Physical Layer Authentication for Wireless Communications

by

Pinchang Zhang

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(The School of Systems Information Science)
in Future University Hakodate
March 2020

To my family

ABSTRACT

Physical Layer Authentication for Wireless Communications

by

Pinchang Zhang

Authentication serves as a critical property of secure communication to verify the identity of the entity involved in the communication. With the rapid development of wireless technologies, the flexible and cost-effective authentication is becoming an increasingly urgent demand for future wireless networks. This is because on one hand, the open and broadcast natures of wireless communications make wireless networks more vulnerable to spoofing attacks, where an unauthorized transmitter may impersonate as a legitimate one. On the other hand, with the wide deployment of Internet of things (IoT) and continuous evolvement of wireless technologies toward the fifth generation (5G) and beyond networks, it is foreseeable that future wireless networks will be consisted of a large number of heterogeneous devices, making cryptographic authentication techniques in wireless networks a challenging issue. Recently, physical layer authentication techniques, which exploit intrinsic and unique features of physical layer for authentication, has drawn a considerable attention to enhance and complement conventional cryptography-based authentication solutions. This thesis focuses on the study of physical layer authentication for wireless communications.

We first explore the channel-based authentication solution taking hardware impairments into account and thus propose a new channel-based authentication scheme for massive multiple input-multiple-output (MIMO) systems with non-ideal hardware. In particular, based on signal processing theory, we formulate channel estimation under hardware impairments and determine error covariance matrix to assess the quantity caused by hardware impairments on authentication performance. With the help of hypothesis testing and matrix transformation theories, we are able to derive exact expressions for the probabilities of false alarm and detection under different channel covariance matrix models. Extensive simulations are carried out to validate theoretical results and illustrate the efficiency of the proposed scheme. Impacts of system parameters on performance are revealed as well.

We then propose a novel authentication solution which not only exploits location-specific wireless channels but also utilizes transmitter-specific hardware impairments for authentication, and thus propose an improved channel-based scheme jointly utilizing channel gain and phase noise in heterogeneous MIMO systems. Three properties of the proposed scheme: covertness, robustness, and security, are analyzed in detail. By using a maximum-likelihood estimator (MLE) and extended Kalman filter (EKF), we estimate channel gains and phase noise, and formulate variances of estimation errors. We also quantize the temporal variations of channel gains and phase noise through the developed quantizers. Based on theories of hypothesis testing and stochastic process, we then derive the closed-form expressions for false alarm and missed detection probabilities with the consideration of quantization errors. Simulations are carried out to validate theoretical results of the two probabilities. Based on theoretical models, we further demonstrate that the proposed scheme makes it possible for us to flexibly control performance by adjusting parameters (such as channel gain threshold, phase noise threshold, and decision threshold) to achieve a required authentication performance in specific MIMO applications.

Finally, we focus on the study of physical layer authentication in a dual-hop wireless network with an untrusted relay and propose an end-to-end (E2E) channel-based authentication scheme. This scheme fully utilizes wireless channel feature (i.e., channel impulse response in the dimensions of amplitude and path delay), and adopts artificial jamming technique, so that it is not only resistant to impersonate attack from an unauthorized transmitter but also resilient to replay attack from the untrusted relay. Theoretical analysis is conducted to derive expressions for false alarm and missed detection probabilities. Finally, numerical and simulation results are provided to illustrate both the efficiency of these theoretical results and the E2E performance of dual-hop wireless networks.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my advisor Professor Xiaohong Jiang. It has been a great honor for me to be one of his Ph.D. students and the Ph.D. experience under his supervision is definitely life-changing for me in Future University Hakodate. I appreciate all his contributions of time and ideas that made my Ph.D. experience productive and exciting. He has been teaching me, both consciously and unconsciously, the important skills required to be a good researcher and the great personality traits that make a better man. I would also like to express my heartfelt gratitude to Professor Jiang's wife, Mrs Li, for her countless care.

Besides my advisor, I would like to thank the rest of my thesis committee: Professor Yuichi Fujino, Professor Hiroshi Inamura, and Professor Masaaki Wada. Their valuable advice and feedback on my dissertation research help me improve this thesis.

I would also like to give my sincere gratitude to Professor Bin Wu of Tianjin University, China, who helped me a lot in improving both the quality and the clarity of dissertation research. He showed me the way to be an excellent researcher.

My sincere thanks also go to other members in our laboratory Xuening Liao, Xiaolan Liu, Xiaochen Li, Wenhao Zhang, Ji He, Ahmed Salem, Shuangrui Zhao, Huihui Wu, Ranran Sun, Yeqiu Xiao, and Chan Gao for their contributions in some way to this thesis.

Last but not the least, I would like to thank my family: my wife, daughter, parents, brother and sister. Words cannot express how grateful I am to them for all of the sacrifices they have made for me.

TABLE OF CONTENTS

LIST OF FIGURES

Figure

LIST OF TABLES

Table

LIST OF APPENDICES

Appendix

CHAPTER I

Introduction

In this chapter, we first introduce the background of physical layer authentication and then we present the objective and main work of this thesis. Finally, we give the outline and main notations of this thesis.

1.1 Physical Layer Authentication

Authentication is a key security service verifying the claimed identity of a legitimate transmitter and rejecting an adversarial impersonation to secure communications [1]. Therefore, providing flexible and cost-effective non-cryptography authentication paradigms is becoming more and more important and challenging for emerging networks (e.g., 5G and IoT networks). This is mainly due to the following two reasons. The first one is that the broadcast nature of wireless medium makes communication systems more vulnerable to various attacks such as impersonation and replay attacks [2]. The other one is that mobile devices randomly join in or leave the network at anytime, resulting in a challenging issue on the distribution and management of secret keys for cryptographic methods for emerging networks [3].

Conventionally, authentication is implemented based on the cryptographic technique [4–6], where it is usually assumed that a secret key is shared in advance between the transmitter and receiver. Nevertheless, the authentication relying on this assump-

tion is increasingly being questioned in emerging network scenarios such as IoT, low power wide area networks [7], and 5G wireless systems [8]. This is mainly due to the reasons that distribution and management of secret keys become troublesome and even impossible in such large-scale heterogeneous networks. Also, the distributed nature of these scenarios makes the stored secret keys vulnerable to physical attacks. E.g., an attacker may capture a legal device and break the keys via hardware level attacks.

Recent works in authentication exploit intrinsic and unique features of physical layer. This draws considerable attentions to both research and academic communities on the development of novel physical layer authentication schemes to complement conventional cryptography-based solutions. Such an authentication approach allows a receiver to quickly differentiate between legitimate and illegitimate transmitters, without having to complete higher-layer processing [9]. Therefore, physical layer authentication is considered as a promising authentication solution for wireless communications, in which terminal devices might not be able to decode each others' higher-layer signaling, because they have different powers and computational capabilities at different levels of the hierarchical architecture [10, 11].

Lots of research efforts have been devoted to the design of effective physical layer authentication schemes, such as channel-based authentication and hardware impairments-based authentication. The fundamental principle of channel-based authentication is that wireless channels are spatially decorrelated between different geographic locations, i.e., characteristics of channels between different transmitter-receiver pairs are significantly different [12–14]. Hardware impairments-based authentication identifies transmitters by using inherent transmitter-specific hardware imperfections (e.g., phase noise, frequency error) [15, 16].

1.2 Objectives and Main Contributions

This thesis exploits intrinsic and unique features of physical layer to authenticate transmitters for wireless communications. Our objective is to design flexible and cost-effective authentication schemes to ensure the security of wireless communications. Towards this end, we first focus on authenticating transmitters in massive MIMO systems with non-ideal hardware, designing a new channel-based authentication scheme with consideration of hardware impairments. We then develop a new authentication scheme, which jointly utilizes two physical layer features (such as wireless channel and hardware features). Finally, we examine the E2E physical layer authentication in a dual-hop wireless network with an untrusted relay and propose an E2E channel-based scheme which utilizes wireless channel feature (i.e., channel impulse response in the dimensions of gain and path delay). Three commonly-used authentication performance metrics are of particular interest, which are false alarm (FA), missed detection (MD), and successful detection (SD) probabilities. Here, FA occurs when a frame transmitted by legitimate transmitter is mistakenly regarded as unauthentic; while MD occurs when a frame originated from illegitimate transmitter is wrongly judged as authentic; and SD occurs when a frame originated from illegitimate transmitter is successfully judged as authentic. The main contributions of this thesis are summarized in the following subsections.

1.2.1 Physical Layer Authentication for Massive MIMO Systems with Hardware Impairments

It is demonstrated that the presence of hardware impairments not only limits capacity but also deteriorates channel estimation accuracy in the high-power regime [17, 18]. Therefore, channel estimation accuracy is affected by hardware impairments, thermal noise, and multiuser interference. It is worth noting that for overall system

performance, considering aggregate effect of all impairments has more substantial benefits than considering separately individual behavior of each hardware module. Recently, increased attention has been focused on a novel system model with aggregate residual hardware impairments which are characterized by independent additive distortion noises at base station and user terminals [17, 19–21].

Hardware impairments need to be deliberately considered in the design of future effective physical layer authentication schemes in massive multiple-input multiple-output (MIMO) systems, which will serve as an essential technology in meeting the continuously increasing throughput demands and spectrum efficiency for the fifth generation (5G) and beyond networks. Based on this background, this work studies transmitter authentication in massive multiple-input multiple-output (MIMO) systems with non-ideal hardware for 5G and beyond networks. The main contributions of this work are summarized as follows:

- By utilizing location-specific property of wireless channels and considering hardware impairments to authenticate transmitters, we first develop a new channel-based authentication scheme for massive MIMO systems with non-ideal hardware.

- To calculate the quantity caused by hardware impairments on authentication performance, we formulate channel estimation under hardware impairments and determine error covariance matrix based on linear minimum mean square error technique.

- Using the quantization result, matrix and hypothesis testing theories, we analytically model FA and SD probabilities under different channel covariance matrix models. Simulation results are also provided to validate theoretical modeling of the two probabilities.

- Through the theoretical models, we further examine how different levels of hard-

ware impairments impact authentication performance, and also determine how to set antennas correlation pattern and the number of base station antennas to achieve a required authentication performance.

1.2.2 Physical Layer Authentication Jointly Utilizing Channel and Phase Noise in MIMO Systems

Extensive research efforts have been devoted to the study on joint estimation of channel and phase noise in MIMO systems [22–24]. The problem of joint estimation of channel and phase noise is considered using data-aided and decision-directed weighted least-squares approaches in MIMO systems [24]. These works mainly focus on joint estimation of channel and phase noise without taking important security issue into account in MIMO systems [22–25]. To the best of the authors' knowledge, how to develop a flexible and cost-effective authentication scheme by jointly utilizing the wireless channel and hardware features, has not been considered. Based on the above background, we explore physical layer authentication by jointly taking wireless channel and hardware features into account for authentication in heterogeneous coexist MIMO systems. The main contributions of this work are summarized as follows:

- By utilizing two physical layer features in terms of location-specific channel gains and transmitter-specific phase noise to authenticate transmitters, we propose a simple and flexible physical layer authentication scheme in MIMO systems to differentiate between legitimate and illegitimate transmitters. We analyze three properties of this scheme: covertness, robustness, and security, which are three important aspects to assess authentication schemes.

- To formulate variances of estimation errors in terms of channel gains and phase noise, we adopt a maximum-likelihood estimator (MLE) to estimate channel gains and soft-input extended Kalman filter (EKF) to track phase noise over a

frame, and then quantize the temporal variations of channel gains and phase noise through the developed quantizers.

- By using quantization results and theories of hypothesis testing and stochastic process, we derive the closed-form expressions for FA and MD probabilities with a careful consideration of quantization errors. Simulation results are also provided to validate theoretical models for the two probabilities.

- through theoretical models, we further investigate how thresholds (for channel gain, phase noise, and decision) can impact the authentication performance. Guidelines for properly setting these parameters are also provided to achieve a desired authentication performance.

1.2.3 End-to-End Physical Layer Authentication for Dual-Hop Wireless Networks

Existing works mainly focus on one-hop physical layer authentication, where transmitters and receivers can communicate with each other directly. In the large-scale distributed wireless networks such as IoT and 5G wireless systems [8], E2E communication is usually conducted with the help of relay(s) [26–28]. Due to transmission efficiency, delay and secrecy constraints, the multi-hop E2E physical layer authentication is an important research issue in wireless communication scenarios, where relay only needs to amplify and forward the signals transmitted by the transmitter to the legitimate receiver, or to decode the signals and then forward them to the legitimate receiver. To the best of our knowledge, the multi-hop E2E physical layer authentication is still not well-explored yet. Notice that the available one-hop physical layer authentication schemes can not be directly extended to multi-hop E2E physical layer authentication mainly due to the following challenges. First, the cascade channels between the transmitter and receiver become much more dynamic and complicated,

making multi-hop E2E physical layer authentication more challenging [29]. Second, the relay can be potential adversary to record the received signals and initiate replay attacks, bringing new threat to the E2E physical layer authentication.

As one step towards the study of E2E multi-hop physical layer authentication, this work focuses on the channel-based E2E physical layer authentication in a dual-hop wireless network with an untrusted relay. This is because the dual-hop wireless networks are simple and serve as a foundation for the study of general multi-hop wireless networks. By carefully exploiting the highly dynamic properties of the dual-hop cascade channels, we develop an efficient E2E physical layer authentication scheme to discriminate transmitters at different locations. The main contributions of this work are summarized as follows.

- We propose a new E2E physical layer authentication scheme for dual-hop wireless networks with an untrusted relay. This scheme utilizes the location-specific features of both channel gain (CA) and delay interval (DI) of cascaded channels to discriminate transmitters, and adopts the artificial jamming technique to resist against possible replay attack from the untrusted relay.

- Using statistical signal estimation theory and the two-dimensional quantizers, we can qualify the temporal variations of CA and DI of cascaded multipath channel.

- Based on the hypothesis test theory, theoretical analysis is then conducted to derive the expressions for FA and MD probabilities, such that E2E authentication performance under the proposed E2E physical layer authentication scheme can be fully depicted.

- Finally, extensive numerical/simulation results are provided to validate theoretical results for FA and MD probabilities and to illustrate performance for the proposed scheme.

Towards this end, we first focus on authenticating transmitters in massive MIMO systems with non-ideal hardware, designing a new channel-based authentication scheme with hardware impairments taken into account. We then develop a flexible and cost-effective authentication scheme, which jointly utilizes two physical layer features (such as wireless channel and hardware features). Finally, we examine the E2E physical layer authentication in a dual-hop wireless network with an untrusted relay and propose a corresponding physical layer authentication scheme which utilizes wireless channel feature (i.e., channel impulse response in the dimensions of gain and path delay).

1.3 Thesis Outline

The remainder of this thesis is outlined as follows. Chapter II introduces the related works of this thesis. In Chapter III, we focus on the study of channel-based authentication under taking hardware impairments into account. Chapter IV presents the work on an improved channel-based authentication scheme jointly utilizes two physical layer features and Chapter V introduces the work regarding the E2E channel-based authentication scheme in a dual-hop wireless network with an untrusted relay. Finally, we conclude this thesis in Chapter VI.

