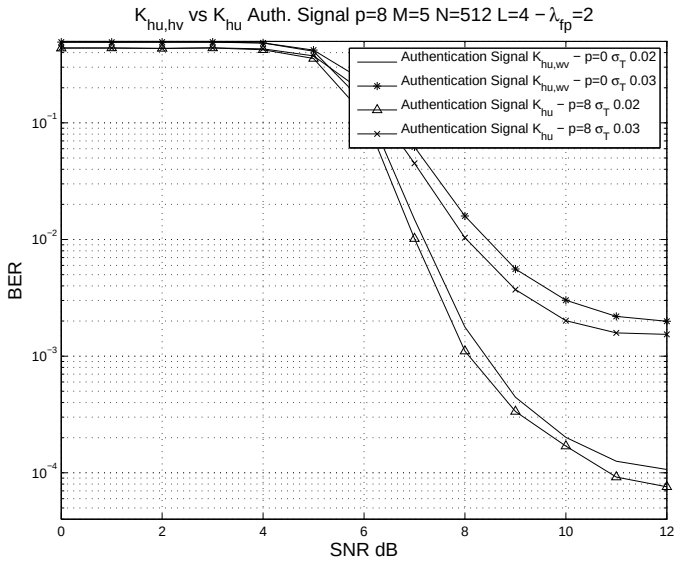
(b) $\mathbb{K}_{hu,wu}$ design vs σ_T^2



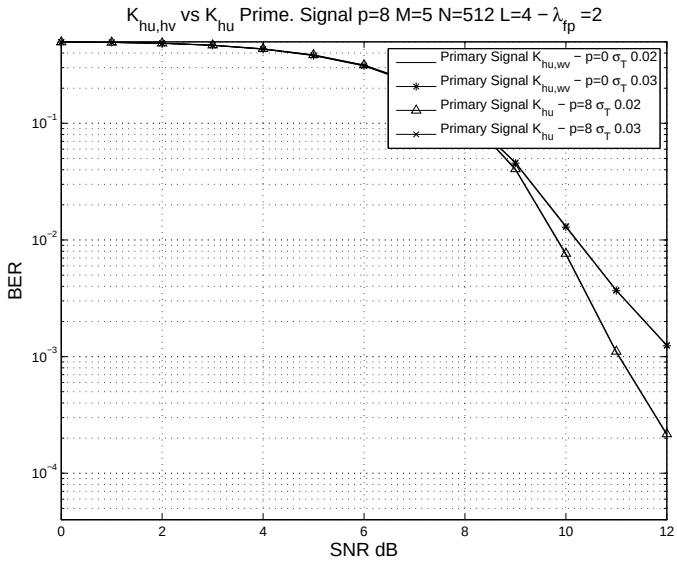
(c) $\mathbb{K}_{hu}$ design vs λ_{fp}

(d) $\mathbb{K}_{hu}$ design vs σ_T^2

Figure 7: BER of the primary transmission and the fingerprint signal with different schemes



(a) Fingerprint BER



(b) Primary BER

Figure 8: Comparison between $\mathbb{K}_{hu,wu}$ design and $\mathbb{K}_{hu}$

Chapter 4

Active Sensing for Dynamic Spectrum Access

4.1 Overview

We present a physical-layer mechanism for dynamic spectrum access (DSA) applications that takes an active approach to spectral sensing. In wireless shared spectrum scenarios, such as the Wireless Regional Area Networks (WRAN) proposed by IEEE 802.22, licensed digital television stations take a *primary user* role and are given explicit first-right-of-access to the television spectrum. When a primary user signal is not present, broadband users are allowed to use shared spectrum, and assume a *secondary user* role. While various detection and classification techniques using cyclostationary behavior, modulation characteristics, signal bandwidth, and spectral shape have been proposed to discern users of various roles in DSA theaters, these methods have limited detection performance, are susceptible to channel distortions such as multipath effects, and are vulnerable to malicious Primary User Emulation (PUE) attacks. In PUE attack scenarios, passive spectrum sensing techniques have poor performance or fail completely, when a secondary user mimics some or all of the features of a primary user. In this chapter we assume that a secondary user is capable of mimicking all features typically employed by passive identification techniques, and instead propose embedding an explicit unforgeable physical-layer fingerprint signature into wireless transmissions for disambiguating and authenticating users. When applied to orthogonal frequency-division multiplexing

signals in time-varying multipath channels, we demonstrate that our "channel-like" overlay fingerprint authentication scheme achieves 99.99 percent detection accuracy and 100 percent classification accuracy, with SNR as low as 5-6 dB.

4.2 Introduction

As wireless communication has become a ubiquitous part of every-day life, access to the electromagnetic spectrum has become increasingly competitive. To facilitate efficient use of limited spectral resources, an arbitration method known as Dynamic Spectrum Access (DSA) was recently proposed [4]. Inspired by a FCC ruling regarding Wireless Regional Area Networks (WRAN), IEEE 802.22 [20] has dramatically increased interest in spectrum sensing and shared spectrum technologies. Under IEEE 802.22, limited access to the unused spectrum between Digital Television (DTV) channels, or the white space spectrum, is granted to next-generation wireless broadband equipment. In particular, licensed DTV stations, or primary users , are given explicit first-right-of-access to television spectrum, while broadband users known as secondary users , are allowed access to the shared spectrum only when a primary user is not transmitting. While DSA shows promise in facilitating efficient spectrum access, accurate signal classification algorithms are required to facilitate the robust operation of next-generation wireless radios and the interoperability of DSA equipment.

To ensure efficient use of white space spectrum under IEEE 802.22, spectral allocations must first be tested to guarantee that primary users are not present before secondary users are granted access to an allocation. Since the accurate detection of primary users

is required for correct utilization of the shared spectrum, this difficult issue has become known as the Primary User Authentication (PUA) problem.

Traditional approaches to signal identification involving the computation of statistical properties [44] or cyclostationary features [22, 57, 68] have been proposed. These approaches can be considered *passive* signal characterization methods since the transmitter does not explicitly participate in the detection and classification process, nor does it modify characteristics of its signal to aid the detection and classification process. Classification approaches using these features in DSA scenarios have also been discussed, including machine learning [23] and policy-based classification engines [67]. These works have demonstrated the utility of machine learning approaches in signal classification applications; however, recent work [14, 48] has shown potential weaknesses in these approaches. In non-cooperative environments, adversaries can easily manipulate the learning process by fooling passive signal characterization methods, exposing DSA systems to a number of identity-based attacks.

While passive approaches readily admit to low complexity implementations, they are also prone to manipulation and forgery, allowing secondary users to masquerade as primary users by simply mimicking basic features of a primary user signal. Once a secondary user has been incorrectly classified as a primary user , the user can unfairly gain unfettered access to the spectrum. These attacks have become known as Primary User Emulation (PUE) attacks [13, 12], and to prevent malicious behavior the proposed physical authentication mechanism embeds an explicit and unambiguous (PHY) layer message into the signal to be transmitted, providing stronger signal identification and user authentication capabilities than those afforded by passive approaches.

We present a three faceted approach to creating a holistic signal authentication mechanism that can be applied to DSA theaters. First, we leverage the orthogonal frequency division multiplexing (OFDM) fingerprint overlay embedding scheme presented in [28] to embed a multi-bit, digital message at the PHY-layer, that is broadcast to all users within reception range of the transmission. Since the message is broadcast to all users, allowing every user to authenticate the fingerprinted transmission, we make no attempt to establish a covert or secret transmission. Secondly, we demonstrate that a digital message, embedded using the "channel-like" fingerprinting method, can achieve a bit error rate (BER) much lower than the original signal. Since we consider the embedding of multi-bit digital digital messages, we use BER of the received fingerprint message as a measure of performance. We describe an example digital authentication message to illustrate our authentication message, leveraging proven cryptographic primitives and best design practices in the security literature, and discuss how such mechanisms prevent PUE attacks. Lastly, we combine our PHY-layer approach with existing passive spectrum sensing techniques to create a hybrid spectrum sensing metric that outperforms passive sensing techniques, when passive techniques are used alone.

This chapter is organized as follows. Section 4.3 introduces our system model, briefly describes the fingerprinting method described in [28], and introduces the "two role" user model. Section 4.4 considers the bit-level content of the fingerprint message, and discusses how the cryptographic devices of the message prevent typical authentication attacks. In Section 4.5 we present simulation results for the proposed method, where we focus on disambiguating a homogeneous set of signals that would cause traditional passive classification techniques to fail. In Section 4.6 we present our conclusions.

4.3 System Model and Problem Formulation

We consider an OFDM system where the transmission is subjected to a linear time-domain channel response $g(t)$, given as

$$g(t) = \sum_{c} A_c(t)\delta(t - \Delta\tau_c), \tag{4.1}$$

where Δ is the sampling interval, τ_c are the delays for each channel component, and $A_c(t)$ are the complex valued delay-spread gains at time t for multipath component c. The OFDM system is modulated using an N-point discrete-time inverse Fourier transform ($IDFT_N$), and then subsequently demodulated using an N-point discrete-time Fourier transform (DFT_N).

While [28] presents a taxonomy of channel-like OFDM fingerprinting overlay designs using various amounts of previous channel state information (CSI) to improve fingerprint detection performance, in this work we leverage only overlay designs that require zero CSI knowledge, such as the $\mathbb{K}_{hu,wv}$ design, since these designs feature the lowest computational complexity, are the simplest to implement, and require zero CSI feedback overhead.

While the advantage of the $\mathbb{K}_{hu,hv}[l]$ design is that knowledge of previous CSI and CSI feedback is not required to design the overlay, the main disadvantage of this design is that linear time-varying channel distortions will interfere with the fingerprint signal, since knowledge of such distortions are not considered when designing the left and right spreading bases of the overlay.

4.3.1 Fingerprint Overlay Variance

In [28] it was demonstrated that an unbiased estimate for the fingerprint overlay $\mathbb{K}[l]$ can be obtained from the block aggregate channel response $\hat{\mathbb{Q}}[l]$ at the receiver. It can be shown that the variance for the recovered overlay estimate is

$$\mathrm{Var}\left[\hat{\mathbb{Q}}[l]\right] = \left(\mathbb{P}[l] + \mathbb{K}[l]\right)^2 \left(\frac{E\left[\hat{\mathbb{N}}^2\right] + \mathbb{L}^2[l]}{E\left[\hat{\mathbb{N}}^2\right] + \mathbb{L}^2[l-\epsilon]} - \frac{\hat{\mathbb{L}}^2[l]}{\hat{\mathbb{L}}^2[l-\epsilon]} \right), \tag{4.2}$$

where each element of $E\left[\hat{\mathbb{N}}^2\right]$ is the noise variance of the channel.

4.3.2 PHY-layer Threat Model

We assume that adversaries are capable of generating a transmission using the same methods that generate $\mathbf{Y}[k]$, and we further assume that adversaries have full knowledge of the fingerprint overlay design and how the fingerprinting function $\mathbf{f}[k]$ is applied, but do not have knowledge of the fingerprint signal vector $[\xi_p\ \xi_{p+1}\ \ldots\ \xi_{P-1}]_l$. In Section 4.4, we will give examples for digital multi-bit authentication messages that incorporate cryptographic security features to prevent forgery and reply of the authentication message to be signaled via $[\xi_p\ \xi_{p+1}\ \ldots\ \xi_{P-1}]_l$. Thus, adversaries are able to mimic all primary user features but are unable to forge the fingerprint signal vector $[\xi_p\ \xi_{p+1}\ \ldots\ \xi_{P-1}]_l$. When $\mathbb{K}[k]$ is designed to sufficiently spread the fingerprint over frequency and time, all signal characterization methods relying on passive features such as cyclostationary behavior, modulation characteristics, signal bandwidth, and spectral shape, will fail when attempting to discern users.

Fabrication of an authentic transmission would require forgery of $\mathbb{F}[k]$. We will protect the digital multi-bit authentication message conveyed by $[\xi_p\ \xi_{p+1}\ \ldots\ \xi_{P-1}]_l$ and

its associated $\mathbf{F}[k]$ from forgery and replay by using cryptographic primitives and best security design practices, to be discussed in Section 4.4. Thus, an attack attempting to forge $\mathbb{F}[k]$, and an authentic looking fingerprinted transmission $\mathbb{Y}[k]$, is not possible unless the cryptographic keys used to sign the digital multi-bit message conveyed by $[\xi_p\ \xi_{p+1}\ \ldots\ \xi_{P-1}]_l$ are compromised.

In Section 4.4 we will demonstrate that the digital multi-bit authentication message conveyed by $[\xi_p\ \xi_{p+1}\ \ldots\ \xi_{P-1}]_l$ is designed to incorporate self-verifying information about the transmitted signal $\mathbf{Y}[k]$. The entire multi-bit message conveyed by $[\xi_p\ \xi_{p+1}\ \ldots\ \xi_{P-1}]_l$ is then signed using cryptographic keys to authenticate the digital message and the self-verifying information fields. By including self-verifying information in the authentication message, if an adversary were to ever compromise the keys used to sign the digital multi-bit transmission conveyed by $[\xi_p\ \xi_{p+1}\ \ldots\ \xi_{P-1}]_l$, and is able to completely replicate $\mathbf{F}[k]$, a forged transmission would be limited to the frequencies and times prescribed by the compromised key since any deviation in these parameters would implicate the transmission of forgery.

4.4 The Digital Fingerprint Message

To modulate a multi-bit digital authentication message, the elements $[\xi_p\ \xi_{p+1}\ \ldots\ \xi_{P-1}]_l$ can be selected by the transmitter as symbols from a typical Pulse Amplitude Modulation (PAM) signal constellation, and using an appropriate bit-to-symbol mapping the receiver can recover the digital authentication message from $[\xi_p\ \xi_{p+1}\ \ldots\ \xi_{P-1}]_l$ after extracting these statistics. While the vector $[\xi_p\ \xi_{p+1}\ \ldots\ \xi_{P-1}]_l$ is only $P-p$ symbols long, the

concatenation a number of consecutive vectors over ω fingerprinted blocks will yield a digital authentication message of length $\rho = \omega(P - p)$ symbols long. For example, for the values $P = 128,\ p = 8\ \omega = 10$ a digital authentication message of $\rho = 1200$ bits can be employed. Recovery of the fingerprint signal vector $[\xi_p\ \xi_{p+1}\ \ldots\ \xi_{P-1}]$ at the receiver is discussed in [28], where the receiver projects the received equalizer channel estimate data onto the left and right signaling bases, $\left[\mathbf{0}\quad \mathbb{U}_2\right]_{l-\epsilon}^{H}$ and $\left[\mathbf{0}\quad \mathbb{V}_2\right]_{l-\epsilon}$, respectively, to de-spread the fingerprint signal.

4.4.1 Basic Authentication Message

To address the needs of DSA applications, the digital authentication message embedded in each node's transmission should contain basic self-verifying information such as the frequency, location, and time the signal is authorized for transmission. We will denote these fields as F, L, and T, respectively. A message hash of these parameters is then digitally signed using a secret key owned by the transmitter and included in the message, while a timestamp denoted TS is also included with the authentication message to prevent future replay of the message by malicious users. The timestamp allows for the enforcement of an expiration deadline on the content of the message, and in the event an authentication message is received with a timestamp that has passed the expiration deadline it will be discarded by the receiver. The authentication message for a primary user U_j, denoted $msg_{U_j,A}$, is given as

$$msg_{U_j,A} = \left\{TS, F, L, T, {K_A}^{+}, \left[Hash_m\left[TS, F, L, T\right]\right]_{K_A}\right\}, \tag{4.3}$$

where $[\cdot]_{K_A}$ is a digital signature of the content within $[\cdot]$ using the private key owned by the primary user group, the subscript A is used to denote that user U_j is a member of the primary user group A, ${K_A}^+$ is the public key of the primary users group, and $Hash_m[\cdot]$ is message digest of length m for the content within $[\cdot]$.

The hash algorithm $Hash_m[\cdot]$ can be any of a number of widely used collision-resistant hash algorithms, such as MD5 or SHA-1 [21], which provide reasonable security against malicious fabrication of messages. Depending on the details of the DSA application, an authentication message similar to (4.3) could be established for secondary users , or to reduce the implementation complexity of secondary user transmitters the fingerprint message $\mathbf{F}[k]$ could be omitted entirely from secondary user transmissions and thus any transmission that does not contain a valid primary user fingerprint will be classified as a secondary user .

4.4.2 Decoding of the Basic Authentication Message

To decode the authentication message, the receiver first recovers the embedded multi-bit digital fingerprint message and then extracts parameters from the payload of the message. Once each field has been extracted, the authenticity of the primary user group's public key, ${K_A}^+$, is verified from a mutually accepted trust anchor or certificate authority (CA). The receiver then independently verifies $[Hash_m[TS, F, L, T]]_{K_A}$ using the primary user group's public key, ${K_A}^+$, which is embedded in the authentication message. Malicious forgery of the authentication message is prevented through the signature process, and by including this signature as part of the authentication message. The modi-

fication of any subset of the authentication message parameters TS, F, L, and T, would cause the message signature to fail validation when it is received, enabling the receiver to detect and discard modified messages.