1.4 Notations

The main notations of this thesis are summarized in Table 1.1.

Table 1.1: Main notations

Symbol	Definition
$A/B/E/R$	Alice/Bob/Eve/Relay
X	an unknown transmitter

$P_F/P_M/P_D$	false alarm/missed detection/successful detection probability
δ_h/δ_θ	channel gain/phase noise threshold
SNR	signal-to-noise ratio
SINR	signal to interference plus noise ratio
$h_{i,j}$	channel coefficient between entity i and j
$\mathbb{E}[\cdot]$	expectation operator
$\mathbf{Pr}(\cdot)$	probability operator
$(\cdot)^*$	conjugate operator
$(\cdot)^T$	transpose operator
$(\cdot)^H$	conjugate transpose operator
$\lvert\cdot\rvert$	absolute value operator
$\mathbb{C}^{M\times K}$	set of complex-valued $M \times K$ matrices
$\mathrm{Cov}(\cdot)$	covariance operator
$\det(\cdot)$	determinant operator
$\triangleq$	definition operator
$\mathrm{tr}(\cdot)$	matrix trace function
$\mathrm{diag}[\lambda_1, ..., \lambda_M]$	diagonal matrix with $\lambda_1, ..., \lambda_M$ on main diagonal
$\exp(\cdot)$	exponential function
$\mathbf{\Gamma}_{\chi_i^2}(\cdot)$	the right-tail probability function for a χ_i^2 random variable with i degrees of freedom
H_0	null hypothesis
H_1	alternative hypothesis

CHAPTER II

Related Works

This chapter introduces the existing works related to our study of the thesis, including wireless channel-based authentication, hardware impairments-based authentication, and tag-based authentication solutions.

2.1 Wireless Channel-based Authentication

The main idea of channel-based authentication is that channel state information is location-specific according to the radio propagation theory [30]. It is difficult for an adversary to precisely build the same channel that is being used by a legitimate transmitter-received pair. The authors in [31] presented a channel-based authentication scheme exploiting the spatial variability of channel frequency response over time-varying channels in a rich scattering environment. The authors in [12] further explored the channel-based authentication by using the temporal channel variations of channel impulse response to authenticate transmitters at different locations in frequency-selective Rayleigh channels. An new physical layer authentication framework was designed in [32] based on the hypothesis testing under a multiple wiretap channels with correlated fading. The authors studied the optimal attack strategy for the cases of both single attempt and multiple repeated trials under some degree of among correlated wireless channels. The authors in [33] examined a single-carrier time

domain method through either residual testing or time-domain wireless channel state information (CSI) comparison. The authors in [34] proposed a physical layer authentication scheme by using the unique CSI of a legitimate transmitter to authenticate subsequent transmissions (frames) from the claimed entity. This scheme relies on comparing two random CSIs to ascertain whether they have identical power spectral densities. Based on the comparison of channel estimates obtained from the received messages, an outer bound on the type I/II error probability region was investigated in [35]. Here, multivariate Gaussian vectors were utilized to model channel estimates when only some side information on the channel estimates is available at the adversary. The authors explored the attacking strategy that presents the tightest bound on the error region. The authors in [14] proposed a novel authentication scheme over time-varying multipath channels by jointly using the location-specific properties of both amplitude and multipath delay of wireless channels to authenticate transmitters. The authors in [13] further proposed a logistic regression-based authentication exploiting channel state information and multiple landmarks to improve the spoofing detection accuracy.

2.2 Hardware Impairment-based Authentication

Hardware impairments-based authentication identifies transmitters by using inherent transmitter-specific hardware imperfections (e.g., phase noise and frequency error, in-phase/quadrature (I/Q), and carrier frequency offset (CFO)). The authors in [36] explored various non-cryptographic mechanisms for device authentication in wireless networks through physical layer features or information. Merits and demerits of these authentication solutions and the practical implementation issues are also discussed. The scheme proposed in [15] leverages minor hardware impairments to identify a frame's device-of-origin by analyzing radio-frequency signals. As an attempt toward a model-based method using statistical models of radio frequency (RF)

device components, the authors in [37] designed an algorithms based on statistical signal processing methods to utilize non-linearities of wireless devices for authentication. They also examined the practical variations of device chain components through simulations, measurements and manufacturers' specifications. The authors in [38] further showed that time domain analysis of a pair of distortion signals caused by imperfections of manufacturing processes can be used to discriminate wireless devices. By using device-specific hardware impairment I/Q imbalance, the authors in [39] proposed a new relay authentication scheme to secure amplify-and forward relay networks. In [39], the generalized likelihood ratio test for classical linear model and a two-parameter hypothesis testing were formulated to improve the authentication performance in differentiating delicate difference between I/Q imbalances. By utilizing oscillators in each transmitter-and-receiver pair, the authors in [40] developed an authentication scheme based on radio frequency time-varying CFO associated with each pair of wireless devices.

Radio-frequency distinct native attribute (RF-DNA) fingerprinting was developed to authenticate transmitters in [41]. By exploiting RF-DNA fingerprints consisting of higher order statistical features, e.g., instantaneous amplitude, phase, and frequency responses, transmitter authentication can be implemented. The authors showed device classification with dimensional reduction analysis (DRA) feature subsets. The multiple discriminant analysis (MDA) classification models are used to assess verification accuracy in [42]. [43] The authors developed a technique acquiring actively and passively wireless devices fingerprint through information emitting by devices. This fingerprint is a function of different device hardware impairments and variations in devices' clock skew. Then, the fingerprint is exploited to identify physical device and device type. By considering diverse hardware impairments (such as circuits, antenna, and environments), the authors in [16] explore a reliability and differentiability of physical layer authentication by means of theoretical modeling and experiment val-

idation. It is notable that the above authentication schemes exploit either intrinsic features of wireless channels or inherent hardware impairments to authenticate transmitters separately.

2.3 Tag-based Authentication

Note that tag-based PHY-layer authentication, which embeds tag signals to modulated signals for identifying devices, is regarded as a promising authentication solution. Such a method has two major advantages over conventional authentication technologies. First, it enables a legitimate receiver to quickly identify transmitters without having to complete higher-layer processing. Second, embedding authentication tag into message signals and simultaneously transmitting them through wireless channels allow adversaries to obtain only the noisy observation of authentication tag [44–48]. Tag-based PHY-layer authentication has been extensively explored in traditional wireless network architectures. The authors in [44] investigated a cryptography secure low-power authentication method that hides tag signals in the modulated signals for authentication. The authors in [45] presented an improved tag-based authentication scheme, where tag conveys much less information of the secret key to adversaries. In [46], the authors implemented extensive experiments in software defined radio system in order to illustrate the authentication performance of the tag-based authentication. The authors in [47] proposed a blind tag-based authentication scheme, which adopts the techniques of blind known interference cancellation and differential processing to conduct authentication. The authors in [48] proposed a slope tag-based PHY-layer authentication scheme which is covert to the unware receiver, robust to interference, and secure for authentication.

It is notable, however, that there are some problems for the above aforementioned authentication solutions.

1) These authentication solutions mainly focus on authenticating in non-massive

MIMO systems. It is demonstrated that the presence of hardware impairments not only limits capacity but also deteriorates channel estimation accuracy in the high-power regime. Therefore, channel estimation accuracy is affected by hardware impairments, thermal noise, and multiuser interference. It is worth noting that for overall system performance, considering aggregate effect of all impairments has more substantial benefits than considering separately individual behavior of each hardware module. It is important to design a new channel-based authentication approach by taking aggregate effort of hardware impairments for MIMO systems into account.

2) How to develop a flexible and cost-effective authentication scheme jointly utilizing the wireless channel and hardware features has not been considered.

3) The above available works mainly focus on one hop physical layer authentication, where transmitters and receivers can communicate with each other directly. In the large-scale distributed wireless networks like 5G wireless systems, ad hoc networks and wireless sensor networks, the E2E communication is usually conducted with the help of relay(s), making the muti-hop E2E authentication an important research issue.

CHAPTER III

Physical Layer Authentication for Massive MIMO Systems with Hardware Impairments

Hardware impairments need to be deliberately considered in the design of future effective physical layer authentication scheme in massive multiple-input multiple-output (MIMO) systems, which will serve as an essential technology in meeting the continuously increasing throughput demands and spectrum efficiency for the fifth generation (5G) and beyond networks. In this chapter, we focus on authenticating transmitters in massive MIMO systems with non-ideal hardware. We propose a new channel-based authentication scheme with hardware impairments being taken into account. In particular, based on signal processing theory, we first formulate channel estimation under hardware impairments and determine its error covariance matrix. With the help of hypothesis testing and matrix transformation theories, we are then able to derive exact expressions for the probabilities of false alarm and detection under different channel covariance matrix models. Finally, extensive simulations are carried out to validate theoretical results and illustrate the efficiency of the proposed scheme. Impacts of system parameters on performance are revealed as well.

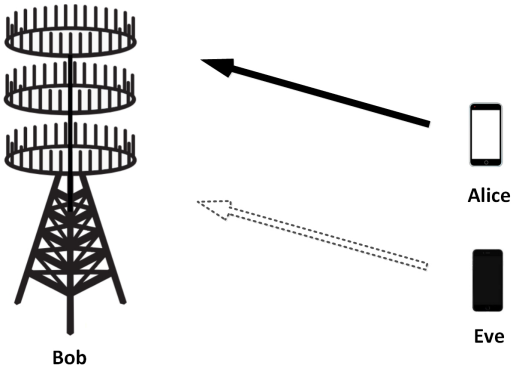


Figure 3.1: System model.

3.1 SYSTEM MODEL

3.1.1 Network Model

As illustrated in Fig. 3.1, we consider an uplink massive MIMO system consisting of three different entities: one M-antenna base station (namely Bob), two single-antenna mobile terminals (namely Alice and Eve). To ensure independent fading channels, any two entities are assumed to be far away from each other, with a distance far more than spatial separation of a wavelength (e.g., 6 cm for a typical 5 GHz RF system). This assumption is reasonable because when the distance between entities is less than one wavelength, they will fail to work well due to strong interference [12, 14]. Alice is a legitimate transmitter to the intended receiver Bob. Eve serves as an adversary who attempts to steal some useful information and/or to inject his own aggressive signals into the network by impersonating Alice. Suppose that Bob receives two messages (also referred to as frames) at time $k-1$ and time k. We assume that the first one is confirmed being from Alice by using a standard higher-layer protocol [12], and Bob stores the channel connecting Alice with him. The other one, received by Bob at time k, is either from Alice or Eve. Therefore, the objective for Bob is to differentiate between Alice and Eve. The message to be authenticated is not expected to be sent continuously but it is necessary to ensure the continuity of

authentication process by probing the channel at time intervals smaller than channel coherence time [31].

3.1.2 Channel Model

We first introduce the following definitions on fading channels:

- **Spatial channel correlation:** A fading channel $\mathbf{h} \in \mathbb{C}^{M\times 1}$ is spatially uncorrelated, if channel gain $\|\mathbf{h}\|^2$ and channel direction $\mathbf{h}/\|\mathbf{h}\|$ following uniform distribution over unit-sphere in $\mathbb{C}^{M\times 1}$ are uncorrelated random variables. Otherwise, it is spatially correlated.

- **Temporal channel correlation:** A fading channel $\mathbf{h} \in \mathbb{C}^{M\times 1}$ is temporally correlated, if each channel component remains constant over one frame and is continuously varying from one frame to the next due to the relative motion between entities and such temporal variations are correlated.

Similar to the work in [14, 31, 47], we consider that channels from the same transmitter-receiver pair are temporally correlated and follow Rayleigh fading channel. The temporally correlated channel may be either spatially independent or spatially correlated.

We use $\mathbf{h}_X(k) = [h_{X,1}(k)\cdots h_{X,M}(k)]^T \in \mathbb{C}^{M\times 1}$ to denote channel vector between X and Bob at time k, and then we have $\mathbf{h}_X(k) \sim \mathcal{CN}(\mathbf{0}, \mathbf{R}_X)$ where $\mathbf{R}_X = \mathbb{E}\{\mathbf{h}_X(k)\mathbf{h}_X^H(k)\} \in \mathbb{C}^{M\times M}$ is a symmetric positive semi-definite matrix. Following existing related literature [12], it is assumed that the statistical information of channel is available at Bob. This assumption is generic and has been adopted in the literature [12, 49].

Here, we exemplify temporal channel variations. We first focus on the time-autocorrelation of channels, which is caused by the Doppler rate. Similar to [31, 50], we assume that the temporal variations of the channel between Alice and Bob are

mutually independent and the normalized maximum Doppler frequencies are identical. Let f denote the normalized maximum Doppler frequency. According to the well-known Jakes' model [30], the time-autocorrelation matrix of $\mathbf{h}_A(k)$ for an arbitrary time lag k_s can be written as $\mathbf{\Psi}_A[k_s] = \mathbb{E}\{\mathbf{h}_A(k)\mathbf{h}_A^*[k+k_s]\} = \mathbf{R}_A J_0(2\pi f k_s)$, where $J_0(\cdot)$ is the zero$^{\text{th}}$ order Bessel function of the first kind. Similar to [47, 50], a first-order Gauss-Markov process is employed to model the fluctuation of channel state. According to [47, 50], correlation coefficient matrix of $\mathbf{h}_A(k)$ can be defined as $\mathbf{\Psi}_A(k_s)\mathbf{R}_A^{-1}$. Thus, we have

$$\mathbf{h}_A(k) = \alpha \mathbf{h}_A(k-1) + \sqrt{1-\alpha^2}\mathbf{e}_A(k), \tag{3.1}$$

where α is temporal correlation coefficient and $\mathbf{e}_A(k) \sim \mathcal{CN}(0, \mathbf{R}_A)$ is independent of $\mathbf{h}_A(k-1)$.

3.1.3 Communication Model with Hardware Impairments

In practical applications, transceivers always suffer from hardware impairments. The impact of hardware impairments on signals mainly includes two aspects: 1) the signal that is actually generated and transmitted does not agree with the intended one; 2) the received signal is distorted during reception processing. Such impairments are treated as the additional distortion noise which are in general relevant to signal power as well as channel gain. Various sources of impairments (e.g., I/Q imbalance and phase noise) may result in distortion noise [17].

In order to characterize non-ideal hardware impairments more accurately, we adopt the communication model with the aggregate residual hardware impairments, which are characterized by independent additive distortion noises at the transmitter and receiver as in [17]. Considering the authentication performance for a system, this is reasonable because considering the aggregate effect of all the residual hard-

ware impairments is more significant than considering residual hardware impairments separately/individually.

Frame-by-frame transmission is considered. A transmission frame consists of deterministic pilot symbols used for channel estimation and stochastic data symbols. Suppose an unknown mobile transmitter X tries to send a frame to be authenticated to Bob at time k. Let $s(k) \in \mathbb{C}$ denote the deterministic pilot signal transmitted by X at time k and let $p = \mathbb{E}\{|s(k)|^2\}$ denote the average power of $s(k)$. Let $\boldsymbol{\nu}(k) \in \mathbb{C}^{M\times 1}$ denote an ergodic process comprised of zero-mean complex additive white Gaussian noise (AWGN) $\boldsymbol{\nu}_N(k) \sim \mathcal{CN}(\mathbf{0}, \sigma_N^2\mathbf{I})$ and interference from other simultaneous transmissions $\boldsymbol{\nu}_I(k) \sim \mathcal{CN}(\mathbf{0}, \sigma_I^2\mathbf{I})$, which is independent of $s(k)$. Then, the signal received by Bob at time k can be written as

$$\mathbf{y}_{BX}(k) = \mathbf{h}_X(k)(s(k) + \eta_X(k)) + \boldsymbol{\eta}_B(k) + \boldsymbol{\nu}(k), \tag{3.2}$$

where $\eta_X(k) \in \mathbb{C}$ and $\boldsymbol{\eta}_B(k) \in \mathbb{C}^{M\times 1}$ denote the independent additional distortion noises at X and Bob at time k, respectively. According to [17, 20], ergodic stochastic processes can model the aggregate residual impairments at X and Bob. Note that distortion noise caused by hardware impairments is irrelevant to $s(k)$, but statistically depend on channel realizations. Also, this distortion noise follows a complex Gaussian distribution for a given channel realization, which is verified experimentally and supported by several theoretical results [17, 20]. Specifically, under a given $\mathbf{h}_X(k)$ the conditional distributions are $\eta_X \sim \mathcal{CN}(0, \varsigma_X)$ and $\boldsymbol{\eta}_B \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{\Upsilon}_B)$, respectively, wherein ς_X can be modeled as $\varsigma_X = \kappa_X p$ and $\boldsymbol{\Upsilon}_B$ can modeled as $\boldsymbol{\Upsilon}_B = \kappa_B p \text{diag}[|h_{X_1}(k)|^2, ..., |h_{X_M}(k)|^2]$, where both κ_X, $\kappa_B \geq 0$ characterize levels of hardware impairments at X and Bob, respectively. They commonly remain constants and are closely related to error vector magnitude (EVM), which is in general used to measure the quality of hardware. The relationship between EVM and κ-parameters

Table 3.1: EVM requirements for different modulation methods

Modulation scheme	**Required EVM**
QPSK	0.175
16-QAM	0.125
64-QAM	0.080
256-QAM	0.035

is illustrated by an example: EVM at X can be formulated as

$$\mathbf{EVM}_X = \sqrt{\frac{\mathbb{E}\{|\eta_X(k)|^2\}}{\mathbb{E}\{|s(k)|^2\}}} = \sqrt{\kappa_X}. \tag{3.3}$$

Remark 1 *A small EVM result is required in the transmitter and receiver for correct demodulation when modulation density increases. Table 3.1 illustrates how 3GPP LTE standard EVM requirements for terminal equipment get tighter as modulation density increases. We also notice that for QAM (quadrature amplitude modulation) in 5G (256-QAM initially and up to 1024-QAM in the future), the constellation points are much closer to each other, so a better EVM performance is required. However, this work focuses on the impact of different levels of hardware impairments (for different modulation densities) on authentication performance. Therefore, we set* κ*-parameters in the range* $[0, 0.15^2]$ *(large* κ*-parameters correspond to low-cost constrained devices) to clearly present authentication performance of the proposed scheme.*

Remark 2 *Modeling of the aggregate residual hardware impairments has been supported and validated by many theoretical investigations and measurements (see e.g., [17, 20, 21], and references therein).*

3.2 Proposed Physical Layer Authentication Scheme

The basic principle for the proposed scheme is that channels are location-specific, which has been widely adopted for transmitters authentication to complement and improve traditional security approaches [12, 14, 31, 51]. Most importantly, this is supported by the well-known Jakes model [30], which states that the received signal rapidly decorrelates over a distance of half a wavelength, and that spatial separation of one to two wavelengths leads to independent fading channels. Therefore, it is difficult (if not impossible) for an attacker to generate or accurately model the signal that is transmitted and received by entities. In other words, the channels between different geographic locations decorrelate rapidly in space due to path loss and fading [30, 31, 36]. Moreover, Eve cannot arrive at Alice's previous location for a typical moving speed 1 m/s and time interval of probing channel 3 ms (please refer to [12]). Consequently, the channel between Alice and Bob is independent of that between Eve and Bob, i.e., $\mathbf{h}_A(k)$ is independent of $\mathbf{h}_E(k)$. Meanwhile, the channel for the same transmitter-receiver pair is correlated over time. Hence, location-specific channel can be used to authenticate transmitters. The proposed scheme includes two processes: Channel estimation with hardware impairments process and decision criterion process.

3.2.1 Channel Estimation

If $\mathbf{R}_{X,\text{diag}} = \text{diag}[r_{11}, ..., r_{MM}]$ consists of diagonal elements of $\mathbf{R}_X$, the covariance matrix of $\mathbf{y}_{BX}(k)$ according to (3.2) is denoted as

$$\mathbf{R}_{\mathbf{y}_{BX}} = \mathbb{E}\{\mathbf{y}_{BX}(k)\mathbf{y}_{BX}^H(k)\} = p(1+\kappa_X)\mathbf{R}_X + p\kappa_B\mathbf{R}_{X,\text{diag}} + (\sigma_I^2 + \sigma_N^2)\mathbf{I}. \quad (3.4)$$

Let $\hat{\mathbf{h}}_X(k)$ denote the estimation of $\mathbf{h}_X(k)$ and then by using linear minimum mean square error estimator [17] we have $\hat{\mathbf{h}}_X(k) = s^*(k)\mathbf{R}_X\mathbf{R}_{\mathbf{y}_{BX}}^{-1}\mathbf{y}_{BX}(k)$. Then, we can establish the following lemma on channel estimation with hardware impairments. The

proof is straightforward, and a similar one can be found in [17].

Lemma 1 *$\hat{\mathbf{h}}_X(k)$ can be decomposed as*

$$\hat{\mathbf{h}}_X(k) = \mathbf{h}_X(k) - \boldsymbol{\epsilon}_X(k). \tag{3.5}$$

where $\boldsymbol{\epsilon}_X(k) \in \mathbb{C}^{M\times 1} \sim \mathcal{CN}(\mathbf{0}, \mathbf{R}_{\boldsymbol{\epsilon}_X})$ is estimation error vector and uncorrelated to $\mathbf{h}_t[k]$; and $\mathbf{R}_{\boldsymbol{\epsilon}_X}$ is given by

$$\mathbf{R}_{\boldsymbol{\epsilon}_X} = \mathbb{E}\{\boldsymbol{\epsilon}_X(k)\boldsymbol{\epsilon}_X^H(k)\} = \mathbf{R}_X - p\mathbf{R}_X\mathbf{R}_{\mathbf{y}_{BX}}^{-1}\mathbf{R}_X. \tag{3.6}$$

As observed from (3.4) and (3.6), levels of hardware impairments of different transmitter-receiver pairs lead to different error covariance matrices under the same AWGN and interference. More precisely, a larger level of hardware impairments will lead to a worse estimation error. It is also notable that when κ equals zero, i.e., for ideal hardware, estimation error only comes from AWGN and interference.

3.2.2 Decision Criterion

Based on the above results, Bob can utilize a binary hypothesis test to decide whether the current message is still from legitimate transmitter Alice. In other words, it helps to test whether the current channel estimation at time k is analogous to the previous ones at time $k-1$. Therefore, the hypothesis test can be formulated as

$$\begin{aligned} H_0 &: \ \hat{\mathbf{h}}_X(k) = \hat{\mathbf{h}}_A(k), \\ H_1 &: \ \hat{\mathbf{h}}_X(k) = \hat{\mathbf{h}}_E(k), \end{aligned} \tag{3.7}$$

where the null hypothesis H_0 represents that the current transmitter is still Alice, i.e., $X = A$. In contrast, the alternative hypothesis H_1 represents that the current transmitter is adversary Eve, i.e., $X = E$.

The proposed scheme utilizes location-specific channels to authenticate transmitters, by comparing the difference between the previous and the current channel amplitude with a threshold. This work considers that Bob receives two messages (i.e., frames) at time $k-1$ and time k. The first one received by Bob at time $k-1$ is validated as from Alice by using a standard higher-layer protocol, and thus Bob estimates the channel connecting Alice with him. At time k, Bob can estimate channel connecting a current transmitter (i.e., Alice or Eve) with him through pilot signals. Although the proposed scheme relies on other higher-layer protocols to validate the identity of the previous legitimate transmitter, for subsequent authentication it enables a receiver to quickly differentiate between legitimate and illegitimate transmitters without complete higher-layer processing. In this work, both channel covariance matrices (statistical CSI) associated with Alice and Eve are available for Bob by using some techniques such as geographical information systems and remote sensing information of interest. Then, Bob will implement authentication by comparing the difference between $\hat{\mathbf{h}}_A(k-1)$ and $\hat{\mathbf{h}}_X(k)$ with a threshold.

To achieve effective authentication, it is of great significance to establish the likelihood ratio test (LRT) for the developed hypothesis test. For notational convenience, let $\mathbf{x} = [x_1 \cdots x_M]^T$ denote the difference between the current and previous channel estimations with x_m representing the m^{th} component, i.e., $\mathbf{x} = \hat{\mathbf{h}}_X(k) - \hat{\mathbf{h}}_A(k-1)$, where $\hat{\mathbf{h}}_A(k-1)$ is stored by Bob at time $k-1$. We use $\mathbf{C}_i$ $(i = 0, 1)$ to denote covariance matrices of $\mathbf{x}$ on the two hypotheses.

Lemma 2 *The LRT for the hypothesis test in* (3.7) *is defined as*

$$\mathcal{L}(\mathbf{x}) \triangleq \mathbf{x}^H \Delta \mathbf{Q} \mathbf{x} \underset{H_0}{\overset{H_1}{\gtrless}} \delta, \tag{3.8}$$

$$\mathbf{C}_i = \begin{cases} 2(1-\alpha)\mathbf{R}_A + 2(\mathbf{R}_A - p\mathbf{R}_A\mathbf{R}_{\mathbf{y}_{B,A}}^{-1}\mathbf{R}_A), \ i=0, & \text{(3.10a)} \\ 2\mathbf{R}_A - p\mathbf{R}_A\mathbf{R}_{\mathbf{y}_{BA}}^{-1}\mathbf{R}_A + 2\mathbf{R}_E - p\mathbf{R}_E\mathbf{R}_{\mathbf{y}_{BE}}^{-1}\mathbf{R}_E, \ i=1. & \text{(3.10b)} \end{cases}$$

where $\mathcal{L}(\mathbf{x})$ is sufficient statistic and δ is decision threshold, and $\Delta\mathbf{Q}$ can be given by

$$\Delta\mathbf{Q} = \mathbf{C}_0^{-1}\mathbf{K}\mathbf{C}_1^{-1}, \tag{3.9}$$

where $\mathbf{C}_i(i=0,1)$ is given in (3.10), $\mathbf{C}_1 = \mathbf{C}_0 + \mathbf{K}$, and $\mathbf{K}$ is given by

$$\mathbf{K} = 2\mathbf{R}_E - p\mathbf{R}_E\mathbf{R}_{\mathbf{y}_{BE}}^{-1}\mathbf{R}_E + p\mathbf{R}_A\mathbf{R}_{\mathbf{y}_{BA}}^{-1}\mathbf{R}_A - 2(1-\alpha)\mathbf{R}_A. \tag{3.11}$$

Proof 1 *See Appendix A.1.*

It is important to note that $\mathcal{L}(\mathbf{x})$ is a function of $\mathbf{x}$ and $\Delta\mathbf{Q}$, which has the property that $\mathcal{L}_0(\mathbf{x})$ can be determined as a function of $\mathcal{L}(\mathbf{x})$. Thus, based on the value of $\mathcal{L}(\mathbf{x})$, Bob can discriminate between Alice and Eve.

Remark 3 *To meet extreme data demand growth, it is a promising solution for future wireless systems (e.g., 5G networks) and mmWave communication systems to operate in the frequency range of 30–300 GHz. Higher frequencies adopted in these systems will require shorter inter-site distances to ensure message transmissions, causing changes in fading characteristics. The proposed scheme utilizes location-specific channels to authenticate transmitters. Therefore, slower fading or without fading might contribute to improving authentication performance. This will be proved by numerical results in Section 3.4.3.*

Remark 4 *In massive MIMO systems, spatial diversity leads to channel hardening, meaning that a fading channel behaves as if it were a non-fading channel (please refer to [52] for details). Channel hardening has two significant advantages. One is the*

improved reliability of having a nearly deterministic channel. The other is almost little estimation error for channels realization. Therefore, these advantages allow us to completely exploit location-specific wireless channels to differentiate between the legitimate transmitter and illegitimate one, by taking aggregate residual hardware impairments into account. As shown in Section 3.4.3, less fluctuation in channel gain (i.e., tending to hardening) will obtain better authentication performance.

3.3 Modeling of FA and SD Probabilities

In this section, we first explore the behaviors of the LRT in (3.8) for diverse channel covariance models, and then utilize these behavior results to derive analytical expressions for P_F and P_D.

According to Section 3.1.2, the channel for the same transmitter-receiver pair can be either spatially independent (uncorrelated) or correlated. Against this background, we need to analyze each case in detail to find analytical expressions for P_F and P_D.

3.3.1 Spatially Independent Channel

For spatially independent case, channel components may be independent and identically distributed (IID) or independent but with unequal variances (IUV). We give the following lemmas on distributions of eigenvalues of $\mathbf{C}_i$ under IID and IUV cases.

When the temporally correlated channel components are spatially IID (i.e., spatio-temporal), $\mathbf{R}_X$ can be denoted as $\mathbf{R}_X = \sigma_X^2\mathbf{I}$, where σ_X^2 is the variance of $h_{X,m}$. Then, by substituting $\mathbf{R}_X$ into (3.4), $\mathbf{R}_{\mathbf{y}_{BX}}$ becomes

$$\mathbf{R}_{\mathbf{y}_{BX}} = \lambda_{\mathbf{y}_{BX}}\mathbf{I}, \tag{3.12}$$

where $\lambda_{\mathbf{y}_{BX}} = (p(1+\kappa_X+\kappa_B)\sigma_X^2 + \sigma_I^2 + \sigma_N^2)$.