If the authenticity of ${K_A}^+$ and the message signature were both deemed valid, and the operating signal is within the specifications of F, L, and T, it will then be recognized as an primary user . This authentication messaging system relies on the existence of a *trust anchor*, sometimes referred to as a Certificate Authority (CA), to verify the authenticity of ${K_A}^+$. While primary users in the PUA problem posed by IEEE 802.22 are digital television transmissions adhering to the American Television Standards Committee (ATSC) specification which does not use OFDM signaling, in this chapter we consider DSA applications employing OFDM due to its widespread adoption in modern waveforms [35, 1].

The establishment of a trust anchor is still required for correct operation of the system. In the United States, for example, the FCC could establish a CA to sign certificates for individual licensed transmitters, and the address of the CA could be distributed to wireless devices requiring accurate authentication of DSA users. Smaller wireless networks could establish a similar centralized trust anchor for the purpose of validating user's public keys. A number of decentralized methods for establishing a CA have been discussed [38, 3] to overcome shortcomings of centralized schemes, however discussion of these methods are outside of the scope of this chapter .

4.4.3 Secure Group Entry

While an authentication message such as (4.3) provides a mechanism for nodes within an established system to differentiate transmissions by users of various roles, it does not address initial network entry, also known as the *bootstrapping problem*, that exists when a new user needs to enter the system for the first time. While in practice wireless networks with any number of user roles are conceivable, for the sake of discussion we consider the two-role wireless network consisting of a group of primary users , a group of secondary user nodes, and a designated CA. A secure system addressing the bootstrapping problem must provide a mechanism for nodes to securely enter the network as a member of one of these two groups. To ensure that the system is highly available, the bootstrapping mechanism should allow new nodes to enter into the network via a number of points of access.

We consider a replicated bootstrapping service, where a percentage of nodes within the network are designated as trusted *bootstrapping agents*, and are allowed to bootstrap new nodes into the network. We denote the group of bootstrapping agents as $Q_i,\ i = 0, \cdots, Q_N$, where Q_N is the number of bootstrapping agents distributed throughout the network. It is assumed that Q_i have been vetted by the network and CA, and are trusted agents for facilitating initial network entry. When a new network node U_j attempts to enter the network for the first time as a primary user , it first sends a bootstrapping broadcast request to locate one of the bootstrapping agents. To locate a bootstrapping agent, U_j sends a broadcast message, denoted $boot_{U_j,A}$, which is given as

$$boot_{U_j,A} = \left\{ \left[A, U_j, T, {K_A}^+ \left(RP, K_{U_j,A}\right)\right]_{U_j}, \left[U_j, {K_{U_j}}^+\right]_{CA} \right\}, \tag{4.4}$$

where the subscript A denotes that the user U_j is attempting to join the primary user group, T is a message timestamp designed to establish an expiration deadline to validate the message, $[\cdot]_{CA}$ denotes that the enclosed fields have been signed by the CA, $K_{U_j,A}$ is the shared key that U_j wants to securely transmit to the group, and RP is a secret *Replay Pad* which will be used by Q_i to encrypt $K_{U_j,A}$ in subsequent messages. The $\left[U_j, {K_{U_j}}^+\right]_{CA}$ component becomes a certificate denoting that the group's trust anchor has verified the authenticity of ${U_j}'s$ public key.

In this example basic message, ${K_A}^+ \left(RP, K_{U_j,A}\right)$ denotes that the relay pad RP and ${U_j}'s$ primary user shared key $K_{U_j,A}$ have been encrypted with the *primary users group* public key ${K_A}^+$, therefore only the group's private key ${K_A}^-$ can decrypt these fields. By encrypting this message using the group's public key, node U_j is protected from malicious key exchange attacks, where an attacker poses as an authentic bootstrapping node. The proceeding steps of the bootstrapping process will securely convey ${K_A}^-$ to the joining user, allowing U_j to decrypt and use ${K_A}^-$ in future transmissions.

When one of the bootstrapping nodes, Q_i, eventually receives a bootstrapping request and verifies that U_j is allowed to join group A, it will respond to node U_j and proceed to the next step of the bootstrapping process. The reply message sent by Q_i, denoted $bootreply_{Q_i,U_j}$, becomes

$$bootreply_{Q_i,U_j} = \left[U_j, N, CK_A \oplus RP, CK_A \left({K_A}^-\right)\right]_{Q_i}, \tag{4.5}$$

where N is a nonce, or a randomly generated piece of information, CK_A is the group's *shared secret* which will be used to encrypt the primary user group's private key, ${K_A}^-$ is the primary user 's group's private key, $CK_A \oplus RP$ denotes that the group's shared secret

CK_A has been xor'ed with the replay pad RP, and $[\cdot]_{Q_i}$ denotes that the enclosed fields have been signed with $Q_i's$ private certificate. By applying the xor operation to CK_A and RP, Q_i can securely transmit the group shared secret CK_A to U_j. The nonce N will be used in subsequent transmissions to verify that U_j has successfully decrypted $K_A{}^-$.

Once U_j receives the $bootreply_{Q_i,U_j}$ message, it uses the RP which it has secretly retained to recover the group shared secret CK_A. After CK_A has been recovered, node U_j can use the group's shared secret to decrypt the group's private key $K_A{}^-$, which it will use in future group communications. To notify that Q_i that the $bootreply_{Q_i,U_j}$ has been successfully processed, U_j will send the final confirmation message, $bootconfirm_{U_j,Q_i}$, to Q_i, i.e.

$$bootconfirm_{U_j,Q_i} = K_{U_j,A}\,(N), \tag{4.6}$$

which is simply the nonce N from the $bootreply_{Q_i,U_j}$ message encrypted with $U_j's$ primary user shared key, $K_{U_j,A}$, which Q_i has had in it's possession since the $boot_{U_j,A}$ message. After receiving this message, Q_i decrypts the nonce, and compares it to the value N which it sent to U_j in (4.5). If the two values match, Q_i now has confirmation that U_j has successfully joined the group.

The use of the nonce ties $U_j's$ confirm message, $bootconfirm_{U_j,Q_i}$, to $bootreply_{Q_i,U_j}$, and together with use of the replay pad between the $boot_{U_j,A}$ and $bootreply_{Q_i,U_j}$ all of the messages in the exchange are tied together as one authentic bootstrapping session.

Using the given secure group entry method, any number of groups representing various roles in a DSA theater could be derived. For the two role system, another secondary user group UA could be created, in addition to the primary user group A in the previous

example.

4.4.4 Message Security Evaluation

To be considered secure, cryptographic protocols need to be robust against forgery, modification, deletion, and replay. Since we are considering a broadcast authentication system, where every user is able to decode and subsequently verify the fingerprint message, all notions of privacy are non-applicable since we want every user to have the ability to extract the authentication message.

Provided the hashing algorithm used is collision-resistant, forgery is not possible. Modification of (4.3) and the self-verifying fields is prevented through use of the signature itself. If any of the fields, TS, F, L, T, or K_A are changed, then the signed hash is no longer valid. By leveraging proven cryptographic primitives in the design of keys, message signatures, and message hashes, the probability of making an authentication error is reduced to the probability of a hash collision. A well designed hash algorithm such as SHA-1 will feature a collision probability which is nearly zero in all practical applications, thus preventing the acceptance of incorrect authentication messages. For example, when using a 64-bit message hash a malicious node would require approximately 5.1×10^9 attempts to achieve one collision using a brute force 'birthday' attack. Current best practices when using secure hashing algorithms suggest using at least a 256-bit hash, i.e. SHA-256, further decreasing the probability of an authentication error and making the probability of accepting a forgery in the unlikely event an attacker were to fabricate a hash collision, virtually impossible.

Since the authentication message $msg_{U_j,A}$ is transmitted as a multi-bit digital signal, the probability of a fingerprint detection miss is the same as the probability of receiving the entire authentication message with one or more bit errors. Since a single bit error in either the authentication message or the signature will cause the authentication to fail, the probability of missing the authentication message is the same as the probability of a at least one bit error in the message. Therefore for an uncoded binary transmission, the probability that the received authentication message is received in error is simply

$$P[\hat{msg}_{U_j,A} \neq msg_{U_j,A}] = 1 - (1 - P_e)^{M+N}, \tag{4.7}$$

where P_e is the probability of a bit error in the authentication signal, $M = length\{TS, F, L, T\}$ and $N = length\{[Hash_m\,[TS, F, L, T]]_{K_A}\}$. The use of forward error correction (FEC) on the authentication signal, combined with a continuously repeated message (i.e. repetition encoding), can further decrease the probability of an authentication miss.

The authentication message in (4.3) also includes the frequency F that the transmitter is allowed to transmit on, which would presumably be associated with the transmitter's key and recorded by a CA like the FCC. Therefore even if we assume that an adversary can compromise an primary user 's key and forge $\mathbf{F}[k]$ at the PHY-layer, the attacker will be constrained to the frequency or frequencies prescribed by the compromised key. Using a forged $\mathbf{F}[k]$ on a frequency other than the original frequency prescribed by the key will implicate the transmission as a forgery when validating the credentials of the key against the CA's records.

In the secure group entry method discussed in Section (4.4.3), the authenticity of

messages are guaranteed using signatures and shared keys, while session coherence and replay attacks are prevented through the use of the replay pad and nonce.

4.5 Simulation Results

In this section we present simulation results for the $\mathbb{K}_{hu,wv}$ fingerprint overlay design, using the intrinsic time-varying channel model described in [64] and applying the same assumptions and parameters as used in [28]. Through simulation we quantitatively compare the detection performance for primary and secondary user signals in complex time-varying channels, and measure any degradation experienced by each user's signal due to the presence of the embedded fingerprint. To measure user signal degradation, we compare the BERs of the primary signal with and without the fingerprint present.

To simulate the embedding of both user fingerprint signals, a full OFDM system and accompanying channel simulator were created in Matlab. The generated OFDM signal uses a 512-point FFT with 430 occupied subcarriers and 41 left and right guard subcarriers. A BPSK-modulated preamble occupies the first symbol of each frame using a $2\times$ time-domain repetition, while the following payload symbols are modulated using QPSK. To allow for periodic sampling of the intrinsic channel response, the odd-even differential modulation scheme discussed in [28] was used, where odd frames are fingerprinted and even frames are left unfingerprinted to aid in channel sounding; a process that is required to measure and reverse distortions of the intrinsic time-varying channel.

Each OFDM frame was then subjected to a simulated time-varying channel by applying $g(t)$ in the time domain using a transversal filter, according to (4.1). Timing jitter

was also added to test the fingerprint's affect on typical frame synchronization algorithms, and the Schmidl and Cox algorithm [51] was used for course timing recovery. The LS channel estimator was applied to estimate the aggregate channel distortion using the frame preamble as training data, and the resulting estimate was then up-sampled using $2\times$ sinc interpolation.

A number of channel estimation techniques have been considered for OFDM systems, including the minimum mean-squared error (MMSE) estimator, and the least-squares (LS) estimator. These estimators, with some improvements, are discussed in [59] for the system model described above. A discussion of particular channel estimation techniques is beyond the scope of this chapter , therefore without loss of generality, we use least-square (LS) channel estimator [59] in Section 4.5, due to its widespread adoption in OFDM systems, ease of implementation, and low computational complexity.

The channel sounding symbol and fingerprinted symbol were then equalized independently, and $\hat{\mathbf{q}}[l-\epsilon]$ and $\hat{\mathbf{h}}[l]$ were recorded for each frame for later use during the fingerprint recovery phase. Both the BPSK preamble data and QPSK payload data are demodulated into bits, and the bit errors for fingerprinted and non-fingerprinted frames were recorded.

We model a DSA system using the two-node model consisting of primary and secondary users . Primary users embed the message $msg_{P_j,A},\ j=1,\ldots,N_A$ into their transmission, where N_A are the number of users in the primary users group A, while secondary users embed $msg_{P_k,U},\ k=1,\ldots,N_U$ into their transmission, where N_U is the number of users in the secondary user group. Two other model parameters, namely the fingerprint strength, denoted λ, and the excitation noise variance of the time-varying channel pro-

cess, σ_T, are both discussed in depth in [28]. A detailed discussion of these parameters is omitted here for space considerations, therefore we refer the reader to this work for a detailed channel model and description of these parameters.

In consideration of a worst-case PUE attack scenario, all other aspects of both the primary user and secondary user 's transmissions, such as the number of subcarriers, the bandwidth of the transmission, signal strength, modulation details, and preamble structure, are identical, while the exact contents of the multi-bit digital fingerprint message transmitted by each user is different and assumed to be unknown by the opposing group.

We assume that secure group assignment has been performed, therefore cryptographic keys have been securely transfered such that secondary users cannot fabricate a primary user 's message, and vice-versa. While a message containing all fields suggested by (4.3) might be several hundred bits long, for the sake of discussion we select both the primary user message, $msg_{U_j,A}$, and the secondary user message, $msg_{U_k,UA}$, to be $\rho = 224$ bit sequences generated by a pseudo-random number (PN) generator, therefore making both messages completely uncorrelated PN sequences.

To correctly detect either $msg_{P_j,A}$ or $msg_{P_k,UA}$ we require that the entire 224 bit sequence be received without bit errors, while reception of a message with at least one bit error will constitute a detection miss. While longer authentication messages may require use of forward error correction (FEC) to ensure that the entire message is received without bit errors, FEC was not considered in our simulations.

The BER plots for the fingerprinted signal with and without the fingerprint present, for fingerprint strength values $\lambda = [2,\ 6]$ and time-varying channel excitation noise variance $\sigma_T = .001$, are depicted in Figure 4.5, along with the BER of the authentication

fingerprint signal. From Figure 4.5 we note that the authentication signal is received with a substantial BER advantage over the user signal, and that this advantage increases with the strength of the fingerprint, i.e as λ increases. Detection performance for the 224 bit

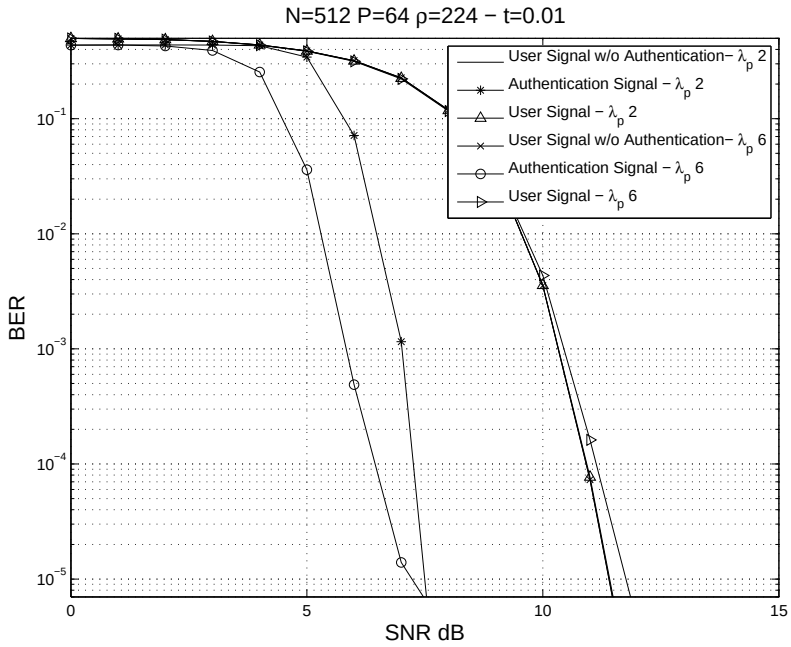


Figure 1: Original Signal and Auth. Signal BER with and without Fingerprint Present, various λ_i

authentication message transmitted by both the primary and secondary users is depicted in Figure 4.5, for values of $\lambda = [2,\ 6]$. For comparison, we also include the message detection performance for a message of the same length sent via the payload of the OFDM signal. We note from Figure 4.5 that the detection performance for both users is identical, as the series for the primary and secondary users are completely overlapping for both values of λ. We also note that the detection performance increases with respect to λ, and that increasing the strength of the fingerprint, by increasing λ, allows the fingerprint to be detected in lower SNR. From Figure 4.5 we notice that the threshold between the 0% de-

tection rate and 100% detection rate is very steep, and that this threshold occurs between 5 and 7 dB SNR, for the values of λ chosen. This proves that the signal can be authenticated nearly 100 percent of the time, in SNR conditions as low as $6 - 7$ dB. Additionally we note that the authentication message sent using fingerprint embedding outperforms a similar message sent via the payload of the signal by a margin of nearly 5 dB.

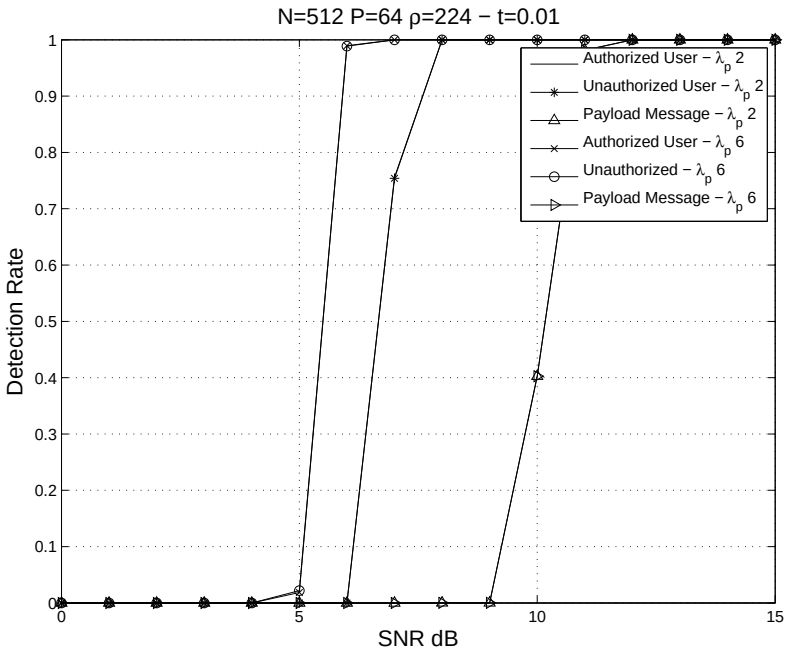


Figure 2: Detection Rate of Primary and Secondary Users for Various λ_i, (overlapping series)

For comparison, we plot the message detection rate results of 4.5 with predicted message detection results based on the BER of the authentication signal using (4.7). These results are depicted in 4.5, where we observe that the simulated detection rates are very close to those predicted using BER rates.

Detection performance for both user messages is also depicted in Figure 4.5, for values of $\sigma_t = [.01,\ .015]$. From 4.5 we see that the detection performance for both

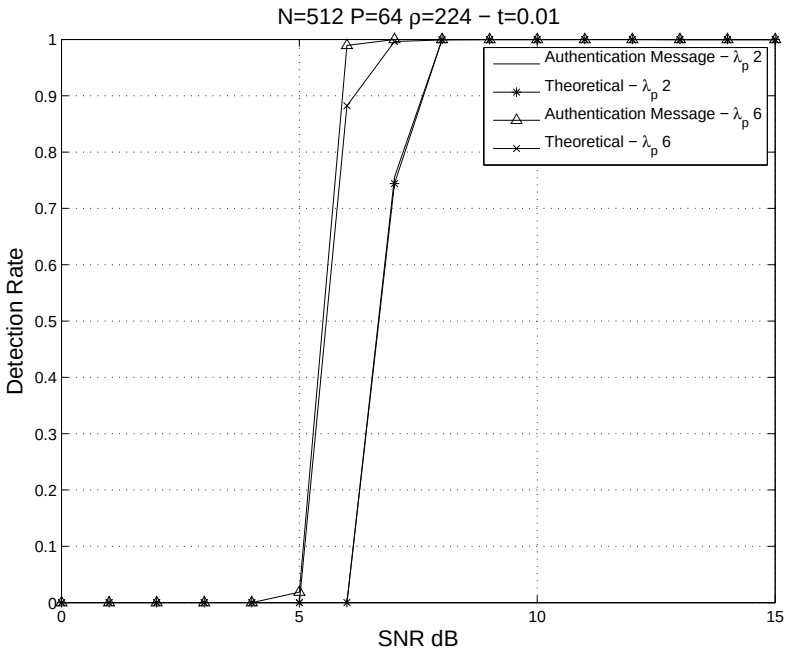


Figure 3: Simulated and Predicted Detection Rates for Various λ_i, (overlapping series)

users is again identical, as the series for the primary and secondary users are completely overlapping for both values of σ_T, and we also note that the detection performance is relatively invariant of σ_T, for the values of chosen. We note from 4.5 that the threshold between the 0% detection rate and 100% detection rate is again very steep, and that this threshold occurs at 6 dB SNR, for both values of σ_T. We again include the message detection performance for a message of the same length sent via the payload of the OFDM signal, and note that in this case the fingerprint authentication message detection rate outperforms the payload authentication message by 4 dB.

We plot the message detection rate results of 4.5 with predicted message detection results based on BER, observing that the simulated detection rates are also very close to the predicted values.

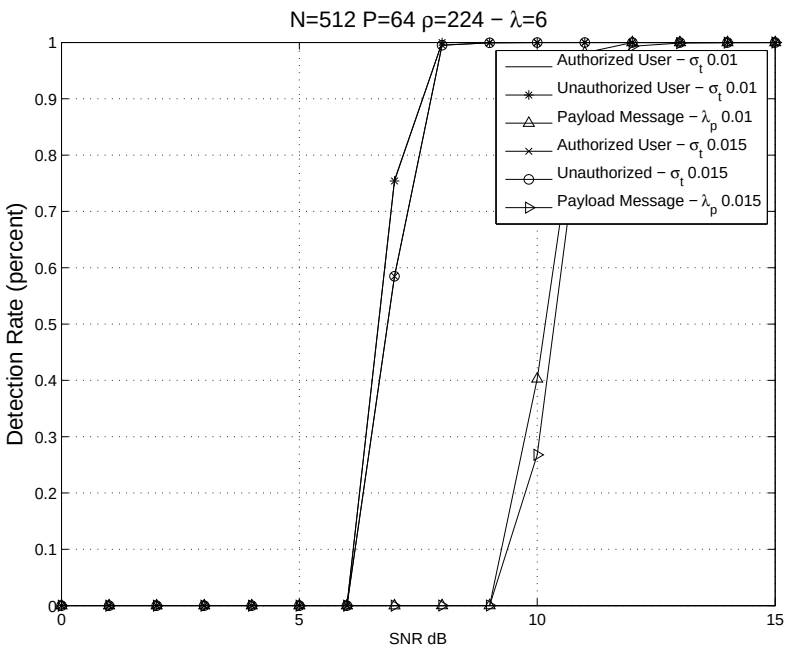


Figure 4: Detection Rate of Primary and Secondary Users for Various σ_t, (overlapping series)

4.6 Conclusion

In this chapter we have discussed a holistic authentication system for OFDM wireless transmissions that leverages channel-like fingerprint embedding techniques. We have extended the basic embedding approach by considering a digital authentication message for the dual-role system, to illustrate our approach. Additionally, the bootstrapping problem was addressed by presenting a mechanism for providing secure group entry. A detailed analysis of the security features of the authentication message, and the secure group entry messages, were discussed. Simulation results were presented for detecting the authentication fingerprint message for the two-role system. We have shown that our authentication scheme achieves 99.99 percent detection accuracy and 100 percent classification

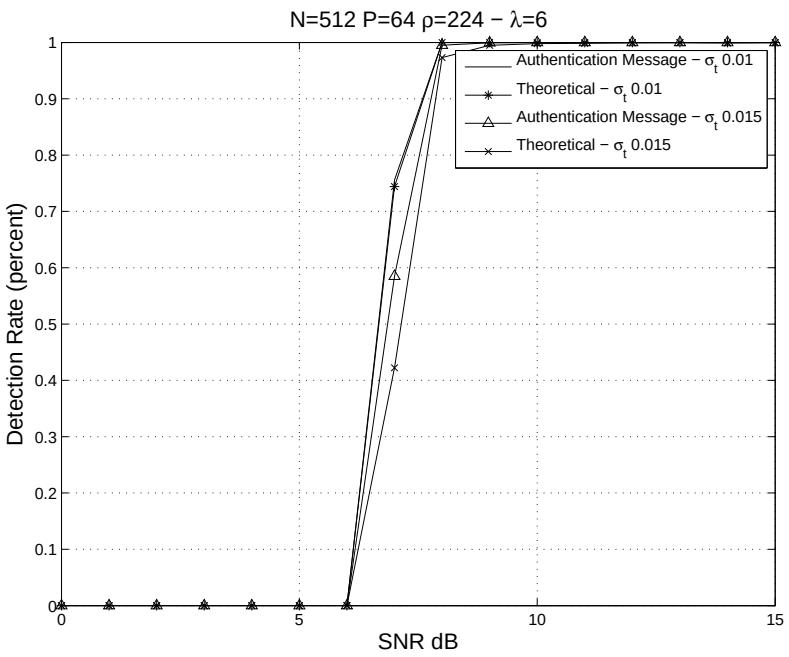


Figure 5: Simulated and Predicted Detection Rates for Various σ_t, (overlapping series)

accuracy, with SNR as low as 5-6 dB.

Chapter 5

Best-Effort Cooperative Relaying

5.1 Overview

Traditional cooperative communications consider dedicated-relays, while often such relays may not be available. In this chapter, we consider wireless transceivers that relay signals in addition to their own primary communication mission. We consider a best-effort delivery policy, where a node is not obligated to devote energy to cooperatively relay signals, nor does it provide a guarantee of signal quality on retransmissions. Instead the relay sacrifices energy at its own discretion, with priority given to the primary communication mission. We consider one best-effort delivery problem: a system that transmits an additional relay signal within its original transmission energy budget while inducing minimal degradation to the primary-user's signal. To maintain this constraint, we consider the feasibility of reallocating energy from pilot signals used for channel estimation toward the relaying service, when channel conditions are stationary. We demonstrate that transmitter energy may be dynamically allocated between a relay component and a pilot component of the transmission using best-effort delivery. This power allocation is critical to system performance, since both the primary-user and the secondary-user may require pilot energy to correctly decode transmitted signals. Sub-optimal power allocation rules with respect to primary-user channel estimate mean-square error and pairwise error probability are derived.

5.2 Introduction

Cooperative communication through the judicious use of relay-nodes has proven extremely advantageous in decreasing outage probabilities, [45, 42] and frame error rates [43] when the channel between the source and destination is of poor quality. Traditional relay schemes typically consider only nodes dedicated to relaying signals [15]. However, when a dedicated-relay is not available, nodes with their own communication mission must assist in relaying signals. We consider a relaying system where nodes provide relay services to a secondary-user in addition to their own primary communications, and the relay service is provided on a best-effort basis. Under the proposed scheme, a node with its own communication mission may also offer relay services while maintaining its original energy constraint. We extend the traditional three node relay problem considered in [15] to include the secondary-user node, and focus on the power allocation problem presented to the relay in this scenario. The best-effort relay system and the relationship between the secondary-user, or relay-user, and primary-user channels is depicted in Fig. 1, where the three node system of [15] is augmented to include the secondary-user.

The terms primary-user and secondary-user have been used numerous times in the literature to describe precedence roles in cognitive radio and cooperative communications [12]. When a constrained resource is shared by a number of users, it is the primary-users that typically have first-right-of-access to the shared resource, and are given higher priority than the secondary-users. Our use of the terms in this work is similar.

As a motivating example for adaptive best-effort relaying we consider the scenario of stationary channel conditions for the primary-user channel, when less energy is re-

quired for channel estimation purposes. A number of studies have attempted to characterize the correlation of time-varying channel estimates, including the works of [34] and [60]. These studies have shown that in fixed and low-mobility scenarios consecutive channel estimates in multipath scenarios are highly correlated. Under these conditions the relay may choose to devote more power toward relaying services and less power toward pilot signals for the primary-user channel, according to the quality of service (QoS) requirements of the primary-user.

When channel conditions require additional channel estimation energy, as is the case when the primary-user channel is undergoing change, the relay may choose to allocate more power toward the pilot component of the transmission for channel estimation purposes and less power for relay transmissions. The proposed best-effort delivery method introduces a unique power allocation problem in which the relay must select optimal power settings for the pilot and relayed signals. Power allocation for the relayed transmission is not guaranteed, and is rationed to aid the secondary-user only at the discretion of the primary-user. Thus, energy for the purpose of cooperative diversity through the relaying of signals is scavenged, when feasible, and the relayed transmissions are delivered on a best-effort basis.

Traditional pilot-aided channel estimation techniques, such as pilot symbol assisted modulation (PSAM), are discussed in [7]. Channel estimation schemes like PSAM create a composite signal consisting of two components: pilot signals used for channel estimation, and the user's data signal. To implement the relay service we instead consider a composite space-time code (STC) composed three components: the primary-user data signal, the pilot signal, and the relayed signal destined for the secondary-user. The density

of the pilot signals and bandwidth resources for each component are assumed to be preallocated and constant, as we focus solely on the power allocation subproblem in this work. Also, we consider the case where the transmission power used for the primary-user data signal will remain constant in the composite signal; however the power devoted to the pilot signal component and the secondary-user component will be dynamically allocated.

While many pilot-embedding techniques have been proposed before, we formulate our best-effort delivery channel using the pilot-embedding framework proposed in [50], which generalizes how pilot signals may be embedded into data signals using a STC approach. With this embedding method, mutually orthogonal pilot and data signals are combined as a composite STC block before transmission. The STC approach is used as a general method for generating a sequence of symbols with periodically occurring pilot signals comprising a single block code, and allows for maximum flexibility in the design of pilot-aided transmissions for systems with one or more transmit antenna. The method easily extends to broadband signals such as OFDM, through the Space Time Frequency (STF) block code model [55]. Here, we investigate best-effort cooperative relaying using generic pilot embedding when the data-component of a data-bearing pilot block is reserved for relay transmissions.

Cooperative and Cognitive Radio (CR) systems can be complex systems faced with a number of resource allocation problems. Nodes that choose to cooperate must constantly ration resources such as transmission energy and bandwidth when these resources are limited [52, 41, 33]. While the rationing of bandwidth, for example, is an important resource allocation problem, in this chapterwe focus solely on the four node power allocation subproblem suggested by Fig. 1 and leave the extension of this subproblem to

multi-node systems where bandwidth may also be rationed, as future work. The important cooperative system reciprocity issues of altruism and avariciousness are also not considered here; therefore, the motivations of the relay for aiding the secondary user, while interesting, are not applicable as this work focuses on cooperation once the decision to cooperate has been made.

In this chapter, we discuss a number of optimization problems that arise when considering the best-effort delivery of signals, and derive sub-optimal power allocation policies with respect to the upper-bound mean squared error (MSE) of the channel estimate, and the pairwise error probability (PEP) of the primary-user. The power allocation problem and general feasibility of best-effort delivery are considered using these two QoS criterion. We extend our work in [31] to consider the MSE-based power allocation rule and the total bit error rate (BER) of the system, comparing the performance of both allocation rules in greater detail. Additionally, sub-optimal power allocation considering the aggregate capacity of both the primary and secondary users is discussed.

This chapteris structured as follows: Section 5.3 generalizes data-bearing pilot frameworks (DBPF) and briefly introduces previous power-allocation work with secondary-channels. In Section 5.4 the best-effort relay problem is discussed, and a model for analytic system design is presented. Sub-optimal power allocation policies, with respect to the primary-user, are derived taking the upper-bound channel estimate MSE and PEP into consideration. In Section 5.5 simulation results are presented, and in Section 5.6 we give some concluding remarks.

5.3 System Model and Problem Formulation

We briefly describe the channel model used in our best-effort delivery scheme. We assume all nodes are MIMO systems with L_t transmit antenna(s) and L_r receive antenna(s) and use a Space Time (ST) block scheme, where the block transmitted at time index t is described by matrix $\mathbf{U}(t)$ of size $L_t \times M$. The ST block $\mathbf{U}(t)$ is a composite signal composed three components: the primary-user data signal, the pilot signal, and the relayed signal destined for the secondary-user. The composite signal is transmitted across all L_t transmit antennas in M time slots, and is broadcast to both the primary and secondary users. The received block $\mathbf{Y}(t)$ expressed in matrix form is

$$\mathbf{Y}(t) = \mathbf{H}(t)\mathbf{U}(t) + \mathbf{N}(t), \tag{5.1}$$

with the channel coefficient matrix $\mathbf{H}(t)$ of size $L_r \times L_t$ describing the channel conditions experienced by the block at time t. The channel noise $\mathbf{N}(t)$ is modeled as complex white Gaussian noise with zero mean and variance $(\sigma^2/2)\mathbf{I}_{(L_r \times M)}$. We assume Rayleigh block fading for the elements of $\mathbf{H}(t)$, where the entries of $\mathbf{H}(t)$ are independent zero mean complex Gaussian random variables that remain constant over each symbol block. The channel estimation problem is to estimate $\mathbf{H}(t)$ and recover the original ST block $\mathbf{U}(t)$, where the channel estimate is computed from the pilot / training signals. The pilot or training signal is embedded in the original transmission $\mathbf{U}(t)$, using a generalized embedding scheme to be described as follows.