Lemma 3 *When the temporally correlated channel components are spatially IID,* $\mathbf{C}_i$ *given in* (3.10) *can be further written as*

$$\mathbf{C}_i = \begin{cases} \lambda_{\mathbf{C}_0}\mathbf{I}, \text{ if } i = 0, \\ (\lambda_{\mathbf{C}_0} + \lambda_{\mathbf{K}})\mathbf{I}, \text{ if } i = 0. \end{cases} \tag{3.13}$$

where

$$\lambda_{\mathbf{C}_0} = 2(1-\alpha)\sigma_A^2 + 2(\sigma_A^2 - p\sigma_A^4/\lambda_{\mathbf{y}_{BA}}), \tag{3.14a}$$

$$\lambda_{\mathbf{C}_1} = \lambda_{\mathbf{C}_0} + \lambda_{\mathbf{K}}, \tag{3.14b}$$

$$\lambda_{\mathbf{K}} = 2\sigma_E^2 - \frac{p\sigma_E^4}{\lambda_{\mathbf{y}_{BE}}} + \frac{p\sigma_A^4}{\lambda_{\mathbf{y}_{BA}}} - 2(1-\alpha)\sigma_A^2. \tag{3.14c}$$

Proof 2 *When the temporally correlated channel components are spatially IID,* $\mathbf{R}_A$*,* $\mathbf{R}_E$*,* $\mathbf{R}_{\mathbf{y}_{BA}}$*, and* $\mathbf{R}_{\mathbf{y}_{BE}}$ *are diagonal matrices. Based on* (3.5)*,* (3.10)*, and* (3.12)*, one can see that* $\mathbf{C}_i$ *are also diagonal matrices. Substituting* $\mathbf{R}_X = \sigma_X^2\mathbf{I}$ *and* $\mathbf{R}_{\mathbf{y}_{BX}} = \lambda_{\mathbf{y}_{BX}}\mathbf{I}$ *into* (3.10) *yields* (3.13)*.*

When the temporally correlated channel components are spatially IUV, $\mathbf{R}_X$ can be denoted as

$$\mathbf{R}_X = \text{diag}[\sigma_{X,1}^2, ..., \sigma_{X,M}^2]. \tag{3.15}$$

Substituting (3.15) into (3.4), $\mathbf{R}_{\mathbf{y}_{BX}}$ becomes

$$\mathbf{R}_{\mathbf{y}_{BX}} = \text{diag}[\lambda_{\mathbf{y}_{BA,1}}, ..., \lambda_{\mathbf{y}_{BA,M}}], \tag{3.16}$$

where $\lambda_{\mathbf{y}_{BA,m}} = \big(p(1+\kappa_X+\kappa_B)\sigma_{A,m}^2 + \sigma_I^2 + \sigma_N^2\big)$.

Lemma 4 *When the temporally correlated channel components are spatially IUV,* $\mathbf{C}_i$

given in (3.10) *can be written as*

$$\mathbf{C}_i = \text{diag}[\lambda_{\mathbf{C}_i,1}, ..., \lambda_{\mathbf{C}_i,M}], \tag{3.17}$$

where

$$\lambda_{\mathbf{C}_0,m} = (4-2\alpha)\sigma_{A,m}^2 - \frac{2p\sigma_{A,m}^4}{\lambda_{\mathbf{y}_{BA,m}}}, \tag{3.18a}$$

$$\lambda_{\mathbf{C}_1,m} = \lambda_{\mathbf{C}_0,m} + \lambda_{\mathbf{K},m}, \tag{3.18b}$$

$$\lambda_{\mathbf{K},m} = 2\sigma_{E,m}^2 - 2(1-\alpha)\sigma_{A,m}^2 - \frac{p\sigma_{E,m}^4}{\lambda_{\mathbf{y}_{BE,m}}} + \frac{p\sigma_{A,m}^4}{\lambda_{\mathbf{y}_{BA,m}}}. \tag{3.18c}$$

Proof 3 *When the temporally correlated channel components are spatially IUV, all* $\mathbf{R}_A$, $\mathbf{R}_E$, $\mathbf{R}_{\mathbf{y}_{BA}}$, *and* $\mathbf{R}_{\mathbf{y}_{BE}}$ *are diagonal matrices. Thus, based on* (3.5), (3.10), *and* (3.12), *we know that* $\mathbf{C}_i$ *is also diagonal matrix. Substituting* (3.15) *and* (3.16) *into* (3.10), *we can obtain* (3.17).

Based on the above lemmas, P_F and P_D under IID and IUV cases are summarized in the following theorem.

Theorem III.1 *Consider the uplink massive MIMO system with hardware impairments over spatially independent time-varying channel components. Under IID and IUV cases* P_F *and* P_D *can be given in* (3.19) *and* (3.20), *respectively, where* $\lambda_{\mathbf{C}_0}$ *and* $\lambda_{\mathbf{K}}$ *are given in Lemma 3,* $a_m = \frac{\lambda_{\mathbf{K},m}}{\lambda_{\mathbf{C}_0,m}+\lambda_{\mathbf{K},m}}$ *and* $c_m = \frac{\lambda_{\mathbf{K},m}}{\lambda_{\mathbf{C}_0,m}}$, *in which* $\lambda_{\mathbf{C}_0,m}$ *and* $\lambda_{\mathbf{K},m}$ *are given in Lemma 4, and* δ *is a decision threshold.*

Proof 4 *See Appendix A.2.*

These results show that we can calculate P_F and P_D through standard mathematical functions under the temporally correlated and spatially independent channel components. It is interesting that $\Delta\mathbf{Q}$ is a diagonal matrix (since $\mathbf{C}_i$ is a diagonal

$$P_F = \begin{cases} \mathbf{\Gamma}_{\chi^2_{2M}}\left(\left(\frac{\lambda_{\mathbf{C}_0}}{\lambda_{\mathbf{K}}}+1\right)\delta\right), \text{ if IID}, & (3.19a) \\ \sum_{m=1}^{M}\left[\prod_{\substack{i=1 \\ i\neq m}}^{M}\frac{a_m}{a_m-a_i}\right]\exp\left(-\frac{\delta}{a_m}\right), \text{ if IUV}. & (3.19b) \end{cases}$$

$$P_D = \begin{cases} \mathbf{\Gamma}_{\chi^2_{2M}}\left(\frac{\lambda_{\mathbf{C}_0}}{\lambda_{\mathbf{K}}}\delta\right), \text{ if IID}, & (3.20a) \\ \sum_{m=1}^{M}\left[\prod_{\substack{i=1 \\ i\neq m}}^{M}\frac{c_m}{c_m-c_i}\right]\exp\left(-\frac{\delta}{c_m}\right), \text{ if IUV}. & (3.20b) \end{cases}$$

matrix). These analytical results enable us to evaluate the performance of the proposed scheme taking hardware impairment into account under spatially independent time-varying channel components.

3.3.2 Spatially Correlated Channel

In practice, the channels between different antennas are spatially correlated due to the following reasons. First, it is well-known that spatial correlation is relevant to antenna separation, which is rarely larger owing to large-scale nature of massive MIMO systems. Second, channels may tend to a point in some directions [17]. Third, for antenna, there exists spatially dependent pattern when setting short antenna space and large angular spread, causing channels between adjacent antennas spatially correlated [17, 53, 54]. Therefore, for massive MIMO systems, spatial correlation properties of channels between adjacent antennas always exist. We generate channel covariance matrix $\mathbf{R}_X$ ($X = \{A, E\}$) via exponential correlation model in [53]. In

fact, it is expressed by a $M \times M$ complex Toeplitz matrix [55]. That is,

$$\mathbf{R}_X = \sigma_X^2 \begin{bmatrix} 1 & \rho_X^* & \cdots & (\rho_X^*)^{M-1} \\ \rho_X & 1 & \cdots & (\rho_X^*)^{M-2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_X^{M-1} & \rho_X^{M-2} & \cdots & 1 \end{bmatrix}, \tag{3.21}$$

where σ_X^2 and ρ_X (here $0 < |\rho_X| \leq 1$, and when $|\rho_X| = 0$, channel components are spatially uncorrelated) are arbitrary scaling factor and correlation coefficient between adjacent antennas, respectively. Note that the eigenvalue spread in $\mathbf{R}_X$ depends on $|\rho_X|$. Hence, we need to consider different $|\rho_X|$ to derive exact expressions for P_F and P_D. Combining (3.10) and (3.11), we will obtain the following lemma.

When the temporally correlated channel components are fully correlated in space (i.e., $|\rho_X| = 1$), we have $\mathbf{R}_X = \sigma_X^2 \boldsymbol{\rho}_X \boldsymbol{\rho}_X^H$, where $\boldsymbol{\rho}_X = [1 \cdots 1^{M-1}]^T$. We use $\lambda_{X,m}$ to denote the m^{th} eigenvalue of $\mathbf{R}_X$, and then we have $\lambda_{X,1} = M\sigma_X^2$ and the remaining eigenvalues are zero, i.e., $\lambda_{X,2} = \cdots = \lambda_{X,M} = 0$. Thus, we have

$$\mathbf{R}_X = \text{diag}[M\sigma_X^2, 0, ..., 0], \tag{3.22}$$

Thus, we have $\mathbf{R}_A = \text{diag}[M\sigma_A^2, 0, ..., 0]$ and $\mathbf{R}_E = \text{diag}[M\sigma_E^2, 0, ..., 0]$. Substituting $\mathbf{R}_A$ and $\mathbf{R}_E$ into (3.4) yields

$$\lambda_{\mathbf{y}_{BX,1}} = p(1 + \kappa_X)M\sigma_X^2 + p\kappa_B\sigma_X^2 + \sigma_I^2 + \sigma_N^2. \tag{3.23}$$

Lemma 5 *When the temporally correlated channel components are fully correlated in space (i.e., $|\rho_X| = 1$), $\mathbf{C}_i$ given in* (3.10) *becomes*

$$\mathbf{C}_i = \text{diag}[\lambda_{\mathbf{C}_i,1}, 0, ..., 0], \tag{3.24}$$

where

$$\lambda_{\mathbf{C}_0,1} = (4-2\alpha)M\sigma_A^2 - \frac{2pM^2\sigma_A^4}{\lambda_{\mathbf{y}_{BA,1}}}, \tag{3.25a}$$

$$\lambda_{\mathbf{C}_1,1} = \lambda_{\mathbf{C}_0,1} + \lambda_{\mathbf{K},1}, \tag{3.25b}$$

$$\lambda_{\mathbf{K},1} = M\left(2\sigma_E^2 - 2(1-\alpha)\sigma_A^2 - \frac{pM\sigma_E^4}{\lambda_{\mathbf{y}_{BE,1}}} + \frac{pM\sigma_A^4}{\lambda_{\mathbf{y}_{BA,1}}}\right). \tag{3.25c}$$

Proof 5 *When the temporally correlated channel components are fully correlated in space, i.e.,* $|\rho_X| = 1$*, according to* (3.10) *and* (3.22)*,* $\mathbf{C}_i$ *has one non-zero eigenvalue and* $M-1$ *zero eigenvalues. Combining* (3.22) *and* (3.23)*, we can obtain* (3.24)*.*

When $0 < |\rho_{\mathrm{X}}| < 1$, the eigenvalues of $\mathbf{R}_X$ are distinct and can be found numerically. Let the eigendecomposition of $\mathbf{R}_X$ be $\mathbf{R}_X = \mathbf{u}_X\mathbf{\Lambda}_X\mathbf{u}_X^H$, where $\mathbf{u}_X$ is an $M \times M$ matrix [56]; and $\mathbf{\Lambda}_X = \mathrm{diag}[\lambda_{X,1}, ..., \lambda_{X,M}]$ with $\lambda_{X,m}$ denoting the m^{th} eigenvalue of $\mathbf{R}_X$. From (3.4), we can see that the eigendecomposition of $\mathbf{R}_{\mathbf{y}_{BX}}$ is $\mathbf{R}_{\mathbf{y}_{BX}} = \mathbf{u}_X\mathbf{\Lambda}_{\mathbf{y}_{BX}}\mathbf{u}_X^H$, where $\mathbf{\Lambda}_{\mathbf{y}_{BX}} = \mathrm{diag}[\lambda_{\mathbf{y}_{BX,1}}, ..., \lambda_{\mathbf{y}_{BX,M}}]$ with $\lambda_{\mathbf{y}_{BX,m}} = p(1+\kappa_X)\lambda_{X,m} + p\kappa_B\sigma_X^2 + \sigma_I^2 + \sigma_N^2$.

To analyze the behavior of the LRT defined in (3.8) under the non-diagonal channel covariance model, we need to transform $\Delta\mathbf{Q}$ to a diagonal matrix by a two-step transformation due to different correlation coefficients for $\mathbf{R}_A$ and $\mathbf{R}_E$ (i.e., $|\rho_A| \neq |\rho_E|$).

We first do eigendecomposition for $\mathbf{C}_0$, that is, $\mathbf{C}_0 = \mathbf{u}_A\mathbf{\Lambda}_0\mathbf{u}_A^H$, where $\mathbf{\Lambda}_0 = \mathrm{diag}[\lambda_{\mathbf{C}_0,1}, ..., \lambda_{\mathbf{C}_0,M}]$ with $\lambda_{\mathbf{C}_0,m}$ representing the m^{th} eigenvalue of $\mathbf{C}_0$. It is easily to see from (3.21) that the rank of $\mathbf{\Lambda}_0$ is M. We define decorrelating transformation $\mathbf{w}^H \triangleq [\mathbf{\Lambda}_0]^{-\frac{1}{2}}\mathbf{u}_A^H$, and then apply it to $\mathbf{x}$ on H_0 to obtain $\mathbf{x}_\mathbf{w} = \mathbf{w}^H\mathbf{x}$. Since $\mathbf{R}_A$ is Hermitian, we have $\mathbf{u}_A^H = \mathbf{u}_A^{-1}$. The covariance matrix of $\mathbf{x}_\mathbf{w}$ on H_0 is $\mathbf{I}$. On H_1, its

covariance matrix is denoted by

$$\mathbf{R}_{1\mathbf{w}} = \mathbb{E}\{\mathbf{x}_{\mathbf{w}}\mathbf{x}_{\mathbf{w}}^H|H_1\} = \mathbf{w}^H\mathbf{C}_1\mathbf{w} = \mathbf{w}^H\mathbf{D}\mathbf{w} + \mathbf{I}. \tag{3.26}$$

Let $\mathbf{R}_{\mathbf{Dw}} = \mathbf{w}^H\mathbf{D}\mathbf{w}$, and it is a non-diagonal matrix because $\mathbf{D}$ contains $\mathbf{R}_E$ and thus $\mathbf{w}^H$ could not decorrelate $\mathbf{D}$. Therefore, we now need to do an eigendecomposition of $\mathbf{R}_{\mathbf{Dw}}$:

$$\mathbf{R}_{\mathbf{Dw}} = \mathbf{u}_{\mathbf{Dw}}\mathbf{\Lambda}_{\mathbf{Dw}}\mathbf{u}_{\mathbf{Dw}}^H, \tag{3.27}$$

where $\mathbf{u}_{\mathbf{Dw}}$ is an $M \times M$ modal matrix; $\mathbf{\Lambda}_{\mathbf{Dw}} = \text{diag}[\lambda_{\mathbf{Dw},1}, ..., \lambda_{\mathbf{Dw},M}]$ with $\lambda_{\mathbf{Dw},m}$ denoting the m^{th} eigenvalue of $\mathbf{R}_{\mathbf{Dw}}$. It is noticed that $\mathbf{R}_{\mathbf{Dw}}$ may not be full rank matrix. Hence, we augment the eigenvectors if its rank is not M. The eigendecomposition of $\mathbf{R}_{1\mathbf{w}}$ is $\mathbf{R}_{1\mathbf{w}} = \mathbf{u}_{\mathbf{Dw}}[\Lambda_{\mathbf{Dw}} + \mathbf{I}]\mathbf{u}_{\mathbf{Dw}}^H = \mathbf{u}_{\mathbf{Dw}}\mathbf{\Lambda}_{1\mathbf{w}}\mathbf{u}_{\mathbf{Dw}}^H$, where $\mathbf{\Lambda}_{1\mathbf{w}} = \text{diag}[\lambda_{\mathbf{Dw},1} + 1, ..., \lambda_{\mathbf{Dw},M} + 1]$.