We now briefly describe the pilot-embedding framework presented in [50], which provides the edifice for the design of $\mathbf{U}(t)$. The ST block to be transmitted is given as

$$\mathbf{U}(t) = \mathbf{D}(t)\mathbf{A} + \mathbf{P}, \tag{5.2}$$

where $\mathbf{P} \in \mathcal{R}^{L_t \times M}$ is the pilot matrix, ST block data-bearer matrix $\mathbf{D}(t) \in \mathcal{C}^{L_t \times N}$, and data-projection matrix $\mathbf{A} \in \mathcal{R}^{N \times M}$. Here, N is the number of time slots reserved for data transmission, while time slots $M - N,\ N < M$ are reserved for the embedded pilot signals and the relayed signal for the secondary-user. The pilot-embedding structures discussed in [50] have a bandwidth efficiency which is proportional to $(M - L_t)/M$ for the case $M = N + L_t$. The ST channel matrix is assumed to follow the constraint $E[||\mathbf{H}(t)||]^2 = L_t$, or constant energy under the Frobenius norm. The salient point of the data-bearing framework is that pilot-embedding schemes may be generalized through the superposition of the data-bearing structure $\mathbf{D}(t)\mathbf{A}$ and the pilot matrix $\mathbf{P}(t)$. The data-projection and pilot matrix satisfy the following properties:

$$\mathbf{AP}^T = \mathbf{0} \in \mathcal{R}^{N \times L_t}\ ,\quad \mathbf{AA}^T = \beta\mathbf{I} \in \mathcal{R}^{N \times N}\ ,\quad \mathbf{PP}^T = \alpha\mathbf{I} \in \mathcal{R}^{L_t \times L_t}. \tag{5.3}$$

where β is the real-valued power of the relayed signal, and α is a real-valued power coefficient of the pilot portion of the signal. The α coefficient will become an important parameter in the analysis to come, as it represents the fraction of divertable power that is retained for pilot signals. Since the power allocated to the relay and pilot signals is allocated dynamically, it is assumed that both the primary and secondary user receivers employ either independent gain control for each signal, or convey the value of α used for each block to the receivers using a control channel signal that is not considered here.

We assume that the number of time slots M in the ST transmission $\mathbf{U}(t)$ remains constant, and the number of time slots used for data transmission N also remains constant. Using this criteria the number of time slots available for best-effort transmissions also remains constant.

The properties (5.3) of the data-projection matrix **A** and **P** essentially allow **A** to project the data component $\mathbf{D}(t)$ onto the orthogonal subspace of the pilot matrix **P**, allowing for signal demodulation by means of a Maximum Likelihood (ML) receiver. These properties imply that $Rank(\mathbf{A}) = N$, $Rank(\mathbf{P}) = L_t$, and the number of time slots M required of the ST symbol $\mathbf{U}(t)$ is $M = Rank(\mathbf{A}) + Rank(\mathbf{P})$. The pilot structures of (5.3) operate under a power constraint

$$\beta = P_p - \alpha, \tag{5.4}$$

with P_p being the original normalized block transmission power of the node. According to (5.4) the power allocated to the composite signal consisting of the pilot plus secondary-user data signal is equal to the power of the original node transmission. To understand the behavior of the power allocation term α in terms of secondary-user performance, we first note that the normalized block power may be expressed as

$$\begin{aligned} P_p &= \frac{E\left[||\mathbf{U}(t)||^2\right]}{L_t} = \frac{E\left[||\mathbf{D}(t)\mathbf{A}||^2\right]}{L_t} + \frac{E\left[||\mathbf{P}||^2\right]}{L_t}, \\ &= \beta + \alpha = 1. \end{aligned} \tag{5.5}$$

The signal at the receiver becomes

$$\mathbf{Z}(t) = \mathbf{H}(t)(\mathbf{D}(t)\mathbf{A} + \mathbf{P}) + \mathbf{N}(t), \tag{5.6}$$

Three basic structures are discussed in [50] for the design of **A** and **P**, including the Time-Multiplexed (TM) structure which generalizes the previous PSAM pilot embedding techniques. The TM structure, which will be used later in simulation, is given as

$$\begin{aligned} \mathbf{A} &= \sqrt{\beta}\left[\mathbf{0}_{(N\times L_t)}; \mathbf{I}_{(N\times N)}\right] \\ \mathbf{P} &= \sqrt{\alpha}\left[\mathbf{I}_{(L_t\times L_t)}; \mathbf{0}_{(L_t\times N)}\right], \end{aligned} \tag{5.7}$$

Since the relay node must make the primary communication mission top priority, we adopt a conservative policy for prioritizing transmissions. Therefore, upper-bounds are used to ensure that the primary-user QoS constraint is met even in worst-cast channel conditions. The Chernoff upper-bound pairwise error probability (PEP) with respect to an independent Rayleigh distributed channel is expressed by [50]

$$P(\mathbf{d} \to \mathbf{e})_{\hat{\mathbf{H}}(t)} \leq \left(\prod_{i=1}^{L_\Delta} \lambda_i \right)^{-L_r} \left(\frac{\sigma_Q^2}{\frac{4}{N}\left(\frac{N}{\beta} + \frac{L_t}{\alpha}\right)\sigma^2} \right)^{-L_\Delta L_r}, \tag{5.8}$$

where λ_i are the eigenvalues of the code error matrix defined as $C_{p,q} = \mathbf{x}_q^H \mathbf{x}_p$ and $\mathbf{x}_p = (d_1^P - e_1^P, \cdots, d_N^P - e_N^P)^T$ is the pairwise difference between the code and the erroneously detected code. In (5.8), $\sigma_Q^2 = 1 + (\sigma^2/\alpha)$ represents the variance of an element in the estimated channel coefficient matrix $\hat{\mathbf{H}}(t)$, and L_Δ is the rank of the ST code with maximum achievable rank L_t. This PEP expression will be used in the next sections to analyze performance of the primary-user signal and the secondary-user signal, when adapting to different channel conditions. Another important result is the pair-wise error probability for the maximum-likelihood receiver, where the channel state information $\mathbf{H}(t)$ is known exactly. We refer the reader to [49] for this result, and the derivation of channel estimate MSE.

Relay-aided cooperative communication is achieved when a signal transmitted by a source node is received by one or more intermediate relay-nodes, who in turn retransmit the source's signal to the destination. Such systems offer performance advantages in terms of spatial diversity and power gain [45]. In the relaying scheme known as Decode-and-Forward (DF), a relay first receives and then decodes the signal from the source node before retransmitting the signal to the destination. It may be shown that the tight upper-

bound PEP for the DF dedicated-relay scheme in sufficiently high SNR can be expressed as

$$P_{sDF} \approx \frac{N_0^2}{b^2} \cdot \frac{1}{P_1 \delta_{s,d}^2} \left(\frac{A^2}{P_1 \delta_{s,r}^2} + \frac{B}{P_2 \delta_{r,d}^2} \right), \tag{5.9}$$

where N_0^2 is the receiver noise power, $\delta_{s,d}^2$, $\delta_{s,r}^2$, and $\delta_{r,d}^2$ are the instantaneous channel gains between the source and destination, the source and relay, and the relay to destination respectively, and A, B, and b are modulation specific constants defined in [54]. For example, if QPSK is employed A, B, and b take on particular values, and if 16QAM is employed A, B, and b take on another set of values. The real-valued terms P_1 and P_2 represent the optimal power allocations used for transmission by the source and the dedicated DF relay, and are given respectively as

$$\begin{aligned} P1 &= \frac{\delta_{s,r} + \sqrt{\delta_{s,r}^2 + 8(A^2/B)\delta_{r,d}^2}}{3\delta_{s,r} + \sqrt{\delta_{s,r}^2 + 8(A^2/B)\delta_{r,d}^2}} P, \\ P2 &= \frac{2\delta_{s,r}}{3\delta_{s,r} + \sqrt{\delta_{s,r}^2 + 8(A^2/B)\delta_{r,d}^2}} P. \end{aligned} \tag{5.10}$$

We will compare the fixed power allocations for the dedicated-relay DF scheme, (5.9) and (5.10), to the dynamic QoS-based allocation rules presented in Section 5.4.

The power optimization problem can be formulated as finding the minimum pilot power factor α that maintains a QoS level with respect to the primary-user, denoted G_p, according to some QoS rule $G(\alpha, \eta)$. That is minimizing α subject to

$$G(\alpha, \eta) \geq G_p, \;\; 0 \leq \alpha \leq 1, 0 \leq \eta \leq 1, \tag{5.11}$$

where the coefficient η represents the influence of previous primary-channel state information when equalizing the current primary-user block, and will be defined in Section 5.4. One may also interpret η as the confidence of previous channel state information in

modeling current channel state. If channel conditions have deviated from the previous channel estimate and there is less confidence in previous channel state information, the receiver will update its channel estimate using the pilot signals embedded in the transmission. The estimation of channel variances and η may be performed by the receiver and communicated to the transmitter through a side control channel; however the discussion of this control channel is outside the scope of this work.

5.4 Analysis

In the following subsections we present the pairwise error probability, MSE for the channel estimate, and power allocations for the best-effort relaying system. Since the power allocations presented are derived with respect to upper-bound system pairwise error probability and MSE, to ensure that these QoS criterion are unconditionally maintained for system users, the following allocations are sub-optimal.

5.4.1 PEP and MSE for the Best-Effort Delivery Policy

In this section we consider the PEP and channel estimate MSE experienced by the primary and secondary users using the system model presented in Section 5.3. Let us consider the case where channel between the relay and the primary-user, depicted as the bold/solid line in Fig. 1, is stationary over at least two consecutive blocks, and the primary-user is able to detect this event. Such conditions may occur between fixed or immutable nodes, and in this motivating example we assume that transmission of additional pilot signal energy will not dramatically alter or improve the receiver's channel estimate.

When the primary-user channel estimate $\hat{\mathbf{H}}_p(t)$ has perfectly estimated the channel $\mathbf{H}_p(t)$, that is

$$\hat{\mathbf{H}}_p(t) = \mathbf{H}_p(t), \tag{5.12}$$

the Chernoff upper-bound PEP for the primary-user becomes

$$P(\mathbf{d} \to \mathbf{e})_{\mathbf{H}_p(t)} \leq \left(\prod_{i=1}^{L_{\Delta_p}} \lambda_{i_p} \right)^{-L_{r_p}} \left(\frac{P_p}{4\sigma_p^2} \right)^{-L_{\Delta_p} L_{r_p}}, \tag{5.13}$$

where L_{Δ_p} is the rank of the primary-user's channel $\mathbf{H}_p(t)$, λ_{i_p} are the eigenvalues of the primary-channel, L_{r_p} are the number of receive antennas used by the primary-user, and P_p is the normalized power used by the primary/relay-node for transmission. The Chernoff upper-bound PEP for the secondary-user can be expressed as

$$P(\mathbf{d} \to \mathbf{e})_{\hat{\mathbf{H}}_s(t)} \leq \left(\prod_{i=1}^{L_{\Delta_s}} \lambda_{i_s} \right)^{-L_{r_s}} \left(\frac{1 + \frac{\sigma_s^2}{\alpha}}{\frac{4\sigma_s^2}{N} \left(\frac{N}{P_p - \alpha} + \frac{L_{t_s}}{\alpha} \right)} \right)^{-L_{\Delta_s} L_{r_s}}, \tag{5.14}$$

where L_{Δ_s} is the rank of the the channel estimate $\hat{\mathbf{H}}_s(t)$ between the the relay and the secondary-user, λ_{i_s} are the eigenvalues of the secondary-channel, and L_{r_s} are the number of receive antennas used by the secondary-user. We note that in this situation the node is fullfilling primary mission obligations while concurrently acting as a relay for the secondary-user. Thus the channel $\mathbf{H}_s$ is analogous to $\delta_{r,d}$ and its MIMO representation $\mathbf{H}_{r,d}$, referring to the dedicated-relay notation used in Section 5.3.

Under the proposed best-effort delivery policy, power for channel estimation purposes is diverted by the primary-user toward relay transmissions for the secondary-user. It should be noted that the secondary-user also requires energy for proper channel estimation, thus there exists a performance tradeoff for the secondary-user as power to the pilot signals is decreased. A plot of the PEP upper-bound for the secondary-user vs SNR,

with respect to the PEP of the primary-user, for various values of pilot power allocations α is depicted in Fig. 2 for the conditions $L_{r_p} = L_{r_s} = 2$ with $N = 2$ time slots and $L_{t_s} = 2$, for unit gain channels $\Delta_p = \Delta_s = 1$ and $\lambda_{i_s} = \lambda_{i_p} = 1$. In this figure, SNR is defined as $SNR = P_p L_{t_p}/\sigma_p^2$, as suggested by (5.25) and (5.28). From this figure we note that as the value of α decreases the PEP for the secondary-user increases; however the secondary-user is always at a disadvantage with respect to the primary-user, when both users experience identical channel conditions.

We now look at the power allocation problem between the pilot-part and relay-part of the proposed scheme, and how this criteria may change with respect to the needs of the primary-user. This model will be used in the coming subsections to derive sub-optimal power allocation rules. By substituting the power constraint (5.5) into the PEP mismatch equation given in [50], the power optimization problem with respect to the secondary-user becomes

$$\min_{\alpha} \ln \left(\frac{(N - L_t)\alpha + P_p L_{t_p}}{(\alpha + \sigma^2)(P_p - \alpha)} \right). \tag{5.15}$$

Since the primary-user may use its prior channel estimate when current channel conditions do not warrant re-estimation of the channel, the MSE for the primary-user does not depend directly on α because pilot signals are ignored in this case. However, for each block transmission there is a chance that channel conditions will significantly change, requiring the primary-user to update its channel estimate using the pilot signal embedded in the transmission. We model this scenario in simple probabilistic terms as a

two-state model, with channel state ν given as

$$\nu = \begin{cases} 0, & \hat{\mathbf{H}}_p(t-1) = \hat{\mathbf{H}}_p(t) = \mathbf{H}_p(t) \\ 1, & \hat{\mathbf{H}}_p(t-1) \neq \mathbf{H}_p(t). \end{cases} \tag{5.16}$$

We define the model parameter η such that

$$\eta = P(\nu = 0) = 1 - P(\nu = 1), \tag{5.17}$$

thus η is the probability that the previous channel state information is sufficient for equalizing the current block, and $1 - \eta$ is the probability that the primary-user must update its channel estimate based on the pilot signals present. In reality the estimates $\hat{\mathbf{H}}_p(t-1)$ and $\hat{\mathbf{H}}_p(t)$ will never perfectly match current channel conditions $\mathbf{H}_p(t)$; therefore, in practice the assertion of estimates being a 'perfect match' can be described in terms of being within some extremely small threshold of error from $\mathbf{H}_p(t)$, and that this error is negligible.

In state $\nu = 0$, the primary-channel is considered stationary. Thus the channel estimate for the current code remains unchanged from the previous channel estimate, and we assume that the previous channel estimate has converged to match the current channel state. The channel estimate MSE for perfectly estimated channel conditions is given as

$$E\left[MSE_p(t|\nu = 0)\right] = \mathrm{tr}\left\{\mathrm{Cov}\left[\tilde{\mathbf{h}}_p(t)\right]\right\} = \sigma_p^2 L_{t_p} L_{r_p}, \tag{5.18}$$

where $\tilde{\mathbf{h}}_p(t)$ is the vectorization of $\hat{\mathbf{H}}_p(t)$ discussed in [50]. The primary-user PEP is given as

$$P(\mathbf{d} \to \mathbf{e}|\nu = 0)_{\mathbf{H}_p(t)} \leq \left(\prod_{i=1}^{L_{\Delta_p}} \lambda_{i_p}\right)^{-L_{r_p}} \left(\frac{P_p}{4\sigma_p^2}\right)^{-L_{\Delta_p} L_{r_p}}. \tag{5.19}$$

When the channel state is $\nu = 1$, channel conditions for the primary-user have changed substantially, requiring the receiver to update its channel estimate $\hat{\mathbf{H}}_p(t)$. In this state the

receiver will experience an MSE and PEP from (5.14) expressed as

$$E\left[MSE_p(t|\nu=1)\right] = \mathrm{tr}\left\{\mathrm{Cov}\left[\tilde{\mathbf{h}}_p(t)\right]\right\} = \frac{\sigma_p^2 L_{t_p} L_{r_p}}{\alpha}, \tag{5.20}$$

$$P(\mathbf{d}\rightarrow\mathbf{e}|\nu=1)_{\hat{\mathbf{H}}_p(t)} \leq \left(\prod_{i=1}^{L_{\Delta_p}} \lambda_{i_p}\right)^{-L_{r_p}} \left(\frac{1+\frac{\sigma_p^2}{\alpha}}{\frac{4\sigma_p^2}{N}\left(\frac{N}{P_p}+\frac{L_{t_p}}{\alpha}\right)}\right)^{-L_{\Delta_p}L_{r_p}}, \tag{5.21}$$

where parameters L_{Δ_p}, λ_{i_p}, L_{r_p}, and σ_p^2 for the primary-user in state $\nu = 1$ are defined similarly to those of the secondary-user in (5.14) and are independent from those of the secondary-user except for the common factors L_{t_p}, α, and P_p. We notice that while energy allocated to the data part of the primary-user's signal remains constant in the best-effort scheme, the PEP expression for a primary-user using pilot signals to re-estimate channel conditions is dependent on α since both the primary and secondary-users must use the energy in these signals for channel estimation when $\nu = 1$. The expected MSE for the primary-user defined by the two-state model becomes

$$\begin{aligned} E\left[MSE_p(t)\right] &= MSE_p(t|\nu=0)P[\nu=0] + MSE_p(t|\nu=1)P[\nu=1] \\ &= \frac{\sigma_p^2 L_{t_p} L_{r_p}}{\alpha}(\eta\alpha + 1 - \eta). \end{aligned} \tag{5.22}$$

The signal power of the pilot component for the current block is determined by the transmitter's selection of α for that block transmission. The MSE and PEP for the secondary-users will be similar to (5.20) and (5.21) respectively; however the parameters σ_p^2, L_{r_p}, $\hat{\mathbf{H}}_p(t)$, λ_{i_p}, λ_{δ_p} become the parameters of the secondary-channel, σ_s^2, L_{r_s}, $\hat{\mathbf{H}}_s(t)$, λ_{i_s}, and λ_{δ_s}.