Based on the above lemmas, P_F and P_D under the spatially correlated channel are summarized in the following theorem.

Theorem III.2 *Consider the uplink massive MIMO system with hardware impairments over spatially correlated time-varying channel. Under the spatially correlated channel case, P_F and P_D of the proposed scheme can be given in* (3.28) *and* (3.29)*, respectively, where $\lambda_{\mathbf{wu},m} = \frac{\lambda_{\mathbf{Dw}_m}}{\lambda_{\mathbf{Dw},m}+1}$.*

Proof 6 *See Appendix A.3.*

This indicates that we can evaluate the authentication performance of the proposed scheme for the channel following the zero-mean complex Gaussian distribution with an arbitrary covariance matrix. The key to deriving the closed-form expressions for P_F and P_D is that complex eigenvalue corresponds to two equal real eigenvalues. Also,

$$P_F = \begin{cases} \exp\left(-\delta\left(1+\dfrac{\lambda_{\mathbf{C}_0,1}}{\lambda_{\mathbf{K},1}}\right)\right), \text{ if } |\rho_A| = |\rho_E| = 1, & \text{(3.28a)} \\ \displaystyle\sum_{m=1}^{M}\left[\prod_{\substack{i=1 \\ i\neq m}}^{M}\frac{\lambda_{\mathbf{wu},m}}{\lambda_{\mathbf{wu},m}-\lambda_{\mathbf{wu},i}}\right]\exp\left(-\frac{\delta}{\lambda_{\mathbf{wu},m}}\right), \text{ if } 0<|\rho_{\mathrm{A}}|,|\rho_{\mathrm{E}}|<1 & \text{(3.28b)} \end{cases}$$

$$P_D = \begin{cases} \exp\left(-\dfrac{\delta\lambda_{\mathbf{C}_0,1}}{\lambda_{\mathbf{K},1}}\right), \text{ if } |\rho_A| = |\rho_E| = 1, & \text{(3.29a)} \\ \displaystyle\sum_{m=1}^{M}\left[\prod_{\substack{i=1 \\ i\neq m}}^{M}\frac{\lambda_{\mathbf{Dw},m}}{\lambda_{\mathbf{Dw},m}-\lambda_{\mathbf{Dw},i}}\right]\exp\left(-\frac{\delta}{\lambda_{\mathbf{Dw},m}}\right), \text{ if } 0<|\rho_{\mathrm{A}}|,|\rho_{\mathrm{E}}|<1 & \text{(3.29b)} \end{cases}$$

---

utilizing eigendecomposition and diagonalizing operations we can transform an arbitrary channel covariance matrix model to the case in which $\Delta\mathbf{Q}$ is a diagonal matrix whose elements are functions with respect to eigenvalues. By studying various models, we can obtain analytical performance results that enable us to understand how channel models (or channel covariance matrix models) affect authentication performance.

3.3.3 Unknown Parameters

If Bob has no knowledge of parameters such as $\mathbf{R}_A$, $\mathbf{R}_E$, α, κ_A, κ_E, and κ_B, he can exploit the following LRT to identify the current transmitter

$$\mathcal{L}(\mathbf{x}) = \frac{1}{\sigma_N^2+\sigma_I^2}\sum_{m=1}^{M}|x_m|^2 = \frac{1}{\sigma_N^2+\sigma_I^2}\sum_{m=1}^{M}|\hat{h}_{X,m}(k)-\hat{h}_{A,m}(k-1)|^2 \underset{H_0}{\overset{H_1}{\gtrless}} \delta. \qquad (3.30)$$

In this case, we only have numerical results for P_F and P_D (which will be illustrated in Section 3.4.3).

Table 3.2: System parameters affecting authentication performance

Parameter	Description
SINR	Signal to interference plus noise ratio
κ	The ratio of the level of hardware impairment for E and A
γ	The ratio of locally averaged channel gains for E-B and A-B
α	Temporal correlation coefficient of $\mathbf{h}_A$
ρ	Spatial correlation coefficient between adjacent antennas
M	The number of base station antennas

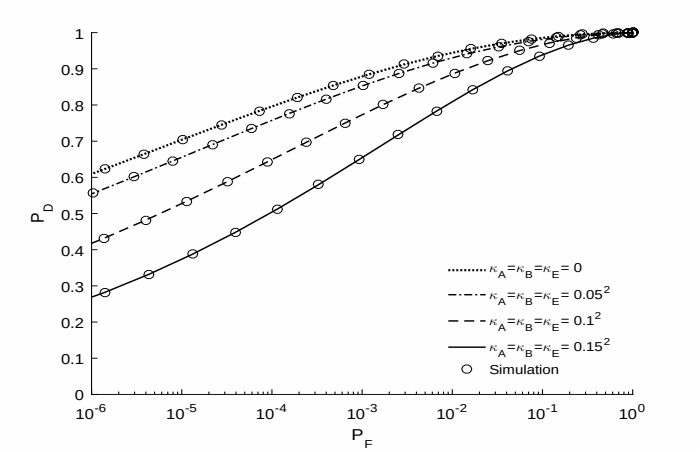
3.4 Numerical Results

In this section, we verify theoretical results through simulations and reveal how system parameters affect the authentication performance of the proposed scheme.

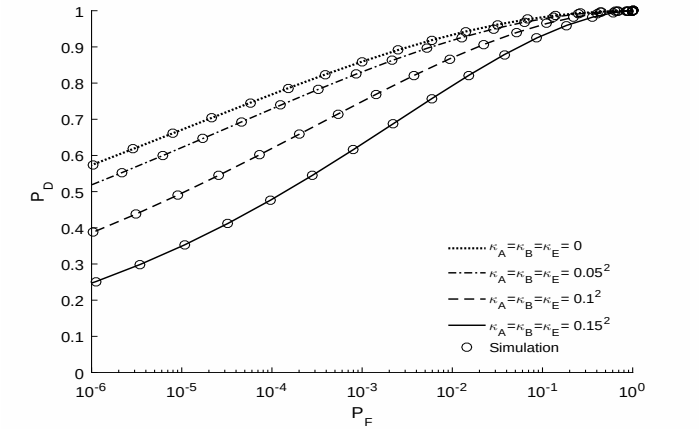
3.4.1 System Parameters and Simulation Settings

System parameters that determine authentication performance (P_F, P_D) are listed in Table 5.1. In particular, signal to interference plus noise ratio (SINR) is defined as $\text{SINR} = p\frac{\text{tr}(\mathbf{R})}{M(\sigma_I^2+\sigma_N^2)}$. The ratio of the levels of hardware impairments for Eve and Alice is defined as $\kappa = \frac{\kappa_E}{\kappa_A}$. According to the EVM ranges introduced in Section 3.1.3, we consider four typical levels of impairments: κ_A, κ_B, $\kappa_E \in \{0, 0.05^2, 0.1^2, 0.15^2\}$. Therefore, if we fix κ_A, we can adjust κ_E to achieve a specified κ. Moreover, $\gamma = \frac{\text{tr}(\mathbf{R}_E)}{\text{tr}(\mathbf{R}_A)}$ denotes the ratio of locally averaged channel gains for Alice-Bob and Eve-Bob. In addition, α is temporal correlation coefficient of $\mathbf{h}_A$, and ρ_A and ρ_E are spatial correlation coefficients between adjacent antennas for $\mathbf{h}_A$ and $\mathbf{h}_E$, respectively. In our simulation, we assume $\rho_A = \rho_E = \rho$.

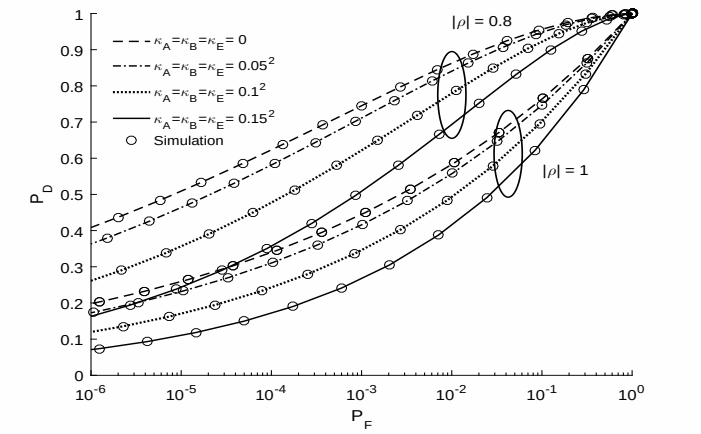
To validate the derived results of P_F and P_D, we develop a dedicated simulator based on Matlab. The simulation method in [57] and exponential correlation model in [53] are exploited to generate time-varying MIMO channels and covariance matrices of such channels, respectively. The quantity of temporal correlation of underlying



(a) IID



(b) IUV



(c) Spatial correlation.

Figure 3.2: ROC curves of the proposed scheme with the settings ($\gamma = 0$ dB, $\kappa = 1.0^2$, $M = 5$, SINR $= 10$ dB, and $\alpha = 0.9$).

channels depends on normalized Doppler frequency, which is determined by the speed of transmitter and carrier frequency. Therefore, for a given carrier frequency, the normalized Doppler frequency is a function of the transmitter speed only. We consider three fading channels (case I: slow-fading with $\alpha = 1$; case II: fast-fading with $\alpha = 0.9$; and case III: faster-fading with $\alpha = 0.8$) [58]. For Monte-Carlo experiments, 10^5 independent trials are conducted to obtain average results.

3.4.2 Model Validation

For simplicity, we assume $\kappa_A = \kappa_B = \kappa_E$. To verify our analytical results, we plot the receiver operating characteristic (ROC) curves in Fig. 3.2. Fig. 3.2 shows that the simulation results match nicely with the theoretical ones for spatially independent (IID, IUV) and spatially correlated channel components, so our theoretical results can be used to accurately model P_F and P_D for an arbitrary channel covariance matrix. As observed from Fig. 3.2 that for three different channel covariance matrix models P_D improves as P_F increases. According to Neyman-Pearson criterion, it is required to make P_D as large as possible for a given P_F constraint (commonly below 10^{-1}).

Also, we can see from Fig. 3.2 that for three channel covariance matrix models, P_D decreases with the levels of impairments when P_F is fixed. In particular, when $\kappa_A = \kappa_B = \kappa_E = 0$ (i.e., ideal hardware), we have the largest P_D for three channel covariance matrix cases; when $\kappa_A = \kappa_B = \kappa_E = 0.15^2$, we have the smallest P_D; for a fixed P_F, the difference between the largest P_D and smallest one can approach 0.3 under the same channel covariance matrix. This clearly reveals that hardware impairments greatly deteriorate authentication performance.

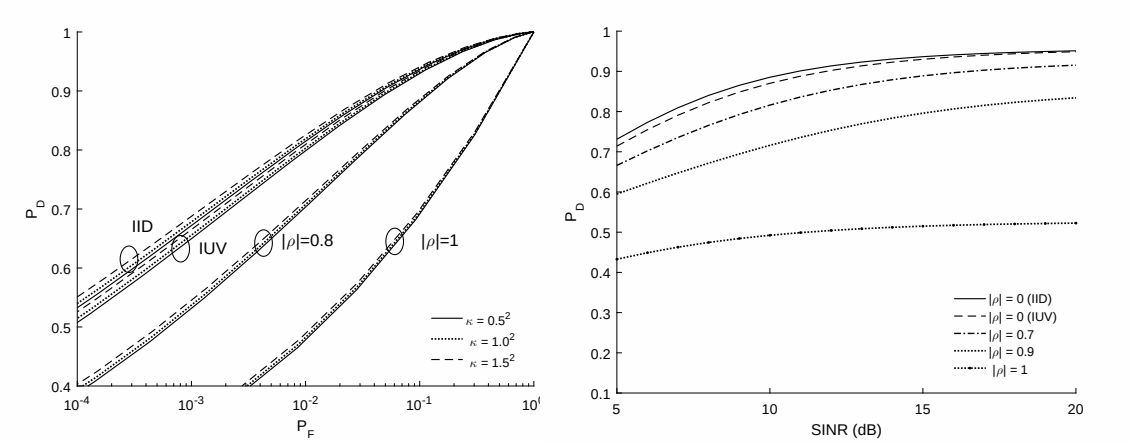
From Fig. 3.2, we see that the choice of covariance model has a significant impact the performance. The reason is that: for the spatially uncorrelated covariance model (Fig. 3.2(a) and Fig. 3.2(b)), we have $2M$ real observations of channel component estimation; decreasing ρ results in lower spatial correlation and thus improves P_D;

while for the spatially correlated covariance model (Fig. 3.2(c)) we have no more than $2M$ real observations, especially when $\rho = 1$ we only have two real observations. It is proved in [59] that the quantity of spatial correlation determines the number of observations for channel component estimation and this is consistent with our results.

3.4.3 Authentication Performance Analysis

Based on theoretical models for P_F and P_D, we explore how system parameters (e.g., κ, SINR, γ, α, and M) affect authentication performance under diverse channel covariance matrix models. Meanwhile, we also examine performance under unknown parameters case via numerical simulations.

We first explore how κ affect the performance for both scenarios (spatially uncorrelated and correlated channel components). We summarize in Fig. 3.3(a) the ROC curves with some representative values of κ for spatially uncorrelated and correlated channel components. As shown in Fig. 3.3(a) that for all channel covariance matrix models, performance monotonically improves as κ increases. In particular, when $\kappa = 1.5^2$, the performance outperforms others; when $\kappa = 0.5^2$, we have the worst performance. In other words, comparing with the legitimate transmitter, the illegitimate one with lager level of impairments is easier to be detected. This tells us that we should choose hardware with smaller level of impairments for secure wireless communications.

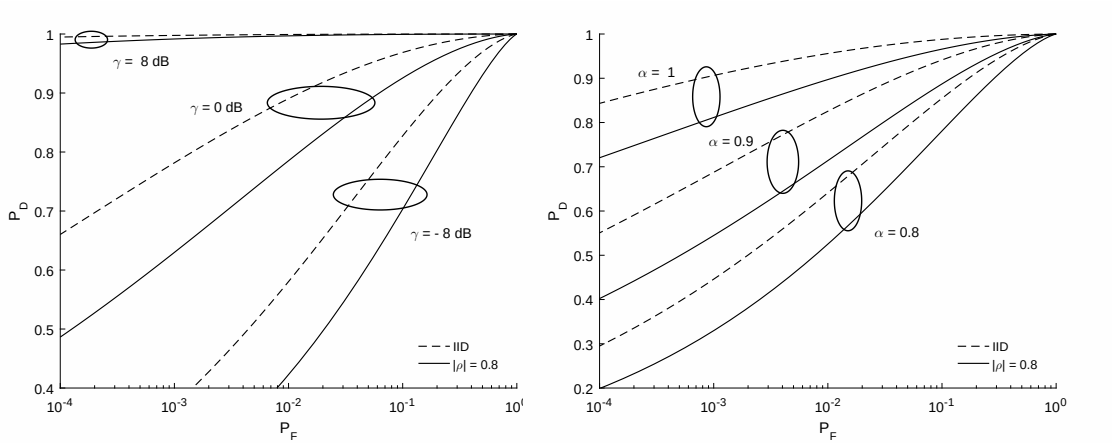


(a) Impact of κ on performance under SINR = 10 dB

(b) P_D vs. SINR under $\kappa = 1.0^2$ and $P_F = 10^{-2}$

Figure 3.3: Authentication performance with the settings ($\gamma = 0$ dB, $M = 5$, SINR $= 10$ dB, $\alpha = 0.9$).