5.4.2 Sub-optimal Power Allocation with Respect to Channel Estimate MSE

We now consider the power optimization problem with respect to a QoS rule limiting the MSE of the primary-user channel estimate. Since channel estimation performance

for the primary-user (5.22) depends on pilot signal energy, proper selection of α to maintain a minimum QoS for the primary-user is critical. For comparison, we first consider the sub-optimal power allocation considering only the relay signal to the secondary-user, i.e. when the primary-user is omitted from Fig. 1. The sub-optimal power allocation considering only the relay transmission, α^*, can be found by taking the derivative of (5.15) and setting to zero, i.e

$$\alpha^* = \begin{cases} \frac{P_p - \sigma_p^2}{2} & N = L_{t_p} \\ \frac{P_p L_{t_p} - \sqrt{P_p N (P_p L_{t_p} + \sigma_p^2 (L_{t_p} - N))}}{L_{t_p} - N} & N \neq L_{t_p}, \end{cases} \tag{5.23}$$

where α^* is the sub-optimal power allocation considering only the secondary-user, P_p is the power allocated to pilot plus relay transmissions, L_{t_p} is the number of transmit antenna which is the same for all transmissions, and σ^2 is the channel variance experienced by the pilot signal. Thus, the sub-optimal power allocation rule considering only the secondary-user is simply $\alpha_s^* = \alpha^*$. We now derive the sub-optimal power allocation rule considering a minimum level of QoS for the primary-user only, denoted α_p^*. The primary-user maximum MSE threshold condition, derived according to the two-state model, is given as

$$MSE_p = \frac{\sigma_p^2 L_{t_p} L_{r_p}}{\alpha} (\eta\alpha + 1 - \eta) \leq T_p, \tag{5.24}$$

where T_p is the maximum channel estimation error allowed for the primary-user. It is worth noting that for the case $N \neq L_t$, the sub-optimum solution for the pilot-power allocation factor α^* in (5.23) exists if and only if $SNR \geq (N - L_{t_p})$, where $SNR = P_p L_{t_p} / \sigma_p^2$. In Section 5.5 we consider the case where $L_{t_p} = N = 2$, thus for the sake of exposition we will consider the case of $N = L_{t_p}$ in our analysis here. We substitute

the sub-optimal power allocation $\alpha = \frac{P_p - \sigma_p^2}{2}$ from (5.23) into (5.24) producing the SNR constraint

$$SNR_p \geq L_{t_p}^2 L_{r_p} \left(\frac{2 - 2\eta - \sigma_p^2 \eta + \frac{T_p}{L_{t_p} L_{r_p}}}{T_p - L_{t_p} L_{r_p} \eta \sigma_p^2} \right), \tag{5.25}$$

where $0 \leq \alpha \leq 1$, and SNR_p is the signal to noise ratio of the primary user as a function of the MSE threshold, T_p, and the noise variance of the primary-user's channel, σ_p^2. Using (5.24) and (5.25) and solving for α, the sub-optimal allocation for the primary-user according to the maximum channel estimate MSE threshold rule, $\alpha_{p_{MSE}}^*$, considering the SNR_p constraint and $MSE_p \leq T_p$ becomes

$$\alpha_{p_{MSE}}^* = \frac{(\eta - 1)(\gamma + T_p \pm \sqrt{\xi})}{\eta(\gamma - T_p \pm \sqrt{\xi})}, \tag{5.26}$$

with

$$\begin{aligned} \gamma &= L_{t_p} L_{r_p} (P_p \eta - 2\eta + 2), \\ \xi &= 4 L_{t_p} L_{r_p} \eta \Big(L_{t_p} L_{r_p} \left(\frac{P_p \eta}{4} + P_p - P_p \eta + \frac{1}{\eta} + \eta - 2 \right) + T_p \left(\frac{1}{\eta} - \frac{P_p}{2} - 1 \right) \Big) + T_p^2. \end{aligned} \tag{5.27}$$

We observe from (5.26), that the power allocation rule $\alpha_{p_{MSE}}^*$ is valid only when ξ is positive. For the sake of exposition, we will only consider the case when $\xi > 0$ here, and will assume that α will take on the values of either zero or one otherwise. Also, since α must be positive and in the range $0 \leq \alpha \leq 1$, only the positive values of (5.26) will be considered as valid allocations.

While the best-effort policy relay-node will optimize its parameters with respect to its own transmissions before secondary-user requirements are considered, for comparison the optimum pilot-power allocation considering only the secondary-user under the same

maximum channel estimate MSE threshold criterion for the case $N = L_t$ is simply

$$SNR_s \geq L_{t_p} + \frac{2L_{t_p}^2 L_{r_s}}{T_s} \tag{5.28}$$

with QoS threshold T_s, secondary-user signal to noise ratio SNR_s, and sub-optimum value of α_s satisfying (5.28) under the maximum MSE rule, ${\alpha_{sMSE}}^*$, becomes

$${\alpha_{sMSE}}^* = \frac{L_{t_p} L_{r_s} P_p}{T_s + 2L_{t_p} L_{r_s}}, \;\; 0 \leq \alpha \leq 1. \tag{5.29}$$

We note that the MSE experienced by the secondary-user is a function of α but not η, and is expressed as

$$MSE_s = \frac{\sigma_s^2 L_{t_p} L_{r_s}}{\alpha} \leq T_s, \;\; 0 \leq \alpha \leq 1. \tag{5.30}$$

To demonstrate behavior of the best-effort power allocation policy, in Fig. 3 we plot the values of $\alpha_{p_{MSE}}^*$ and $\alpha_{s_{MSE}}^*$ with respect to η, according to (5.26) and (5.29) respectively. Here we choose $L_{t_p} = L_{r_p} = L_{r_s} = 2$, $P_p = P_s = 1$, and the select $T_s = T_p = 7$ for the MSE thresholds. These results demonstrate sub-optimal pilot-power allocations for the primary and secondary users under various primary-channel stability scenarios, η.

For comparison, in Fig. 3 we plot the power settings P_1 and P_2 using the dedicated-relay-node criteria (5.10) discussed in Section 5.3. These are plotted in Fig. 3 against the sub-optimal MSE QoS rules (5.26) and (5.29), using $\delta_{s,r}^2 = .001$ and $\delta_{r,d}^2 = .001$. For a typical dedicated decode-and-forward relaying scheme, the relay would allocate $P_1 = .408$ and $P_2 = .372$. From Fig. 3 we note that the sub-optimal power allocation using the MSE rule tends toward the dedicated-relay power allocation, P_2, as $\eta \to 0$. Additionally, we observe that as the channel becomes more stationary, i.e. as $\eta \to 1$, less

pilot power is required to maintain a fixed channel estimate MSE for the primary-user, and α decreases under (5.26) accordingly.

Heuristically, we expect that frequent channel re-estimation will be required when the primary-channel is undergoing change, thus in this scenario the relay-node will be less inclined to sacrifice power for relay transmissions, and α will increase accordingly. Conversely, we expect that when channel conditions require less frequent channel estimation, the relay will behave altruistically and sacrifice energy for relay transmissions. Thus, in a typical power allocation policy α will be a monotonically decreasing function of η. This behavior is demonstrated in the MSE-based power allocation rule depicted in Fig. 3.

Next, we observe the general behavior of the channel estimate MSE for both the primary and secondary users as the pilot-power allocation α varies. In Fig. 4 a plot of the channel estimate MSE for the primary-user, MSE_p, and the secondary-user, MSE_s, is presented for the range $0 \leq \alpha \leq 1$, for fixed values of η. We expect that as α increases, the energy devoted to pilot signals used for channel estimation will also increase, and the MSE of the channel estimate for the secondary-user will decrease. In general, channel estimate MSE for both users will be a decreasing function of α, which is demonstrated in Fig. 4. Also, when η is exactly 1 the channel estimator is a perfect representation of the channel with probability 1, and the error of the channel estimate remains constant and invariant of the choice of α.

Fig. 5 demonstrates the PEP of the primary and secondary users for fixed values of α, with respect to η. The results presented here are for the values of $L_{t_p} = L_{r_p} = L_{r_s} = 2$ and $P_p = 1$. We note that as $\eta \to 1$ the channel becomes increasingly stationary, and the PEP for the primary-user decreases accordingly under the fixed MSE rule. Conversely,

as it is assumed that the secondary-user is unconditionally required to preform channel estimation based on the pilot energy present, the PEP response for the secondary-user under the fixed MSE rules remains constant and independent of η, when α is fixed.

We now observe the behavior of a relay-node operating under the dynamic MSE-based power allocation rule (5.27). Fig. 6 demonstrates the PEP results of (5.13) and (5.14) using $\alpha = \alpha^*_{p_{MSE}}$, with respect to η, for various values of T_p. For comparison, PEP results for both users are also shown for the fixed power allocation policy $\alpha = 0.5$. The best-effort behavior this system is readily discernible, as the primary-user enjoys a general PEP advantage over the secondary-user. As the relay-node sacrifices energy for pilot signals used for channel estimation purposes, that is as $\eta \to 1$ in response to increasingly stationary channel conditions, the PEP of the primary-user improves as the PEP of the secondary-user degrades, for the values $T_p = 3$ and $T_p = 5$. This trend is demonstrated by the PEP offset between the primary and secondary-user PEP curves for $T_p = 3$ and $T_p = 5$. The behavior of the system when $T_p = 7$ will be discussed shortly.

From Fig. 6 we note that as channel conditions become more deterministic, i.e. $\eta \to 1$, the PEP for the secondary-user degrades, a curious result indeed. This phenomena may be explained by the increasingly worsened channel estimate a secondary-user would obtain as energy devoted to pilot signals is drastically decreased, since both the primary and secondary users use the same pilot signals for channel estimation. We observe that the PEP curve for the secondary-user is highly determined by the value chosen for T_p. As less model estimate MSE is permitted in the primary-user's channel estimate, i.e. as value of the threshold T_p is decreased, the primary-user becomes increasingly more conservative with the amount of energy it diverts from channel estimation devices resulting

in a decreased PEP for the primary-user. Since the secondary-user is required to unconditionally use pilot energy for channel estimation independently of η, it will benefit from additional pilot energy when preforming channel estimation; however the signal strength of its data signal will suffer as the relay diverts less energy toward the relay services, and thus, PEP of the secondary-user degrades.

We observe that for larger values of T_p, as is the case $T_p = 7$, increased model error severely degrades the general performance of both receivers resulting in detrimental effects for both users. In particular, we observe that as $\eta \to 1$ the PEP actually increases for both users when larger values of T_p are used. We conclude that the value for T_p must be carefully chosen with respect to the SNR experienced by both users under the MSE-based QoS rule (5.27), to ensure that sufficient pilot energy is retained. The main drawback of the MSE-based QoS rule is that the secondary-user cannot benefit from additional relay assistance when the primary communication mission becomes less difficult, i.e. as $\eta \to 1$, since the PEP for the primary-user decreases while the PEP for the secondary-user dramatically increases, in this scenario.

In this motivating example we use identical channel SNRs for the primary and secondary users, that is $\sigma_p^2 = \sigma_s^2$, and we have shown that an overall benefit to a secondary-user is obtainable in the form of useful relay bandwidth, without significant degradation of service to the primary-user. Thus, we have demonstrated that relay-diversity is achievable when a node with primary transmission responsibilities also employs cooperative relaying techniques, under the channel estimate MSE QoS rule.

5.4.3 Sub-optimal Power Allocation With Respect to PEP constraint

In the previous subsection, the sub-optimal power allocations for the best-effort relaying problem were presented for a power allocation policy that optimizes with respect to channel estimate MSE. We now consider the power allocation problem for relays that instead optimize with respect to a rule limiting the PEP experienced by the primary-user. From (5.19) and (5.21) the Chernoff upper-bound PEP expression under the two-state model can be expressed as

$$\begin{aligned} P(\mathbf{d} \to \mathbf{e})_{\mathbf{H}_p(t)} &\leq P(\mathbf{d} \to \mathbf{e}|\nu = 0)_{\mathbf{H}_p(t)} P(\nu = 0) + P(\mathbf{d} \to \mathbf{e}|\nu = 1)_{\hat{\mathbf{H}}_p(t)} P(\nu = 1) \\ &= \eta Q_p \left[R_p - (1 - 1/\eta)\, S_p(\alpha) \right], \;\; 0 \leq \alpha \leq 1. \end{aligned} \tag{5.31}$$

where

$$Q_p = \left(\frac{1}{4\sigma_p^2} \right)^{-L_{\Delta_p} L_{r_p}} \left(\prod_{i=1}^{L_{\Delta_p}} \lambda_{i_p} \right)^{-L_{r_p}}, \tag{5.32}$$

$$R_p = P_p^{-L_{\Delta_p} L_{r_p}}, \tag{5.33}$$

$$S_p(\alpha) = \left(\frac{1 + \frac{\sigma_s^2}{\alpha}}{\frac{1}{N} \left(\frac{N}{P_p} + \frac{L_{t_p}}{\alpha} \right)} \right)^{-L_{\Delta_p} L_{r_p}}. \tag{5.34}$$

The previous simplifications allow us to observe the behavior of sub-optimal power allocations with respect to the terms Q, R, and S when manipulating α and η. For a fixed channel and antenna arrangement L_{r_p}, L_{t_p}, λ_{i_p}, and σ_p^2, we note that Q, R, and S become constants with respect to a fixed α. For a fixed η, we observe that the only term dependent on α is S, and all other aspects of the result are fixed for a power constraint P_p. We evaluate this system under the constraint that the relay-node must maintain a minimum QoS with respect to the PEP of its primary transmissions. The QoS constraint (5.11) for

the primary-user minimizing α with respect to the PEP of (5.31) is simply

$$G_p(\alpha,\eta) = \eta Q_p\left[R_p - (1-1/\eta)\, S_p(\alpha)\right], \;\; 0 \leq \alpha \leq 1. \tag{5.35}$$

Similarly, the constraint for the secondary-user from (5.14) has a QoS threshold determined by maximum allowable PEP. This constraint is a function of the relay's choice of α, or exactly

$$G_s(\alpha) = \left(\prod_{i=1}^{L_{\Delta_s}} \lambda_{i_s}\right)^{-L_{r_s}} \left(\frac{1+\frac{\sigma_s^2}{\alpha}}{\frac{4\sigma_s^2}{N}\left(\frac{N}{P_p-\alpha}+\frac{L_{t_s}}{\alpha}\right)}\right)^{-L_{\Delta_s}L_{r_s}}. \tag{5.36}$$

The sub-optimal power allocation for the best-effort relay with respect to a maximum allowable PEP for the primary-user is found by solving (5.35) for $\alpha_{p_{PEP}}$, i.e.,

$$\alpha^*_{p_{PEP}} = \frac{-P_p\left(L_{t_p} - \Psi(G_p,\eta)\sigma_p^2\right)}{N - P_p\Psi(G_p,\eta)} \tag{5.37}$$

where

$$\Psi(G_p,\eta) = \left(\frac{\eta P_p^{-L_{r_p}L_{\Delta_p}} - G_p\left(\frac{N}{4\sigma_p^2}\right)^{L_{r_p}L_{\Delta_p}}}{\eta - 1}\right)^{\frac{1}{L_{r_p}L_{\Delta_p}}}. \tag{5.38}$$

The PEP behavior using the PEP-based sub-optimal power application rule (5.37) with respect to η is shown in Fig. 7 for the threshold value $G_p = 1.4e^{-5}$.

We have now proposed two QoS criteria for allocating power under the best-effort relay model, given as $\alpha^*_{p_{PEP}}$ and $\alpha^*_{p_{MSE}}$ derived in (5.27) and (5.37), respectively. If we compare the PEP behavior demonstrated in Fig. 7 with the MSE-based power allocation rule depicted in Fig. 3, we see that for (5.27) the acceptable power allocation range $0 \leq \alpha \leq 1$ is valid over a much wider range of η for the MSE-based rule, when compared to the PEP-based rule (5.37), for the value of G_p selected.

The PEP behavior of a relay-node operating under the power constraint (5.37) is shown in Fig. 8 according to (5.35) and (5.36), with respect to the channel estimate

confidence coefficient η, for various values of G_p. As was demonstrated with the MSE-based power allocation rule, cooperative diversity gains are also obtainable using the PEP rule, as the relay diverts energy to secondary-user transmissions. The same phenomena of increased secondary-user PEP under high channel stationary ($\eta \to 1$) is apparent for the same reasons mentioned in Section 5.4.2. As the relay diverts too much energy away from pilot signals, the PEP for the secondary-user suffers since the performance of this receiver requires sufficient pilot energy for channel estimation. Conversely, the primary-user enjoys an increasingly stationary channel with a high probability while its PEP remains the same. The salient difference between the MSE-based and PEP-based power allocation rules is demonstrated in their PEP behaviors: We expect that a power allocation rule optimizing with respect to specific maximum primary-user PEP threshold will exhibit a constant PEP response over all channel stationary states. This behavior is clearly discernible from Fig. 8, as PEP for the primary-user is constant-valued for the entire range $0 \leq \eta \leq 1$.

The best-effort behavior of the system under this rule is also apparent, as the primary-user consistently enjoys an PEP advantage over the secondary-user. The behavior of the PEP threshold G_p is shown in Fig. 8 and may be compared to that of the MSE-based rule in Fig. 6. As G_p increases, the relay will too readily divert energy from pilot signals and the PEP of the primary and secondary-users suffer accordingly. In general, decreasing G_p has the effect of improving PEP for both the primary and secondary-users, at the cost of decreasing useful capacity for the best-effort channel when channel stability confidence is reduced (i.e. the PEP curve for the secondary-user is shifted to the right). We conclude that the threshold G_p, like T_p, must also be carefully chosen with respect to the SNR

experienced by the primary-user under the PEP-based QoS rule (5.37). In this example identical channel SNRs were used for both channels, demonstrating an overall benefit for the secondary-user using best-effort cooperative relaying.

5.5 Simulation Results

In this section we present simulation results for the proposed system for various values of α, using the minimum mean-squared channel estimator MMSE [5] and 2x2 Alamouti ST codes with $M = 3$ and $N = L_{t_p} = 2$. A block-stationary channel model was used with QPSK constellations and an ML symbol decoder. In Fig. 9, we plot the BER experienced by the primary and secondary-users obtained through Monte Carlo MATLAB simulations using values of α equal to 0.5, 0.6, 0.75, and 0.9. In The BER in Fig. 9 is plotted vs. SNR, where SNR in this figure is defined the same as in Fig. 2. We observe that as power is retained for pilot signals the BER for the primary-user improves; however this improvement is achieved at the expense of decreased SNR for the secondary-user's data signal. We also observe that the BER for the secondary-user when $\alpha = 0.6$ is better than the BER experienced at $\alpha = 0.5$ and $\alpha = 0.75$, suggesting that secondary-user BER is a convex function of α with a BER maximum somewhere between these two values. This behavior is due to the trade-off between data signal energy and pilot signal energy afforded to the secondary-user. A plot of the channel estimate MSE for both users is given in Fig. 10 vs. SNR, where SNR in this figure is defined the same as in Fig. 2. We observe that as the value of α increases more energy is allocated for channel estimation, thus the MSE of the receiver decreases accordingly.

We would like to note that the previous sections presented the behavior of the primary and secondary-nodes as a function of η operating with a best-effort policy using two different power allocation rules. While these results are sub-optimal with respect to the primary-user, they fail to achieve optimal allocation with respect to the cooperative system. We now compare the results of the best-effort policy with a power allocation policy that attempts to maximize overall system capacity. The maximum-capacity QoS rule is defined by a policy that allocates power with respect to the primary and relay transmissions in a way that minimizes the sum of the BERs for both links.

The overall system BER is

$$BER_{tot} = r_p P_p(\mathbf{d} \to \mathbf{e}) + r_s P_s(\mathbf{d} \to \mathbf{e}), \tag{5.39}$$

where r_p and r_s are the bits/code/Hz for the primary and secondary transmissions respectively, and $P_p(\mathbf{d} \to \mathbf{e})$, $P_s(\mathbf{d} \to \mathbf{e})$ are the PEP expressions for the primary and secondary-users respectively. A plot of the PEP for the overall system for various values of α is given for the case $r_s = r_p = .5$ in Fig. 11. The results shown are for channel estimate confidence coefficients $\eta = .25$, $\eta = .5$, $\eta = .75$, and $\eta = .95$, with $L_{r_s} = L_{r_p} = L_t = 2$, with primary signal to noise ratio $SNR_p = 13$ dB, the left hand side of (5.25), and secondary signal to noise ratio $SNR_s = 20$ dB, the left hand side of (5.28). We note that Fig. 11 clearly demonstrates that an optimal value for α exists that minimizes the total BER for the cooperative system, and this value changes with respect to the channel stability η. Unfortunately, an analytical solution for this value is difficult to derive due to the large number of variables in (5.39). We note, however, that tractable analytical solutions for α exist for specific antenna configurations L_{r_s}, L_{r_p} and, L_t and fixed channel conditions

$\lambda_{i_p}, \lambda_{i_s}, \Delta_{L_p}$, and Δ_{L_p}, and the exploration of solutions to (5.39) under various scenarios remains future work.

In this work we always assume that the orthogonal resource for the secondary user data is allocated in the primary user frame. One may question if the assumption of the availability of a pre-allocated resource for the secondary is applicable to conventional CR systems.

Cooperative and Cognitive Radio systems can be complex systems faced with a number of resource allocation problems. Nodes that choose to cooperate must constantly ration resources such as transmission energy and bandwidth when these resources are limited. Choosing when and when not to cooperate is, in itself, a difficult problem. While the dynamic allocation of bandwidth is also an extremely important resource allocation problem, in this work we focus solely on the four node power allocation subproblem suggested by Figure 1 and leave the extension of this subproblem to multi-node systems, where bandwidth may also be rationed, as future work.

In other words, we first focus on the power allocation subproblem for the scenario where a node has already chosen to cooperate and has allocated resources for this purpose, so that future work may use this result to solve larger, more complex problems.

We would like to briefly discuss why the values derived in this section are suboptimal and not optimal. In best effort delivery policy the relay must, at all costs, prioritize the primary-user transmissions over any secondary-user cooperative transmissions. Because the relay must make the primary communication mission top priority, we adopt a conservative policy for prioritizing transmissions. Therefore, upper-bounds are used to ensure that the primary-user QoS constraint is met even in worst-case channel conditions.

Because of this cautious approach the power allocation rules given will be sub-optimal under specific channel conditions, however the derivation of closed-form power allocation rules using a less conservative approach could be presented in future work.

One may wonder why the results in Fig.s 9 and 10 do not present series for the optimum values of α. Figures 9 and 10 demonstrate the effect of manipulating α using the block-based structure presented in the System model, and their purpose is to is to depict BER performance for both users for particular values of α, over a range of channel SNRs. In other words, at this point the more sophisticated model parameters, such as η, are omitted for clarity. Since both of the optimal allocation rules are a function of η, including these results (i.e. for some arbitrary value of η) would be confusing and irrelevant because the other series in the plot are not functions of η. The behaviors of the optimal allocation rules, as functions of η alone, are instead presented in separate plots showing analytical results.

In this chapter the relay node sacrifices some of its own resources to help out the secondary node, however this chapter does not explain the node's motivation for cooperating. In this chapter we focus on the power allocation subproblem for the scenario where a node has already chosen to cooperate and has allocated resources for this purpose, so that future work may use this result to solve larger, more complex problems. Therefore the results presented tackle the power allocation problem by developing optimal transmission strategies, once the relay has determined that it in his interest to cooperate. These results can be leveraged when considering more complex problems that take things like motivation into consideration. The important cooperative system reciprocity issues of altruism and avariciousness are also not considered in this work, therefore the motivations of the

relay for aiding the secondary user, while interesting, are irrelevant in this context as this work focuses on optimal cooperation once the decision to cooperate has been made.

The requirement that the primary source has to decode the secondary source's data before forwarding might be unreasonable for energy constrained mobile terminals. This is an important consideration when implementing the Decode-and-Forward relay strategy in a mobile system. Closed-form solutions considering the alternative Amplify-and-Forward strategy would be useful for comparison but will be reserved for future work.

Since the primary source is assumed to know the presence of secondary source and user, from overall power and bandwidth efficiency points of view, one may question if it makes more sense to treat the system in Fig. 1 as a two-source and two-user cooperative wireless network, where each source relays the information of the other source by apportioning its own resources as and when necessary with the hope that the favor will be returned back by the other source.

Our results focus on solving the power allocation subproblem with respect to a single relay, node once the decision to cooperate has been made. This smaller subproblem can be applied to larger, more complex problems in future work. For example, a new problem could be formulated for the 2 node cooperative system described by applying the results of the 4-node subproblem twice, first from the perspective of user 1 and second from the perspective of user 2. This new problem could be formulated in many ways, such as a phase-based system where user 1 relays in during odd-numbered phases and user 2 relays during even-numbered phases, etc. The problem could also be formulated using concurrent transmissions on orthogonal frequencies. The power allocation subproblem we consider here can be applied in both cases, yielding interesting results for these

more complex problems.

If the primary source is far away from the secondary source then the successful decoding probability in the relaying phase is low. On the other hand, if the primary source is far away from the secondary user then gains of relaying (due to path-loss arguments) are diminished. One may question if this limits the usefulness of the proposed method. This is a very good point, and is considered in depth in [45] for the classic 3-node fixed relay case. In this work, transmissions are broken into 2 phases: In Phase 1, the source transmits and this message is received by both the relay node and the destination, and in Phase 2 the relay transmits to the destination, and both messages are considered by the destination. In this work, we consider only the transmission of Phase 2 and focus on optimal power allocations for the new 4-node 'best-effort' problem. These results could be applied to solving the more complex 2 Phase problem in future work, however, the current work focuses on optimal power allocation using best-effort delivery policy once the relay has decided that relaying the signal is in its best interest. It does guarantee that the relayed signal will be useful to the secondary-user, but instead it participates to the extent it believes is prudent. Therefore all results in our chater focus on the performance of the Phase 2 with respect to the best-effort delivery policy, leaving consideration of the 2 Phase problem as future work.

5.6 Conclusion

We have demonstrated that cooperative diversity can be achieved through a best-effort delivery policy. In one example best-effort relaying scenario, energy scavenged

from pilot signals was re-purposed for relay transmissions when channel conditions accommodate. It was demonstrated that in certain circumstances a node may sacrifice resources for relaying signals while maintaining a level of QoS for the primary-user, allowing the node to cooperate at its own discretion. Allowing nodes with primary communication missions to cooperate on a best-effort basis may lead to increased performance in cooperative communication systems, when compared to systems in which only dedicated-relays are allowed.

In deriving a sub-optimal power allocation policy, the MSE and PEP QoS rules were considered. It was demonstrated that the MSE QoS rule may yield lower PEP for the primary-user than the PEP-based QoS rule, for certain ranges of η, and that the MSE QoS rule yields acceptable relay service over a larger range of channel stationary conditions, when compared to the PEP-based rule. The drawback of the MSE-based QoS rule is that the secondary-user does not receive extra assistance from the relay when the primary communication mission becomes less difficult, i.e. as channel conditions become more stationary for the primary-user, since the PEP for the primary-user increases and the PEP for the secondary-user decreases using this rule. Conversely, the PEP-based QoS rule provides a constant PEP for the primary-user over all values of η while providing extra assistance to the secondary-user via decreased PEP; however it yields acceptable relay performance over a small range of η only.

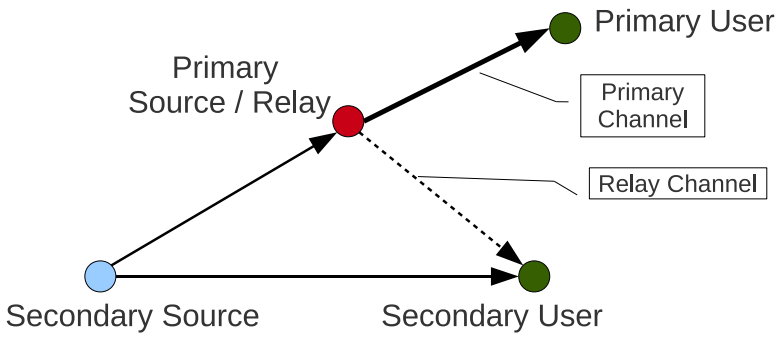


Figure 1: System diagram of a best-effort relay system

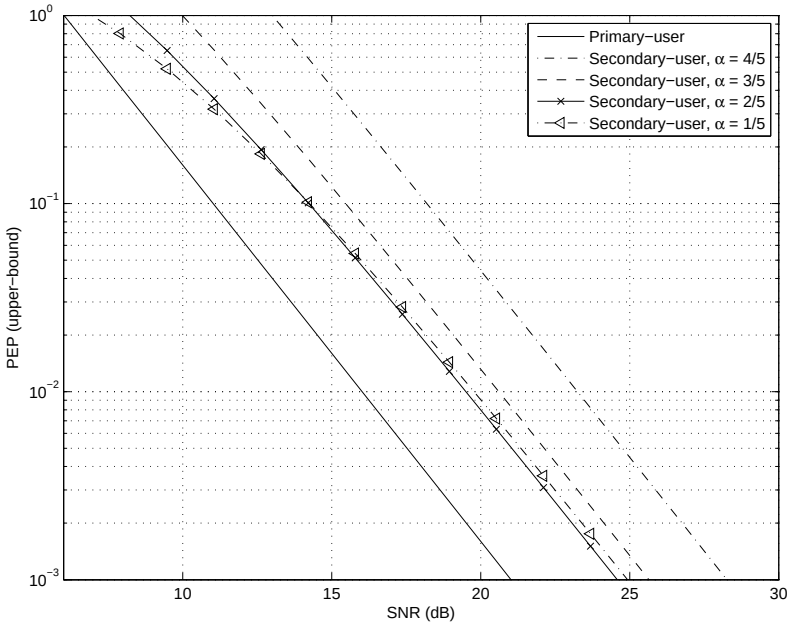


Figure 2: Theoretical upper-bound of PEP of primary and secondary-user for values of α