Next, we investigate the impact of SINR on P_D for a fixed P_F. Fig. 3.3(b) illustrates how P_D varies with SINR with the settings ($\gamma = 0$ dB, $\kappa = 1.0^2$, $M = 5$, and $P_F = 10^{-2}$). We can see that under a fixed P_F, increasing SINR leads to different tendencies of P_D for different channel model. In particular, P_D improves as SINR increases; the curves for spatially uncorrelated channel model (i.e., $|\rho| = 0$ for IID and IUV) have better slope than that for spatially correlated channel model. For spatially correlated channel model, the curves for $\rho = 0.7$ and $\rho = 0.9$ have the same slope while that for $\rho = 1$ exhibits the smallest slope. This is because more concentrated channel components in a lower dimensional subspace lead to insufficient observations. This reveals that better P_D performance is achieved as $|\rho| \to 0$, since channel components are more evenly distributed throughout the M-dimensional observation space. Increasing transmit power can improve performance for both spatially uncorrelated and correlated models. It is notable, however, that for general wireless networks applications, transmit power is limited to a certain level due to energy constraint and interference requirement.



(a) Impact of γ on ROC curve under $\alpha = 0.9$ (b) Impact of α on ROC curve under $\gamma = 0$ dB

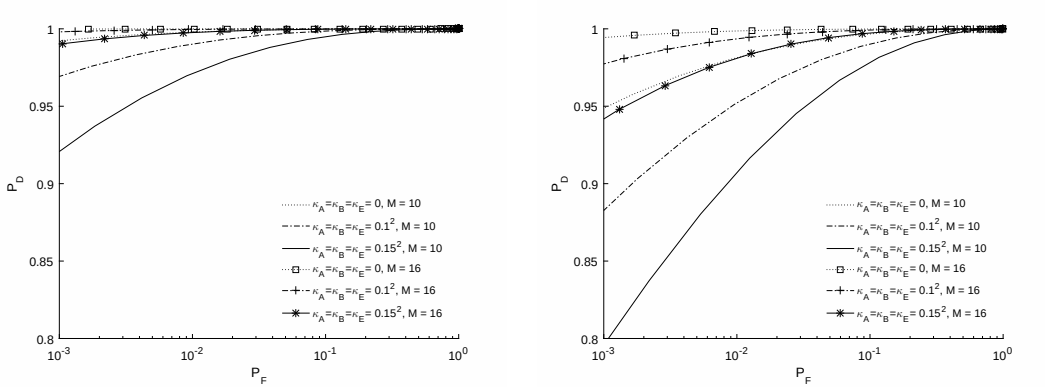
Figure 3.4: Impacts of (γ, α) on ROC curve with the settings (SINR = 10 dB, $\kappa = 1.5^2$, $M = 5$).

Fig. 3.4(a) shows how performance varies with $\gamma \in \{-8 \text{ dB}, 0 \text{ dB}, 8 \text{ dB}\}$, given that SINR = 10 dB, $\kappa = 1.5^2$, $M = 5$, and $\alpha = 0.9$. It is interesting to see from Fig. 3.4(a) that for both channel covariance matrix models, the performance monotonically rises as γ increases. More specifically, when $\gamma = 8$ dB, we have the best performance while when $\gamma = -8$ dB we have the lowest one. This clearly indicates that if Eve is closer to Bob, she might be successfully detected by Bob.

Fig. 3.4(b) demonstrates the impact of channel fading status on authentication performance for spatial independence (IID) and correlation ($|\rho| = 0.8$) models, given that $\gamma = 0$ dB, SINR = 10 dB, $\kappa = 1.5^2$, and $M = 5$. As seen from Fig. 3.4(b), the authentication performance under case I outperforms that under the others (case II and case III), while the scheme under case III provides the worst performance. This indicates that channel-based authentication scheme can effectively differentiate between Alice and Eve, while it might not work well in a highly dynamic environment.

Now, we present in Fig. 3.5 the impact of $M \in \{10, 16\}$ on authentication performance for IID and $|\rho| = 0.8$, given that $\gamma = 0$ dB, SINR = 10 dB, $\alpha = 0.9$, and $\kappa_A = \kappa_B = \kappa_E \in \{0, 0.1^2, 0.15^2\}$. The main observation from Fig. 3.5 is that the

choice of channel covariance model has a large impact on performance. Moreover, for a given covariance model, performance improves as M increases. When $M = 16$ under IID case, the proposed scheme for different levels of hardware impairments has nearly indistinguishable performance (P_D approaching 1), indicating that the degrading of performance due to hardware impairments vanishes asymptotically in large-dimensional vector space.



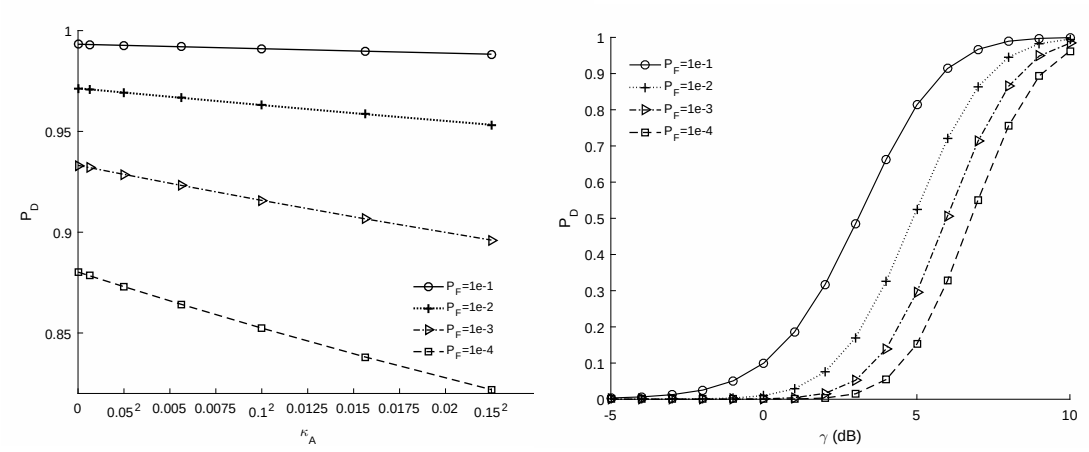
(a) IID (b) $|\rho| = 0.8$

Figure 3.5: Impact of $M \in \{10, 16\}$ on performance, given that $\gamma = 0$ dB, SINR = 10 dB, $\alpha = 0.9$, and $\kappa_A = \kappa_B = \kappa_E \in \{0, 0.1^2, 0.15^2\}$.

Furthermore, we explore how P_D varies with κ_A for IID case with the settings ($\kappa = 1.0^2$, SINR = 5 dB, $\alpha = 0.9$, $\gamma = 0$ dB, and $M = 5$). Fig. 3.6(a) shows that for a given P_F, P_D reduces monotonously when κ_A varies from 0 to 0.15^2. This reveals that within the range of κ_A, aggregate residual hardware impairments can always be utilized to identify transmitters, and a higher aggregate level of impairments leads to a lower authentication performance.

Finally, we investigate the authentication performance of the proposed scheme under the unknown parameter case in Fig. 3.6(b) via numerical simulations. Fig. 3.6(b) demonstrates that P_D vs. γ under the unknown parameter case with the settings (SINR = 10 dB, $M = 5$). At low γ, P_D tends to zero for a given P_F. However, at high γ, P_D rises when γ increases for a given P_F. This means that when Eve is close

to Bob, she might be easily detected by Bob; when being away from Bob, she might impersonate Alice successfully to send possible aggressive message into the network without being detected. In other words, although Bob has no knowledge of system parameters (such as $\mathbf{R}_A$, $\mathbf{R}_E$, α, κ_A, κ_E, and κ_B), she could still identify the current transmitter by using the LRT given in (3.30) when γ is above a certain value. It shows that the proposed scheme has a certain scalability in the case when the base station is "blind" on some systems parameters. We also notice that by setting a high P_F, we can obtain a high P_D for the unknown parameter case. Nevertheless, a high P_F implies low robustness of the proposed scheme. Therefore, we should set P_F properly to achieve a desired authentication performance in specific massive MIMO applications.



(a) P_D vs. κ_A under $\kappa = 1.0^2$, $\gamma = 0$ dB (b) P_D vs. γ under unknown parameter case

Figure 3.6: Authentication performance with the settings (SINR = 10 dB, $M = 5$).

3.5 Summary

We proposed a channel-based authentication scheme for massive MIMO systems with different levels of hardware impairments, and investigated its authentication behaviors. False alarm and detection probabilities were theoretically analyzed with hypothesis testing and matrix transformation approaches. Analytical results were

validated via Monte Carlo simulations, showing that analytical and numerical results match each other well under different channel covariance matrix models. Our results show that authentication performance is clearly deteriorated by hardware impairments, with a nontrivial impact from the choice of antenna patterns.

Notice that multiple hardware impairments (such as I/Q imbalance and phase noise) can be effectively utilized to authenticate transmitters. While in this work, their effects have been taken into account by using κ-parameters. Nevertheless, considering a single specific (rather than the aggregated) hardware impairment (e.g., I/Q imbalance) for authentication is an interesting research topic for our future work to further explore how these hardware impairments can be used to improve the security of massive MIMO systems.

CHAPTER IV

Physical Layer Authentication Jointly Utilizing Channel and Phase Noise in MIMO Systems

In this chapter, we investigate the channel-based authentication solution which not only exploits location-specific wireless channels but also utilizes transmitter-specific hardware impairments for authentication, and propose an improved channel-based scheme jointly utilizing channel gain and phase noise in heterogeneous MIMO systems. Three properties of the proposed scheme: covertness, robustness, and security, are analyzed in detail. By using a maximum-likelihood estimator (MLE) and extended Kalman filter (EKF), we estimate channel gains and phase noise, and formulate variances of estimation errors. We also quantize the temporal variations of channel gains and phase noise through the developed quantizers. Based on quantization results and theories of hypothesis testing and stochastic process, we then derive the closed-form expressions for false alarm and missed detection probabilities with the consideration of quantization errors. Simulations are carried out to validate the theoretical results of the two probabilities. Based on theoretical models, we further demonstrate that the proposed scheme makes it possible for us to flexibly control authentication performance by adjusting parameters (such as channel gain threshold, phase noise threshold, and decision threshold) to achieve a required authentication performance in specific MIMO applications.

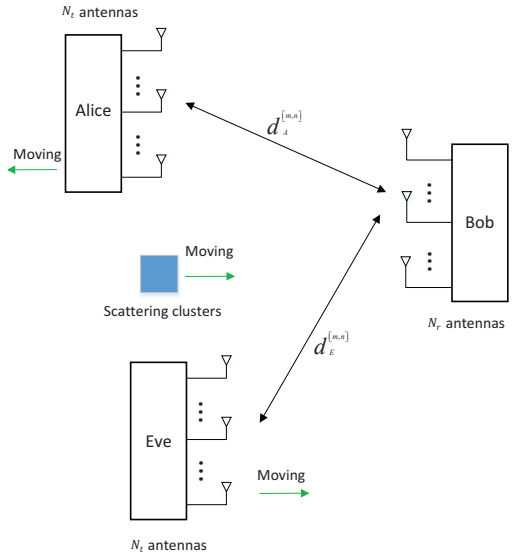


Figure 4.1: A MIMO system consisting of Alice with N_t antennas, Eve with N_t antennas, and Bob with N_r antennas, which are geographically separated and in a rich scattering environment. Entities (e.g., Alice and Eve) and/or scatters are moving.

4.1 System Model

4.1.1 Network Model

Similar to previous works [12–14], we consider a MIMO system consisting of three different entities: Alice and Eve with N_t antennas and Bob with N_r antennas, which are geographically widely separated and in a rich scattering environment, as shown in Fig. 4.1. Alice is a legitimate transmitter and Bob is an intended receiver. While Eve serves as an active attacker, who not only can overhear all the signals transmitted from Alice and Bob, but also can inject aggressive signals or replay signals transmitted from Alice into the network, by using the identity of Alice. Frame-by-frame message transmission is considered. Generally, confidential information such as estimation techniques and authentication schemes employed by Bob cannot be easily obtained by Eve. However, Eve is assumed to know some repeatedly used and publicly known information such as training sequences and pilot symbols [9] due to the wireless broadcasting nature, as well as frame structure by analyzing the transmitted

signals from Alice [13]. Eve cannot arrive at Alice's previous location for the typical moving speed 1 m/s, and time interval of probing channel is set as 3 ms (please refer to [12]). All entities in the network operate in the half-duplex mode. Each antenna is equipped with an independent oscillator.

Suppose that Bob receives two frames at time $k-1$ and time k. The one received at time $k-1$ is from Alice, which is validated through a cryptographic authentication at the application layer [13]. Based on this authentication, Bob measures the channel gains and phase noise parameters at time $k-1$. Bob needs to decide whether the received frame at time k is still from Alice. This chapter focuses on devising a simple and flexible physical layer authentication scheme to fight against spoofing attacks (e.g., impersonation and/or replay attacks), where either Alice or Eve transmits a signal to Bob but simultaneous transmission of both is not considered. Carrier sense multiple access/collision avoidance (CSMA/CA) can be employed to ensure the operation of transmission scheme [60]. Note that if Alice and Eve transmit signals simultaneously to Bob, Bob will discard the composite signal because of failing to decode it [13, 14].

4.1.2 Channel Model

According to the well-known Jakes model [30], the separation of several to tens of wavelengths for any two entities is required to ensure that the channels between different transmitter-receiver pairs are spatially decorrelated. While the channels from the same transmitter-receiver pair are closely correlated. The receive and transmit antennas are located randomly. Due to the amount of scattering and reflection in the environment, there are lots of multipath in each of the resolvable angular bins and there is no direct path between entities [61]. Moreover, antenna separation of any entity is assumed to be not less than half to one carrier wavelength. Hence, channels between different antenna pairs experience spatially independent fading, and thus can

be modeled with mutually independent Rayleigh fading. Throughout this chapter, indices $m = 1, \ldots, N_t$ and $n = 1, ..., N_r$ are used to denote transmit antennas and receive antennas, respectively.

The distance between the m^{th} transmit antenna at Alice (Eve) to the n^{th} receive antenna at Bob is denoted as $d_A^{[m,n]}$ ($d_E^{[m,n]}$). We assume that $d_A^{[m,n]}/c \ll 1/W$ ($d_E^{[m,n]}/c \ll 1/W$), where c and W are the speed of light and transmission bandwidth, respectively. Assume that antenna array sizes are much smaller than the distance between the transmitter and receiver. We use $h_A^{[m,n]}(k)$ ($h_E^{[m,n]}(k)$) to denote baseband channel gain from the m^{th} transmit antenna at Alice (Eve) to the n^{th} receive antenna at Bob at time k. According to [61], baseband channel gain $h_A^{[m,n]}(k)$ ($h_E^{[m,n]}(k)$) can be expressed by $h_X^{[m,n]}(k) = a_X^{[m,n]} \exp(-\frac{j2\pi f_c d_X^{[m,n]}}{c})$, where $a_X^{[m,n]}$ is path attenuation and f_c is carrier frequency, for $X = \{A, E\}$. We invoke the central limit theorem and approximate baseband channel gain $h_A^{[m,n]}(k)$ ($h_E^{[m,n]}(k)$) as a zero-mean complex circular symmetric Gaussian process [61], i.e., $h_X^{[m,n]}(k) \sim \mathcal{CN}(0, \sigma_{h_X}^{2\ [m,n]})$. For large-scale fading channels, $\sigma_{h_X}^{2\ [m,n]}$ can be modeled by applying [62, Chapter 2], as

$$\sigma_{h_X}^{2\ [m,n]} = K \left(\frac{d_0}{d_X^{[m,n]}} \right)^{\beta} \Upsilon_X, \ X = \{A, E\}, \tag{4.1}$$

where K is a reference path gain value; d_0 is a reference distance for antenna far-field; β is a path loss exponent; and Υ_X is a shadowing factor modeled as a log-normal random variable.

The channels from the same transmitter-receiver pair are assumed to remain constant over a frame but to vary continuously from one frame to the next. We adopt a first-order Gauss-Markov process to characterize temporal channel variations [47, 50], and for instance, $h_A^{[m,n]}(k)$ can be mathematically expressed as

$$h_A^{[m,n]}(k) = \alpha h_A^{[m,n]}(k-1) + \sqrt{1-\alpha^2} u_A^{[m,n]}(k), \tag{4.2}$$

where α is channel correlation coefficient for Alice-Bob, and $u_A^{[m,n]}(k) \sim \mathcal{CN}(0, \sigma^{2\ [m,n]}_{h_A})$ is independent of $h_A^{[m,n]}(k-1)$.

4.1.3 Phase Noise Model

Phase noise is generated at both transmitter and receiver sides during the up-conversion of baseband signal to bandpass and vice versa due to the impairments of local oscillator [23, 63, 64]. For free-running oscillators, phase noise is time-varying and can be modeled as a Wiener process. Phase noise remains constant within a symbol duration but evolves from one symbol to the next. For the frame received by Bob at time k, let $\theta_X^{[m]}(i,k)$ and $\theta_B^{[n]}(i,k)$ be the i^{th} sample of phase noise process at the m^{th} transmit and n^{th} receive antennas, respectively, for $i = 1, ..., L$. Therefore, $\theta_X^{[m]}(i,k)$ and $\theta_B^{[n]}(i,k)$ can be expressed by

$$\theta_X^{[m]}(i,k) = \theta_X^{[m]}(i-1,k) + \Delta_X^{[m]}(i,k), \tag{4.3a}$$

$$\theta_B^{[n]}(i,k) = \theta_B^{[n]}(i-1,k) + \Delta_B^{[n]}(i,k), \tag{4.3b}$$

where $\Delta_X^{[m]}(i,k)$ is phase noise innovation for the m^{th} transmit antenna at X (Alice or Eve) and $\Delta_B^{[n]}(i,k)$ is that of the n^{th} receive antenna at Bob. Both $\Delta_X^{[m]}(i,k)$ and $\Delta_B^{[n]}(i,k)$ can be modeled as zero-mean real Gaussian processes, i.e., $\Delta_X^{[m]}(i,k) \sim \mathcal{N}(0, \sigma^2_{\Delta_X^{[m]}})$ and $\Delta_B^{[n]}(i,k) \sim \mathcal{N}(0, \sigma^2_{\Delta_B^{[n]}})$. Let T_s be the sampling time, we have $\sigma^2_{\Delta_X^{[m]}} = 2\pi c_X^{[m]} T_s$ and $\sigma^2_{\Delta_B^{[n]}} = 2\pi c_B^{[n]} T_s$, where both $c_X^{[m]}$ and $c_B^{[n]}$ are constants and represent the one-sided 3-dB bandwidth of the Lorentzian spectrum of the oscillators at the m^{th} transmit and n^{th} receive antennas, respectively [23, 63, 64].

According to [23, 63, 65, 66], phase noise innovation variances are closely relevant to physical properties of oscillators, i.e., both $\sigma^2_{\Delta_X^{[m]}}$ and $\sigma^2_{\Delta_B^{[n]}}$ are determined by the quality of the oscillators being used at the transmitter and receiver, respectively. Therefore, different transmitter-receiver pairs lead to different innovation variances,

and thus result in different phase noise. This characteristic can be utilized to differentiate between transmitters. $\sigma^2_{\Delta_X^{[m]}}$ and $\sigma^2_{\Delta_B^{[n]}}$ are assumed to be known at Bob. This assumption is reasonable and in line with previous studies on phase noise estimation in MIMO systems [23].

4.1.4 Communication Model

Message frame might not be transmitted continuously but it is necessary to ensure the continuity of authentication process by probing channel at time intervals smaller than channel coherence time [1, 31]. Each frame of length L symbols includes a training sequence of L_t symbols, data symbols, and L_p pilot symbols that are periodically inserted in data symbols for tracking phase noise. Mutually orthogonal training sequences of length L_t ($L_t = jN_t$, $j = 1, 2, \cdots$) is simultaneously transmitted by all transmit antennas to Bob, so that they can be used to jointly estimate channel gains and phase noise. Training sequences and pilot symbols are known at Bob.

Similar to [23, 67, 68], different transmit-receive antenna pairs are assumed to have different channel gains and phase noise. Consider that an unknown transmitter X sends a frame to be authenticated. The signal received by Bob at the n^{th} receive antenna at time k can be written as

$$\begin{aligned} y^{[n]}(i,k) &= \sum_{m=1}^{N_t} h_X^{[m,n]}(k) e^{j\theta_{XB}^{[m,n]}(i,k)} s^{[m]}(i,k) + w^{[n]}(i,k) \\ &= \mathbf{s}_\theta^{[n]}(i,k)\mathbf{h}_X^{[n]}(k) + w^{[n]}(i,k),\ n = 1, ..., N_r, \end{aligned} \tag{4.4}$$

where

- $\mathbf{s}_\theta^{[n]}(i,k) = [s^{[1]}(i,k)e^{j\theta_{XB}^{[1,n]}(i,k)} \cdots s^{[N_t]}(i,k)e^{j\theta_{XB}^{[N_t,n]}(i,k)}]$ with $s^{[m]}(i,k)$ denoting the i^{th} sample symbol of the $m^{\mathbf{th}}$ transmit antenna at time k, and consists of both pilots and data symbols; the average power is $p = \mathbb{E}\{s^{[m]}(i,k)s^{H\ [m]}(i,k)\}$;

- $\theta_{XB}^{[m,n]}(i,k) = \theta_{X}^{[m]}(i,k) + \theta_{B}^{[n]}(i,k)$ represents the overall phase noise from the oscillators corresponding to the m^{th} transmit and n^{th} receive antennas;
- $\mathbf{h}_{X}^{[n]}(k) = [h_{X}^{[1,n]}(k) \cdots h_{X}^{[N_t,n]}(k)]^T$;
- $w^{[n]}(i,k)$ is a sequence of independent and identically distributed zero-mean complex AWGN with variance σ_w^2 at the n^{th} receive antenna, i.e., $w^{[n]}(i,k) \sim \mathcal{CN}(0,\sigma_w^2)$. σ_w^2 is readily available at the receiver via the method in [56].

4.2 Proposed Physical Layer Authentication Scheme

The basic principles for the proposed scheme are that wireless channels are location-specific, and phase noise is transmitter-specific [14, 51, 69]. Therefore, channel gains and phase noise can be jointly utilized to differentiate between Alice and Eve. The proposed scheme consists of three processes: channel and phase noise estimation, channel and phase noise quantization, and decision.

4.2.1 Channel and Phase Noise Estimation

4.2.1.1 Channel Estimation

Similar to [70], we can use maximum-likelihood estimator (MLE) to estimate channel gains. Let $\boldsymbol{\theta}_{XB}^{[n]}(i,k) = [\theta_{XB}^{[1,n]}(i,k) \cdots \theta_{XB}^{[N_t,n]}(i,k)]^T$, for $i = 1, ..., L_t$. We use $\hat{\mathbf{h}}_{X}^{[n]}(k)$ and $\hat{\theta}_{XB}^{[n]}(i,k)$ to denote the estimations of $\mathbf{h}_{X}^{[n]}(k)$ and $\boldsymbol{\theta}_{XB}^{[n]}(i,k)$, respectively. Using the mutually orthogonal training sequences of length L_t estimates channel gains. Then, at the i^{th} iteration, given the phase noise estimation $\boldsymbol{\theta}_{XB}^{[n]}(i,k)$, $\hat{\mathbf{h}}_{X}^{[n]}(k)$ can be obtained by using MLE as

$$\hat{\mathbf{h}}_{X}^{[n]}(k) = ((\mathbf{s}_{\theta}^{[n]})^H(i,k)\mathbf{s}_{\theta}^{[n]}(i,k))^{-1}(\mathbf{s}_{\theta}^{[n]})^H(i,k)y^{[n]}(i,k),\ i = 1, ..., L_t. \tag{4.5}$$

To acquire separate estimators, we substitute $\hat{\mathbf{h}}_{X}^{[n]}(k)$ into least-squares objective

to obtain

$$\hat{\boldsymbol{\theta}}_{XB}^{[n]}(i,k) = \arg \max_{\theta_{XB}^{[n]}(i,k)} (y^{[n]})^H(i,k)\mathbf{s}_\theta^{[n]}(i,k)((\mathbf{s}_\theta^{[n]})^H(i,k)\mathbf{s}_\theta^{[n]}(i,k))^{-1}(\mathbf{s}_\theta^{[n]})^H(i,k)y^{[n]}(i,k), \tag{4.6}$$

where $\hat{\boldsymbol{\theta}}_{XB}^{[n]}(i,k) = [\hat{\theta}_{XB}^{[1,n]}(i,k)\cdots\hat{\theta}_{XB}^{[N_t,n]}(i,k)]^T$ and $\hat{\theta}_{XB}^{[m,n]}(i,k)$ is the estimation of $\theta_{XB}^{[m,n]}(i,k)$. Based on (4.6), we can acquire separate phase noise, and then substitute phase noise parameters into (4.5) to get $\hat{\mathbf{h}}_X^{[n]}(k)$. Note that (4.6) is a multi-dimensional minimization problem with a high computational complexity. Some other heuristic methods with a low computational complexity such as majorization-minimization methods based on dimensionality reduction and regularization [71], have been proposed to jointly estimate channel gains and phase noise. In addition, following a similar manner in [72] to jointly estimate channel and carrier frequency offset, we can also obtain the estimations of channel gains and phase noise.

Channel estimation from the m^{th} transmit antenna of X to the n^{th} receive antenna of Bob at time k, denoted by $\hat{h}_X^{[m,n]}(k)$, can be obtained by Bob. In particular, $\hat{h}_X^{[m,n]}(k)$ can be modeled as a sum of $h_X^{[m,n]}(k)$ and estimation error $w_X^{[m,n]}(k)$. Based on [56, Eq.(4.7)], for training sequences of length L_t, we have $w_h(i,k) \sim \mathcal{CN}(0, \sigma_w^2/(pL_t))$. Then, they are related by

$$\hat{h}_X^{[m,n]}(k) = h_X^{[m,n]}(k) + w_h(i,k). \tag{4.7}$$

4.2.1.2 Phase Noise Estimation

After obtaining the estimation of channel, we can track N phase noise parameters over a frame by using a soft-input extended Kalman filter (EKF) proposed in [68]. We use $\boldsymbol{\phi}(i,k) = [{\boldsymbol{\theta}_{XB}^{[1]}}^T(i,k)\cdots{\boldsymbol{\theta}_{XB}^{[N_r]}}^T(i,k)]^T$ and $\mathbf{M}(i,k)$ to denote the unknown state vector and error covariance matrix, respectively. Before starting the EKF recursion,

$\boldsymbol{\phi}(i,k)$ and $\mathbf{M}(i,k)$ should be initialized as $\boldsymbol{\phi}(L_t|L_t,k)$ and $\mathbf{M}(L_t|L_t,k)$ using (4.5) and (4.6), respectively, where L_t corresponds to the last training symbol (please kindly refer to [68] for details).

Similarly, phase noise estimate is formulated as its actual value and a real Gaussian noise (i.e., estimation error). For a frame containing L_p pilot symbols, the estimation of the i^{th} sample for the overall phase noise, denoted by $\hat{\theta}_{XB}^{[m,n]}(i,k)$, can be given by

$$\hat{\theta}_{XB}^{[m,n]}(i,k) = \theta_{XB}^{[m,n]}(i,k) + w_{\theta}^{[m,n]}(i,k),\ i = 1, ..., L_p, \tag{4.8}$$

where $w_{\theta}^{[m,n]}(i,k) \sim N(0, \sigma_{w_\theta}^{2\ [m,n]})$ is the estimation error for phase noise.

Now, we need to calculate $\sigma_{w_\theta}^{2\ [m,n]}$. Let $\Delta_{XB}^{[m,n]}(i,k)$ denote the total phase noise innovation corresponding to $\theta_{XB}^{[m,n]}(i,k)$, and then we have $\Delta_{XB}^{[m,n]}(i,k) = \Delta_{X}^{[m]}(i,k) + \Delta_{B}^{[n]}(i,k)$ and $\Delta_{XB}^{[m,n]}(i,k) \sim \mathcal{N}(0, \sigma^2_{\Delta_{XB}^{[m,n]}})$, where $\sigma^2_{\Delta_{XB}^{[m,n]}} = \sigma^2_{\Delta_{X}^{[m]}} + \sigma^2_{\Delta_{B}^{[n]}}$. By applying [73, Eq.(104)], the limit posterior Bayesian Cramér-Rao bound (BCRB) on the instantaneous phase noise can be expressed by

$$\text{BCRB}(\hat{\theta}) = \frac{\sigma_{\Delta_{XB}}^{2\ [m,n]}}{\varepsilon^{[m,n]}}(-\varepsilon^{[m,n]} + \sqrt{\varepsilon^{2\ [m,n]} + 4\varepsilon^{[m,n]}}), \tag{4.9}$$

where $\varepsilon^{[m,n]} = \zeta^{[m,n]} + \sqrt{\zeta^{2\ [m,n]} + 8\zeta^{[m,n]}}$ and $\zeta^{[m,n]} = \frac{p\sigma_{h_X}^{2\ [m,n]}}{\sigma_w^2}\sigma_{\Delta_{XB}}^{2\ [m,n]}$.