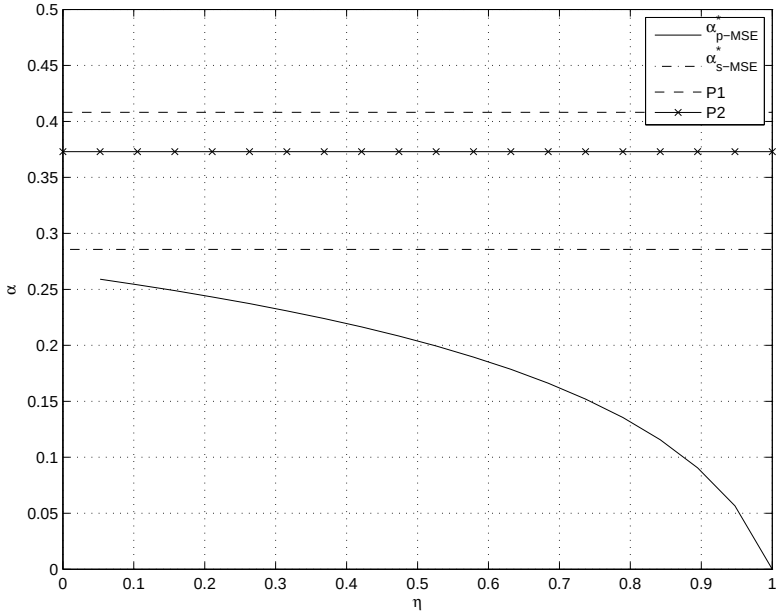


Figure 3: $\alpha^*_{p_{MSE}}$, $\alpha^*_{s_{MSE}}$, $P1$ and $P2$ vs. η

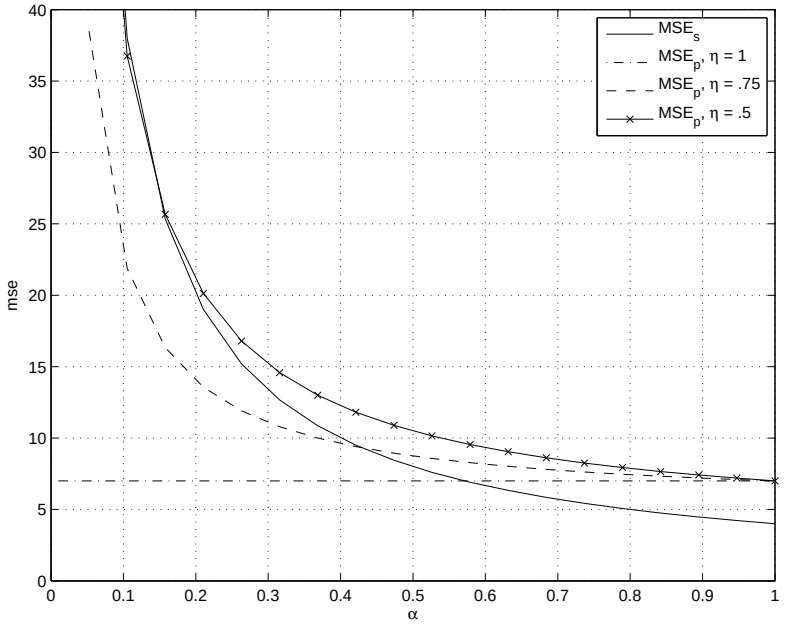


Figure 4: MSE_s and MSE_p vs. α for various η

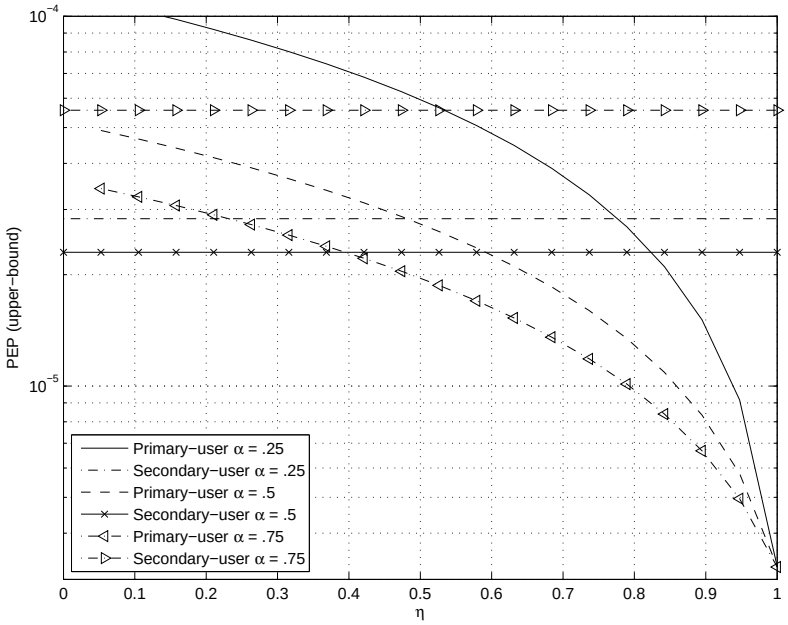


Figure 5: PEP vs. η for fixed α - MSE rule

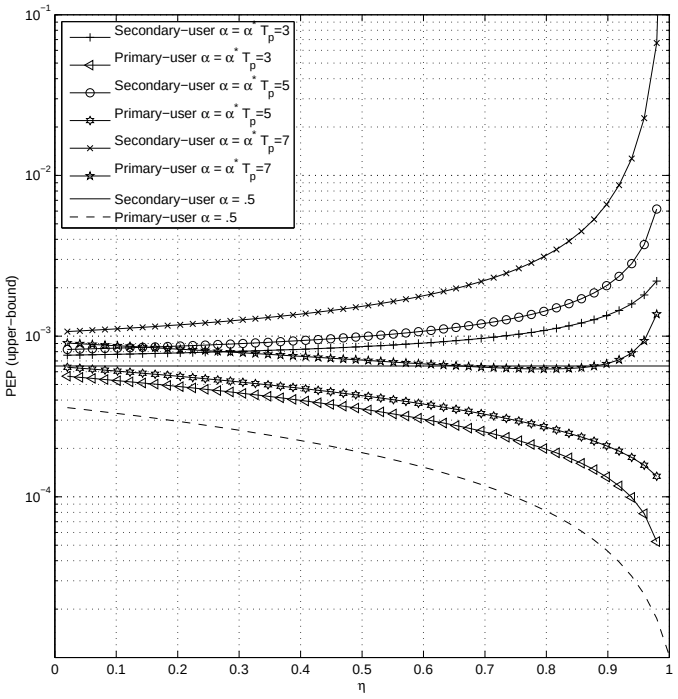


Figure 6: PEP vs. η for values of T_p - MSE rule

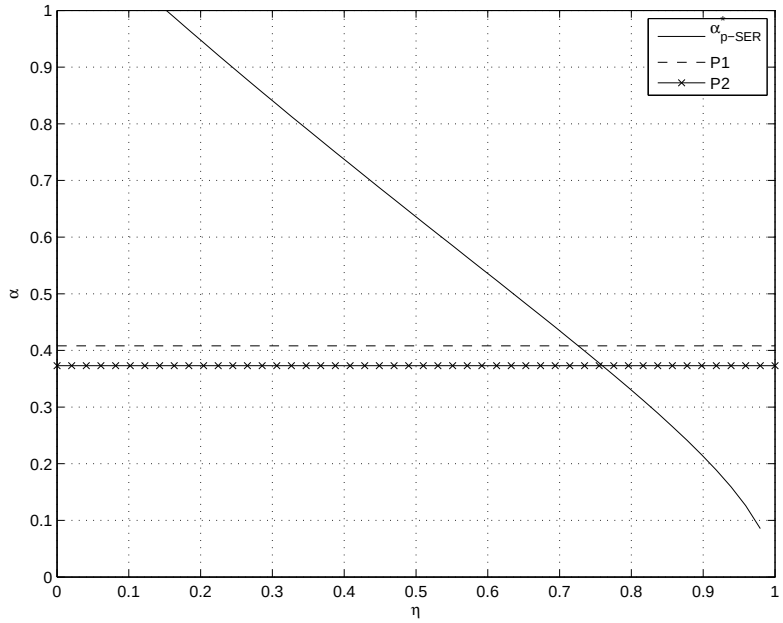


Figure 7: α vs. η for $G_p = 1.4e-5$ - PEP rule

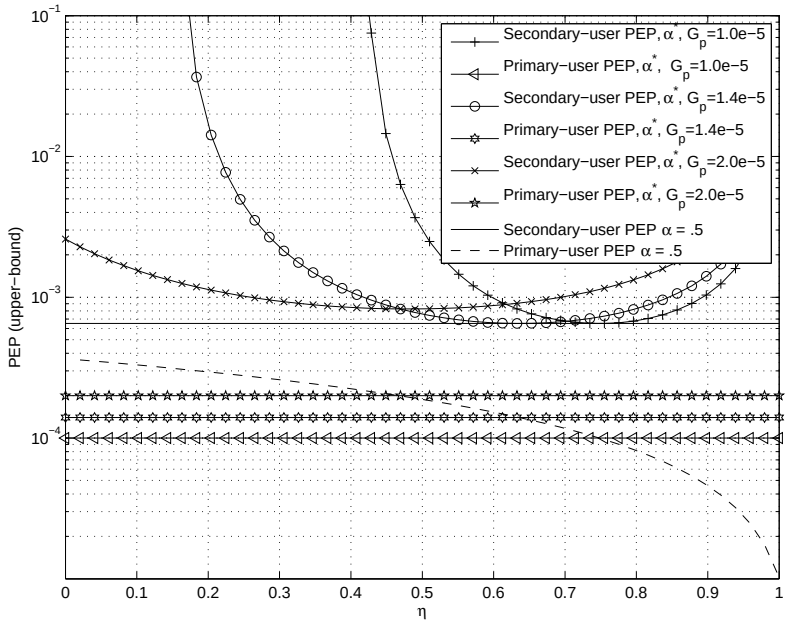


Figure 8: PEP vs. η for values of G_p - PEP rule

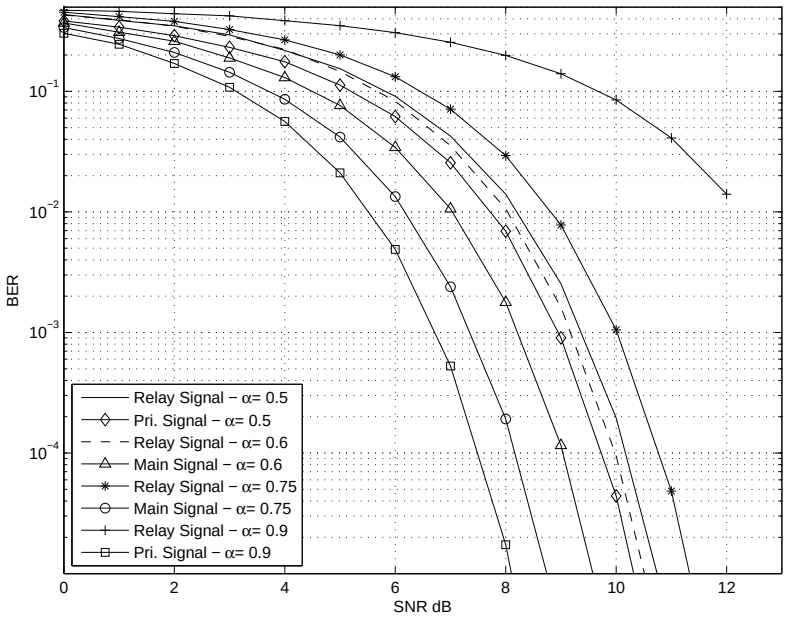


Figure 9: BER for primary and secondary-users for various α

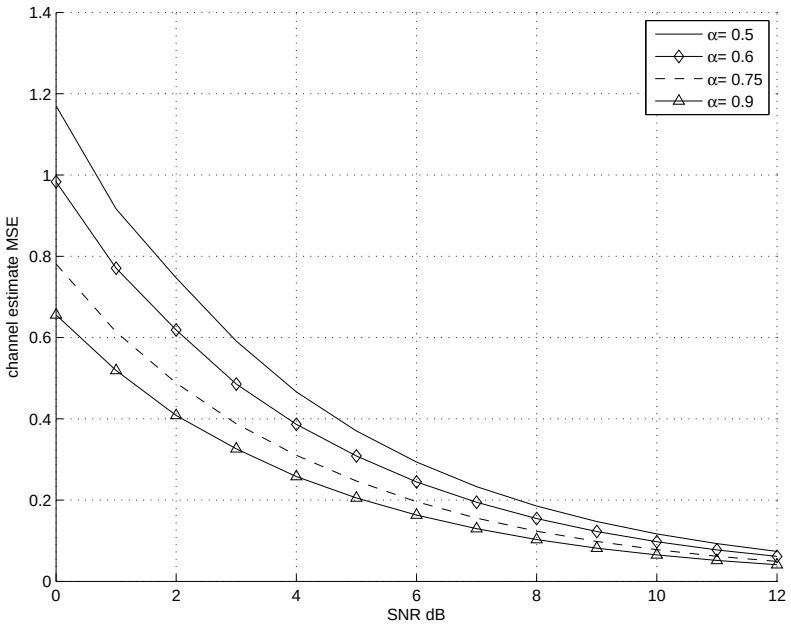


Figure 10: Channel estimate MSE for various α

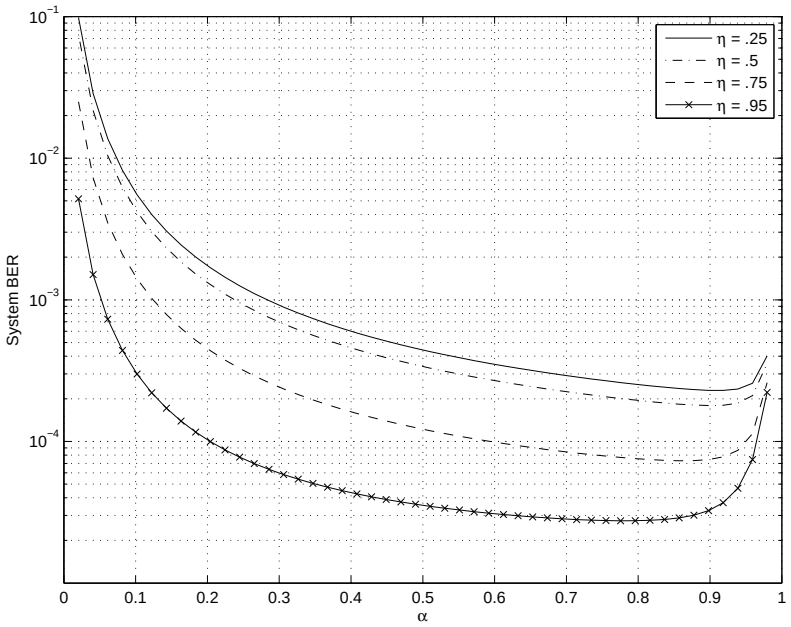


Figure 11: Total System BER vs. α_p for various η

Chapter 6

Concluding Remarks and Future Work

6.1 Concluding Remarks

As wireless communication systems become a ubiquitous part of everyday life, secure authentication systems become increasingly important. In this dissertation we have discussed a wireless fingerprint authentication scheme which can be used prevent a number of identity-based attacks that next-generation wireless systems face. Our wireless fingerprinting scheme has a number of advantages over authentication schemes that embed authentication messages in upper-layer protocols. We have discussed how our method can decrease wasteful processing of unauthenticated and malicious transmissions, by preventing unauthenticated messages from traversing the entire protocol stack before the messages's credentials are verified. We have discussed how potentially malicious transmissions, designed to exploit vulnerabilities in upper-layer protocols, can be avoided using our explicit PHY-layer signal. We have also shown how authentication messages embedded using our fingerprint embedding scheme can achieve detection rates much greater than messages embedded in the transmission's payload, enabling signal authentication in scenarios where original signal itself is unrecoverable due to low signal-to-noise ratio or fading conditions.

To address aforementioned security and robustness problems, our fingerprint embedding scheme provides signal authentication using only PHY-layer characteristics, al-

lowing the signal's source to be verified before the signal is demodulated or decoded. We have demonstrated that our PHY-layer approach creates a completely independent authentication mechanism that is decoupled from upper-layer system design, allowing the designers of wireless systems to create authentication devices that are resistant to upper-level protocol mutability.

While delineating the advantages of PHY-layer authentication approaches, we have described how our fingerprinting method improves on previous works that use blind superposition techniques. We have shown that our method overcomes a number of the disadvantages associated with basic superposition, namely the noise-like appearance of superimposed signals which causes the fingerprint signal to interfere with the original transmission and decrease its reception quality. We have shown how our fingerprint embedding approach can exploit typical receiver preprocessing algorithms, and side information on anticipated channel distortions, to mitigate the undesirable effects associated with blind superposition approaches. By exploiting preprocessing algorithms, our method yields a substantial improvement over prior works.

In addition to improving existing fingerprint superposition methods, we have described how the fingerprinting work presented here can augment and/or improve upon previous fingerprinting schemes that leverage only intrinsic channel information. Our method can be used when channel conditions are not conducive to intrinsic fingerprint recognition, and it addresses many of the drawbacks of intrinsic authentication schemes such as highly correlated multipath profiles and rapidly varying channel conditions. We have described how extrinsic, synthetically generated signals are perceived by receivers, and shown that these extrinsic signals are handled by the receiver in an nearly identi-

cal fashion to natural time-varying channel distortions. Lastly, we have defined a class of *channel-like* fingerprinting methods to describe fingerprint signals designed to manipulate parameters of the transmitted signal using methods that may be modeled as time-varying channel distortions.

By designing our fingerprint signal using time-varying channel models, we have demonstrated how fingerprint distortions are subsequently corrected by the receiver through traditional channel equalization and synchronization practices. This causes the natural intrinsic channel distortion and the synthetic, extrinsic signal, to be partially filtered from the original signal in pragmatic receiver designs, allowing our fingerprint scheme to be applied in heterogeneous systems where legacy and/or unmodified receivers may be present. We have demonstrated how the intrinsic channel distortions become interference when recovering the extrinsic fingerprint signal, and we have shown how the authentication signal must be carefully designed to overcome these distortions.

Leveraging the perceptual model of the wireless receiver has led us to a number of improved fingerprint designs, however additional improvements are obtainable when extra information is available to the transmitter. Open-Loop designs that exploit the receiver perceptual model alone must be created and embedded blindly into the original signal, since side information on channel state is unavailable to the transmitter. Nevertheless, through careful consideration of anticipated channel distortions, the detection performance of our fingerprint designs can be improved further in the Closed-Loop system. We have defined the Open-Loop and Closed-Loop fingerprint design classes to denote the two design approaches.

To demonstrate how Closed-Loop and Open-Loop fingerprints can be created to

suit many transmission schemes and modulations, we have discussed fingerprinting scenarios for a number of modern and traditional wireless transmission designs. We have considered SISO systems using rudimentary pulse-amplitude modulation signaling and blind channel estimation; we have extended this work to MIMO systems, where manipulation of both the time and spacial domains is explored; and we have discussed wideband OFDM fingerprinting scenarios, where the fingerprint can be applied as an overlay in both time and frequency. In all of these transmission schemes an authentication signal BER that is less than the original signal was achieved, with nearly zero perceivable degradation. By leveraging proven, best security practices and cryptographic primitives in the design of the digital authentication message conveyed by the fingerprint, we have shown how to protect the fingerprint from forgery and replay by malicious nodes.

While a number of improvements over prior works have been discussed, our fingerprinting approach has demonstrated another key contribution, above and beyond previous work. We have shown that when the sequence of channel estimates produced by the receiver using embedded pilot signals contains redundant information, our channel-like fingerprinting scheme can exploit these correlations to transmit new information to the receiver. While the impetus for this dissertation has been wireless signal authentication, our fingerprint signal can be leveraged to transmit messages of arbitrary content, to suit a number of purposes. Our work presented in Chapter 5 demonstrates that, via similar assumptions to the authentication work, a relay system using a best-effort delivery policy can be created. Throughout this work we have demonstrated an ulterior connection between the fingerprinting method discussed in Chapters 1 through 4, and best-effort delivery method discussed in Chapter 5. Specifically, we have shown that when the amount

of new information in a sequence of channel estimates decreases due to stationary channel conditions, the information contained in the pilot signals from which the channel estimates are derived decreases accordingly.

By demonstrating that correlations between consecutive channel estimates can be used to convey new information to the receiver, we have shown how bandwidth resources devoted to pilot signals, historically treated as useless transmission overhead, can serve a useful purpose. An ancillary contribution of this work is that when little new information can be gleaned from pilot signals, supplemental information can be added via the same manipulations used in channel-like fingerprinting. We have also discussed how the works of Chapters 1 through 4 relate to the work in Chapter 5, by discussing how intrinsic time-varying channel distortions can interfere with the fingerprint signal, making the fingerprint message a 'best-effort' style delivery as well. We have shown that our method endows pilot signals with the capability of transmitting useful data, while also allowing these signals to fulfill their original purpose. This capability demonstrates an increase in system efficiency and represents an ancillary contribution of this work.

6.2 Future Work

We will now discuss a number of areas for future work for the fingerprinting schemes presented in this dissertation.

In the introductory chapter, we presented our second design goal, namely, the design of optimal, "near-capacity" fingerprints. Since this fingerprint "capacity" will be determined by an almost unlimited set of model parameters jointly describing the charac-

teristics of the original transmission, the embedded pilot signals within the transmission that prescribe the quality of channel estimates that a receiver may obtain, and all of the particulars of the physical time-varying channel between the transmitter and receiver, a complete exploration of optimal fingerprints would be and intractable goal for this work. While a number of designs were discussed in the previous chapters, the shear quantity of fingerprinting scenarios that can be concocted with this number of variables is numerous, making the exploration of optimal channel-like fingerprint designs in scenarios not explicitly addressed in this work open for future research.

We also discussed a forth design goal in the introductory chapter, namely the "equi-energy" fingerprint constraint. This constraint was considered in the work of Chapter 2 and Chapter 5, however, exploration of the equi-energy fingerprint design constraint for the modulation schemes discussed in Chapters 3 and 4 remains future work. We note, however, that fingerprint designs that do not adhere to the equi-energy constraint are very useful and pragmatic, therefore exploration of designs that meet the equi-energy constraint may commence in parallel with designs that do not meet this constraint, and therefore, both design paths remain future work.

Another advantage of physical-layer authentication that was mentioned in passing throughout this work, is that authentication at the physical layer can protect upper-layer protocol processing from unauthenticated messages that can potentially expose the receiver to malformed or malicious transmissions. A whole class of attacks designed to exploit vulnerabilities in the upper-layer protocols, and the implementation of these protocols, exists in the literature. In this work, we have focused on the embedding of the fingerprint message, and the physical layer details pertaining to robust recovery of the

embedded signal. Therefore, the exploration of how physical-layer authentication approaches can protect upper-layer protocol processors from malicious users is outside of the scope of this dissertation and remains future work.

We have also claimed that authentication at the PHY-layer can prevent wasteful processing of unintended, uninteresting, or maliciously fabricated transmissions, allowing nodes to more quickly authenticate legitimate users and implicate malicious users. This statement has not been specifically considered in this work, therefore the potential for physical-layer fingerprints to ameliorate these undue receiver burdens remains future work.

Additionally, the claim that physical-layer embedding approaches allow for completely independent authentication mechanisms which are decoupled from upper-layer authentication devices and protocols, has not been discussed in detail. A detailed delineation of this statement is also outside of the scope of our fingerprint embedding discussion, therefore this claim also remains future work.

While a great deal of the content of Chapter 4 has discussed how proven cryptographic primitives and best security practices can be leveraged to secure the authentication message from attack, the security discussion and example messages given in this chapter serve only to illustrate a complete authentication system. Therefore, a detailed presentation all of the security aspects of the authentication message that is conveyed by the fingerprint would be intractable and outside of the scope of this work. While the authors suggest that system designers adhere to the best design practices routinely discussed in secure systems research circles and the literature that these communities publish, and that to the best of the authors knowledge much of the work in these areas can be directly applied

in the design of the independent authentication message transmission scheme discussed in this work, a detailed discussion of security considerations specifically pertaining to our fingerprint embedding scheme may also be in order, and thus, remains future work.

Lastly, the results of Chapter 3 clearly demonstrate that predictive filtering could be used to improve the performance of a number of fingerprint overlay designs, therefore research into the application of predictive filtering to improve fingerprinting performance in Closed-Loop designs is perhaps the most fruitful area for future work.

Bibliography

[1] 3rd Generation Partnership Project. *Technical Specification Group Radio Access Network; Evolved Universal Terrestrial Radio Access (E-UTRA); LTE Physical Layer*, November 2007. 3GPP TS 36.201 V8.1.0-2007.

[2] 3rd Generation Partnership Project; Technical Specification GroupRadio Access Networks; Deployment aspects (Release 7). *3GPP TR 25.943*, v7.0.0 edition.

[3] William Joseph Adams. *Decentralized Trust-Based Access Control for Dynamic CollaborativeEnvironments*. PhD thesis, Virginia Polytechnic Institute and State University, 2006.

[4] I. Akyildiz, W. Lee, M. Vuran, and S. Mohanty. Next generation/dynamic spectrum access/cognitive radio wireless networks: A survey. *Computer Networks: The International Journal of Computer and Telecommunication Networking*, 50:2127–2159, 2006.

[5] S. Alamouti. A simple transmitter diversity scheme for wireless communications. *IEEE Journal of Selected Areas in Communications*, 16:1414–1458, October 1998.

[6] American Television Systems Committee. *ATSC Digital Television Standard Part 2 RF/Transmission System Characteristics*, January 2007. A/53, Part 2:2007.

[7] B. M. Hochwald B. Hassibi. How much training is needed in multiple-antenna wireless links? *IEEE Transactions on Information Theory*, 49:951–963, April 2003.

[8] P. Bello. Characterization of randomly time-variant linear channels. *IEEE Transactions on Communication Systems*, CS-11:360–393, December 1963.

[9] A. Bender, R. Sherwood, D. Monner, N. Goergen, N. Spring, and B. Bhattacharjee. Fighting spam with the neighborhoodwatch dht. In *IEEE INFOCOM*, April 2009.

[10] T. Bhatt, V. Sundaramurthy, J. Zhang, and D. McCain. Initial synchronization for 802.16e downlink. In *Fortieth Asilomar Conference on Signals, Systems and Computers, 2006. ACSSC '06.*, pages 701–706, 2006.

[11] J. Burbank. Security in cognitive radio networks: The required evolution in approaches to wireless network security. In *International Conference on Cognitive Radio Oriented Wireless Networks and Communications (Crowncom'08)*, May 2008.

[12] R. Chen, J. Park, and J. Reed. Defense against primary user emulation attacks in cognitive radio networks. *IEEE Journal on Selected Areas in Communications*, 26(1):25–37, January 2008.

[13] R. Chen and J. M. Park. Ensuring trustworthy spectrum sensing in cognitive radio networks. *Proc. IEEE Workshop on Networking Technologies for Software Defined Radio Networks IEEE SECON*, September 2006.

[14] T. Clancy and N. Goergen. Security in cognitive radio networks: Threats and mitigation. In *International Conference on Cognitive Radio Oriented Wireless Networks and Communications (Crowncom'08)*, May 2008.

[15] T. M. Cover and A. A. El Gamal. Capacity theorems for the relay channel. *IEEE Transactions on Information Theory*, 25(5), September 1979.

[16] T. M. Cover and J. A. Thomas. *Elements of Information Theory*. Wiley-Interscience, 2006.

[17] I. Cox, M. Miller, and A. McKellips. Watermarking as communications with side information. *Proceedings of the IEEE*, 87:1127–1141, July 1999.

[18] J. B. Dennis and E. C. Van Horn. Programming semantics for multiprogrammed computations. *Communications of the ACM archive*, 9(3):143–155, March 1966.

[19] R. Everson and S. Roberts. Inferring the eigenvalues of covariance matrices from limited, noisy data. *IEEE Transactions on Signal Processing*, 48(7), 2000.

[20] FCC. FCC adopts rules for unlicensed use of television white spaces. http://hraunfoss.fcc.gov/edocs_public/attachmatch/DOC-286566A1.pdf.

[21] Federal Information Processing Standards. *Secure Hash Standard*, 180-1 edition, 1995.

[22] A. Fehske, J. Gaeddert, and J. Reed. A new approach to signal classification using signal correlation and neural networks. In *New Frontiers in Dynamic Spectrum Access Networks (DySPAN'05)*, November 2005.

[23] M. Gandetto, M. Guainazzo, and C. S. Regazzoni. Use of time-frequency analysis and neural networks for mode identification in a wireless software-defined radio approach. *EURASIP Journal on Applied Signal Processing*, 2004:17781790, 2004.

[24] N. Goergen, T. C. Clancy, and T. R. Newman. Physical layer authentication watermarks through synthetic channelemulation. In *New Frontiers in Dynamic Spectrum Access Networks (DySPAN'10)*, April 2010.

[25] N. Goergen, W. Lin, and K. J. R. Liu. Channel-like fingerprinting overlays using predicted channel state(under submission). *IEEE International Confrence on Accoustics, Speech, and Signal Processing (ICASSP)*, 2011.

[26] N. Goergen, W. Sabrina Lin, K. J. R. Liu, and T. Charles Clancy. Authenticating MIMO transmissions using channel-like fingerprinting. In *IEEE GLOBECOM 2010*, December.

[27] N. Goergen, W. Sabrina Lin, K. J. R. Liu, and T. Charles Clancy. Extrinsic channel-like fingerprinting for authenticating MIMO systems (under review). *IEEE Transactions on Wireless Communication*.

[28] N. Goergen, W. Sabrina Lin, K. J. R. Liu, and T. Charles Clancy. Channel-like fingerprinting overlays for authenticating OFDM signals using channel side information. In *APSIPA*, December 2010.

[29] N. Goergen, W. Sabrina Lin, K. J. R. Liu, and T. Charles Clancy. Extrinsic channel-like fingerprinting overlays using channel side-information. In *IEEE Trans. on Info. For. and Sec., Special Issue: Using the PHY-Layer for Securing the Next Gen. of Comm. Sys. (Under Submission)*, 2010.

[30] N. Goergen, K. J. R. Liu, and T. Charles Clancy. Best-effort cooperative relaying (to appear). *IEEE Transactions on Wireless Communication*.

[31] N. Goergen, K. J. R. Liu, and T. Charles Clancy. Best-effort cooperative communication without dedicated relays. *IEEE International Conference on Acoustics, Speech, and Signal Processing*, March 2010.

[32] Nate Goergen, W. Sabrina Lin, and K. J. Ray Liu. Active sensing for dynamic spectrum access. *IEEE Transactions on Information Forensics and Security (in preperation)*, 2011.

[33] S. Haykin. Cognitive radio: brain-empowered wireless communications. *IEEE Journal on Selected Areas in Communications*, 23(2):201–220, 2005.

[34] M. Herdin, N. Czink, H. Ozcelik, and E. Bonek. Correlation matrix distance, a meaningful measure for evaluation of non-stationary MIMO channels. *IEEE Vehicular Technology Conference*, 1:136–140, June 2005.

[35] Institute of Electrical and Electronics Engineers. *IEEE Standard for Local and metropolitan area networks*, March 2004. IEEE 802.16e-2004.

[36] A. O. Kaya, L. J. Greenstein, and W. Trappe. Characterizing indoor wireless channels via ray tracing combined with stochastic modeling. *IEEE Transactions on Wireless Communications*, 8(8):4165–4175, August 2009.

[37] J. E. Kleider, S. Gifford, S. Chuprum, and B. Fette. Radio frequency watermarking for OFDM wireless networks. *ICASSP*, 5:397–400, 2004.

[38] S. Koga and K. Sakurai. Decentralization methods of certification authority using the digital signature schemes. *Second Annual PKI Research Workshop*, April 2003.

[39] T. Kohno, A. Broido, and C. Claffy. Remote physical device finger-printing. *IEEE Symposium on Security and Privacy*, 5, May 2005.

[40] K. Konstantinides and K. Yao. Statistical analysis of effective singular values in matrix rank determination. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 36(5):757–763, May 1988.

[41] J. N. Laneman, D. N. C. Tse, and G. W. Wornell. Cooperative diversity in wireless networks: Efficient protocols and outage behavior. *IEEE Transactions on Information Theory*, 50(12):3062–3080, Dec. 2004.

[42] J. N. Laneman and G. W. Wornell. Distributed space-time coded protocols for exploiting cooperative diversity in wireless networks. *IEEE Transactions on Information Theory*, 49(10):2415–2525, Oct. 2003.

[43] E.G. Larsson and B. R. Vojcic. Cooperative transmit diversity based on superposition modulation. *IEEE Journal Communications Letters*, 9(9):778–780, Oct. 2005.

[44] B. Le, T. Rondeau, D. Maldonado, and C. Bostian. Modulation identification using neural networks for cognitive radios. In *SDR Forum Technical Conference (SDR'05)*, November 2005.

[45] K. J. R. Liu, A.K. Sadek, W. Su, and A. Kwasinski. *Cooperative Communications and Networking*. Cambridge Univ. Press, 2009.

[46] Chun-Shien Lu. *Multimedia Security: Steganography and Digital Watermarking Techniques for Protection of Intellectual Property*. IGI Publishing, Hershey, PA, USA, 2004.

[47] M. Morimoto, M. Okanda, and S. Komaki. A hierarchical image transmission system in fading channel. *IEEE Proceedings 4th IEEE International confrence on Universal Personal Communications*, pages 769–772, November 1995.

[48] T. Newman and T. Clancy. Security threats to cognitive radio signal classifiers. In *Virginia Tech Wireless Personal Communications Symposium*, June 2009.

[49] C. Pirak, Z. J. Wang, K. J. R. Liu, and S. Jitapunkul. Adaptive channel estimation using pilot-embedded data-bearing approach for MIMO-OFDM systems. *IEEE Transactions on Signal Processing*, 54(12), December 2006.

[50] C. Pirak, Z. J. Wang, K. J. R. Liu, and S. Jitapunkul. A data-bearing approach for pilot-embedding frameworks in space-time coded MIMO systems. *IEEE Transactions on Signal Processing*, pages 3966–3979, October 2006.

[51] T. M. Schmidl and D. C. Cox. Robust frequency and timing synchronization for OFDM. *IEEE Transactions on Communications*, 45(12):1613–1621, Dec 1997.

[52] A. Sendornaris, E. Erkip, and B. Aazhang. User cooperation diversity, part I: system description. *IEEE Transactions on Communications*, 51:1927–1938, Nov 2003.

[53] R. Shaukat, S. Khan, and A. Ahmed. Threats identification and their solution in inter-basestation dynamic resource sharing IEEE-802.22. In *International Conference on Convergence and Hybrid Information Technology*, pages 609–614, August 2008.

[54] W. Su, A. Sadek, and K. J. R. Liu. Cooperative communication protocols in wireless networks: Performance analysis and optimum power allocation. *Wireless Personal Communications*, 44:181–217, January 2008.

[55] W. Su, Z. Safar, and K. J. R. Liu. Full-rate full-diversity spacefrequency codes with optimum codingadvantage. *IEEE Transactions on Information Theory*, pages 229–250, January 2005.

[56] S. H. Supangkat, T. Eric, and A. S. Pamuji. A public key signature for authentication in telephone. *APCCAS*, 2:495–498, 2002.

[57] P.D. Sutton, K.E. Nolan, and L.E. Doyle. Cyclostationary signatures for rendezvous in OFDM-based dynamic spectrum access networks. *New Frontiers in Dynamic Spectrum Access Networks, 2007. DySPAN 2007. 2nd IEEE International Symposium on*, pages 220–231, April 2007.

[58] I. E. Telatar and David N. C. Tse. Capacity and mutual information of wideband multipath fading channels. In *IEEE Transactions on Information Theory*, volume 46, July 2000.

[59] J. J. van de Beek, O. Edfors, M. Sandell, S. K. Wilson, and P. O. Borjesso. On channel estimation in OFDM systems. In *Proc. 45th IEEE Vehicular Technology Conf. (VTC)*, pages 815–819, Chicago, IL, July 1999.

[60] J. Wallace and M. Jensen. Time-varying MIMO channels: Measurement, analysis, and modeling. *IEEE Transactions on Antennas and Propagation*, 54:3265–3273, November 2006.

[61] X. Wang, Y. Wu, and B. Caron. Transmitter identification using embedded pseudo random sequences. *IEEE Transactions on Broadcasting*, 50:244–252, September 2004.

[62] Z. Wang and G. B. Giannakis. Linearly precoded or coded OFDM against wireless channel fades. In *IEEE Signal Processing Workshop on Signal Processing Advances in Wireless Communications*, March 2001.

[63] L. Wei. Coded modulation with unequal error protection. *IEEE Transactions on Communication*, 41:1439–1449, October 1993.

[64] L. Xiao, L. Greenstein, N. Mandayam, and W. Trappe. Using the physical layer for wireless authentication in time-variant channels. *IEEE Transactions on Wireless Communications*, 7:2571–2579, July 2008.

[65] L. Xiao, L. J. Greenstein, N. B. Mandayam, and W. Trappe. Channel-based detection of sybil attacks in wireless networks. *IEEE Transactions on Information Forensics and Security*, 4(3):492–503, 2009.

[66] P. Yu, J. Baras, and B. Sadler. Physical-layer authentication. *IEEE Transactions on Information Forensics and Security*, 3:38–51, March 2008.

[67] T. Yücek and H. Arslan. Spectrum characterization for opportunistic cognitive radio systems. *Proc. IEEE Military Communication Confrence*, pages 1–6, October 2006.

[68] Tevfik Yücek and Hüseyin Arslan. A survey of spectrum sensing algorithms for cognitive radio applications. *IEEE COMMUNICATIONS SURVEYS & TUTORIALS*, 11(1):116–130, First Quarter 2009.