From [73], when increasing the number of recursion for phase noise, the EKF mean-square error is gradually close to the BCRB. Thus, to comprehensively investigate the optimum performance of the proposed scheme, we let $\sigma_{w_\theta}^{2\ [m,n]} \approx \text{BCRB}(\hat{\theta})$.

4.2.2 Channel and Phase Noise Quantization

4.2.2.1 Channel Quantization

To quantify temporal channel variations, we use channel gain quantizer [14], which compares the square of the difference between the current and previous channel gain

estimations at adjacent time with a channel gain threshold δ_h. We use Q_h and $O_h^{[m,n]}$ to denote channel gain quantizer and the $[m,n]^{\text{th}}$ output, respectively. Then, channel gain quantization can be formulated as

$$\begin{aligned} O_h^{[m,n]} &\triangleq Q_h[|\hat{h}_X^{[m,n]}(k) - \hat{h}_A^{[m,n]}(k-1)|^2] \\ &= \begin{cases} 1, \ |\hat{h}_X^{[m,n]}(k) - \hat{h}_A^{[m,n]}(k-1)|^2 > \delta_h, \\ 0, \ \text{otherwise.} \end{cases} \end{aligned} \tag{4.10}$$

4.2.2.2 Phase Noise Quantization

Based on (4.3a), (4.3b), and (4.8), we have

$$\hat{\theta}_{XB}^{[m,n]}(i,k) - \hat{\theta}_{XB}^{[m,n]}(i-1,k) = \hat{\Delta}_{XB}^{[m,n]}(i,k), \tag{4.11}$$

where $\hat{\Delta}_{XB}^{[m,n]}(i,k)$ are regarded as the estimation of $\Delta_{XB}^{[m,n]}(i,k)$, which can be given by

$$\hat{\Delta}_{XB}^{[m,n]}(i,k) = \Delta_{XB}^{[m,n]}(i,k) + w_\theta(i,k) - w_\theta(i-1,k). \tag{4.12}$$

From (4.12), we can see that $\hat{\Delta}_{XB}^{[m,n]}(i,k)$ is a zero-mean real Gaussian random variable and its variance can be given by

$$\sigma^2_{\hat{\Delta}_{XB}^{[m,n]}} = \sigma^2_{\Delta_X^{[m]}} + \sigma^2_{\Delta_B^{[n]}} + 2\sigma^2_{w_\theta}, \ X = \{A, E\}. \tag{4.13}$$

To quantify phase noise variations, we just need to quantize phase noise innovation estimation variations. Phase noise quantizer is developed by comparing the sum of L_p squares of difference between phase noise innovation estimations of the current and previous frames with a phase noise threshold δ_θ. Let Q_θ and $O_\theta^{[m,n]}$ denote phase noise quantizer and the $[m,n]^{\text{th}}$ output, respectively. Phase noise quantization can

be formulated as

$$
\begin{aligned}
O_{\theta}^{[m,n]} &\triangleq Q_{\theta}\left[\sum_{i=1}^{L_p}\left(\hat{\Delta}_{XB}^{[m,n]}(i,k)-\hat{\Delta}_{AB}^{[m,n]}(i,k-1)\right)^2\right] \\
&= \begin{cases} 1, \sum\limits_{i=1}^{L_p}\left(\hat{\Delta}_{XB}^{[m,n]}(i,k)-\hat{\Delta}_{AB}^{[m,n]}(i,k-1)\right)^2 > \delta_{\theta}, \\ 0, \text{ otherwise}, \end{cases}
\end{aligned} \tag{4.14}
$$

where $\hat{\Delta}_{AB}^{[m,n]}(i,k-1)$ is stored at Bob. Let D_h denote the output sum of channel gain quantizer and D_θ for phase noise quantizer, i.e., $D_\imath \triangleq \sum_{m=1}^{N_t}\sum_{n=1}^{N_r} O_{\imath}^{[m,n]}$, for $\imath = \{h, \theta\}$. If let $N = N_t N_r$, then both D_h and D_θ are non-negative integers between 0 and N.

4.2.3 Decision

Based on the above quantization results in terms of temporal variations for channel gains and phase noise, a decision criterion can be modeled as a composite hypothesis test to discriminate the identity of the current transmitter. The decision criterion can be formulated as

$$
\begin{aligned}
H_0: \ D = D_h + D_\theta \leq Z, \\
H_1: \ D = D_h + D_\theta > Z,
\end{aligned} \tag{4.15}
$$

where Z is a decision threshold, which is a non-negative integer between 0 and $2N$. The decision criterion enables Bob to decide whether the transmitter is still Alice. In particular, on H_0, the current transmitter is still Alice; on H_1, the current transmitter is Eve.

4.2.4 Properties of the Proposed Authenticate Scheme

4.2.4.1 Covertness

A transmitted signal for the current frame to be authenticated is considered to be anomalous if its power spectral density (PSD) is different from that of normal signal. For the proposed scheme, on one hand, the frame does not contain any special signals such as the proof of authentication (namely tag) [47]. Thus, Eve cannot discover authentication process by analyzing the PSD of the signal transmitted by Alice, since this PSD is normal. On the other hand, signal propagation in wireless channels is generally affected by multiplicative fading of channel and additive background noise. Channel fading is not be dependable for asserting the transmitted signal as anomalous. Nevertheless, one could judge the transmitted signal as anomalous, due to the abnormal distribution of background noise. Since our proposed scheme does not affect background noise, the distribution information of background noise is also normal. Therefore, the proposed scheme jointly using wireless channel and hardware features for authentication is covert to adversaries due to the absence of abnormal PSD and background noise in the whole authentication process.

4.2.4.2 Robustness

The authentication scheme is regarded to be robust if it can resist channel fading, the random locations of entities, and background noise effects without sacrificing performance when many frames from the same transmitter are authenticated together (instead of each frame separately). In particular, the proposed scheme can be resistant to channel fading by using some physical layer techniques, e.g., antenna diversity. This scheme is reasonably robust to the large-scale fading and random locations of entities effects (detailed in Section 4.4.4). When an unknown transmitter (Alice or Eve) sends a frame cluster consisting of many frames to be authenticated to Bob within the

channel coherence time, Bob can authenticate them without sacrificing performance in low mobility scenarios, by adjusting three thresholds (for channel amplitude, phase noise, and decision, respectively). In particular, the number of frames received allows Bob to know phase noise increment variances based on (4.11), and the arrival time of the last frame enables Bob to calculate the temporal correlation channel coefficient [47]. As a result, Bob can independently adjust the three thresholds to authenticate a frame cluster from the same transmitter.

4.2.4.3 Security

When Eve is close to Alice, she still fails to impersonate Alice for injecting aggressive signals into the network. This is because phase noise solely relies on oscillator properties of transmitter-receiver pair. When Eve is far away from Alice, she cannot succeed in replaying the received signals. The reason is that each transmitter distorts signals in its own way. In summary, the proposed authentication scheme is an effective solution to resist against impersonation and replay attacks.

4.2.5 Analysis of Communication Overhead and Computational Complexity

Notice that the proposed scheme is based on both channel gains and Wiener phase noise. Compared with the existing methods that only rely on channel gains (e.g., the one in [12]), our proposed scheme has higher communication and computational overheads. This is because tracking phase noise needs more pilot symbols and thus requires a larger bandwidth and latency, and also the tracking phase noise based on the EKF algorithm involves extra addition and multiplication operations and thus causes an increase in computational complexity. If we use C^{CM} and C^{ADD} to denote the number of complex multiplications and the number of additions in the EKF algorithm, then the computational complexity C of the EKF algorithm required to

update phase noise parameters is determined as $C = C^{CM} + C^{ADD}$ [23]. Here, C^{CM} and C^{ADD} are given by

$$C^{CM} = N_t N_r(3N_t N_r^2 + 4) + N_r^2[N_r(2(1 + N_t^2) + 1) + N_t], \tag{4.16}$$

$$C^{ADD} = (N_t N_r)^2(N_t N_r + 1) + N_t N_r^3(3N_t + 2) - N_t N_r(3N_r + 1) + N_r. \tag{4.17}$$

Since smart devices in heterogeneous coexistence MIMO systems (e.g., 5G networks) have powerful communication and computational capabilities, such additional communication and computational overheads in our proposed scheme are in general acceptable for engineering practices.

4.3 Modeling of FA and MD Probabilities

In this section, we first derive some basic results regarding the probabilities that the output of each quantizer under two hypotheses is 1, and then use the results to analytically model P_F and P_M. For simplicity, we use $P_\imath^{H_\jmath}$ to denote the probability that $Q_\imath$ outputs 1 on $H_\jmath$, for $\imath = \{h, \theta\}$ and $\jmath = \{0, 1\}$.

It is notable that there always exist quantization errors due to the presence of thermal noise and/or interference. In order to exactly model false alarm and missed detection probabilities, it is necessary to take the inevitable quantization errors resulting from the two quantizers into account. When an error occurs, regarding of channel and/or phase noise quantization, a 1 is wrongly quantified as a 0, and vice versa. Suppose such errors are equally probable. For concreteness, the probability that a temporal channel variation (resp. temporal phase noise variation) is quantified with error is denoted by P_{e_h} (resp. P_{e_θ}); hence, the temporal channel variation (resp. phase noise variation) is quantified without error is denoted by 1-P_{e_h} (resp. 1-P_{e_θ}).

4.3.1 False Alarm Probability

The following lemmas are dedicated to present $P_h^{H_0}$ and $P_\theta^{H_0}$.

Lemma 6 *For a given δ_h, $P_h^{H_0}$ can be evaluated as*

$$P_h^{H_0} = P_{e_h} + (1 - 2P_{e_h}) \exp\left(-\frac{\delta_h}{2(1-\alpha)\sigma_{h_A}^{2\ [m,n]} + 2\sigma_w^2/(pL_t)} \right). \tag{4.18}$$

Proof 7 *On H_0, the current transmitter is still Alice, that is, $X = A$. Let $\Lambda h_{A,A}^{[m,n]} = \hat{h}_A^{[m,n]}(k) - \hat{h}_A^{[m,n]}(k-1)$, so*

$$\begin{aligned} \Lambda h_{A,A}^{[m,n]} &= \hat{h}_A^{[m,n]}(k) - \hat{h}_A^{[m,n]}(k-1) \\ &= (\alpha - 1) h_A^{[m,n]}(k-1) + \sqrt{1-\alpha^2} u_A^{[m,n]}(k) + w_h(k) - w_h(k-1). \end{aligned} \tag{4.19}$$

From (4.19), one can see that $\Lambda h_{A,A}^{[m,n]}$ is a zero-mean circularly symmetric complex Gaussian random variable with variance $2(1-\alpha)\sigma_{h_A}^{2\ [m,n]} + 2\sigma_w^2/(pL_t)$, since all $h_A^{[m,n]}(k-1)$, $u_A^{[m,n]}(k)$, $w_h(k-1)$, and $w_h(k)$ are zero-mean circularly symmetric complex Gaussian random variables and statistically independent of each other. Hence, $|\Lambda h_{A,A}^{[m,n]}|^2$ follows exponential distribution and its cumulative distribution function (CDF) can be written as

$$F_{|\Lambda h_{A,A}^{[m,n]}|^2}(x) = 1 - \exp\left(-\frac{x}{2(1-\alpha)\sigma_{h_A}^{2\ [m,n]} + 2\sigma_w^2/(pL_t)} \right). \tag{4.20}$$

According to (4.10), $P_h^{H_0}$ can be evaluated as

$$\begin{aligned} P_h^{H_0} &\triangleq \mathbf{Pr}(O_h^{[m,n]} = 1 \mid H_0) \\ &= \mathbf{Pr}(O_h^{[m,n]} = 1, |\Lambda h_{A,A}^{[m,n]}|^2 > \delta_h \mid H_0) + \mathbf{Pr}(O_h^{[m,n]} = 1, |\Lambda h_{A,A}^{[m,n]}|^2 \le \delta_h \mid H_0) \\ &= (1 - \mathbf{Pr}(|\Lambda h_{A,A}^{[m,n]}|^2 \le \delta_h))(1 - P_{e_h}) + \mathbf{Pr}(|\Lambda h_{A,A}^{[m,n]}|^2 \le \delta_h) P_{e_h}. \end{aligned} \tag{4.21}$$

Substituting (4.20) *into* (4.21)*, we can obtain* (4.18)*.*

Lemma 7 *For a given* δ_θ*,* $P_\theta^{H_0}$ *can be evaluated as*

$$P_\theta^{H_0} = P_{e_\theta} + (1-2P_{e_\theta})\mathbf{\Gamma}_{\chi^2_{L_p}}\left(\frac{\delta_\theta}{2(\sigma^2_{\Delta_A^{[m]}}+\sigma^2_{\Delta_B^{[n]}}+2\sigma^2_{w_\theta})}\right), \tag{4.22}$$

where $\mathbf{\Gamma}_{\chi^2_{L_p}}(\cdot)$ *denotes the right-tail probability function for a* $\chi^2_{L_p}$ *random variable with* L_p *degrees of freedom, which can be expressed as (please refer to [56] details)*

$$\mathbf{\Gamma}_{\chi^2_{L_p}}(x) = \begin{cases} 2\Psi(\sqrt{x}),\ L_p = 1, & \text{(4.23a)} \\ \exp(-\frac{1}{2})\sum_{j=0}^{\frac{L_p}{2}-1}\frac{\frac{x}{2}}{j!},\ L_p = 2,4,\cdots & \text{(4.23b)} \\ 2\Psi(\sqrt{x}) + \frac{\exp(-\frac{1}{2}x)}{\sqrt{\pi}}\sum_{j=1}^{\frac{L_p-1}{2}}\frac{(k-1)!(2x)^{j-\frac{1}{2}}}{(2j-1)!},\ L_p = 3,5,\cdots & \text{(4.23c)} \end{cases}$$

where $\Psi(\cdot)$ *denotes the right-tail probability function of Gaussian distribution in engineering texts.*

Proof 8 *Let* $\Lambda\theta_{A,A}^{[m,n]} = \sum_{i=1}^{L_p}\left(\hat{\Delta}_{AB}^{[m,n]}(i,k) - \hat{\Delta}_{AB}^{[m,n]}(i,k-1)\right)^2$*. Since* $(\hat{\Delta}_{AB}^{[m,n]}(i,k) - \hat{\Delta}_{AB}^{[m,n]}(i,k-1))$ *is real zero-mean Gaussian random variable with variance* $2\sigma^2_{\hat{\Delta}_{AB}^{[m,n]}}$*,* $\Lambda\theta_{A,A}^{[m,n]}/2\sigma^2_{\hat{\Delta}_{AB}^{[m,n]}}$ *follows central chi-square distribution with* L_p *degrees of freedom, i.e.,* $\Lambda\theta_{A,A}^{[m,n]} \sim \chi^2_{L_p}$*.*

For a given δ_θ*,* $P_\theta^{H_0}$ *can be evaluated as*

$$\begin{aligned} P_\theta^{H_0} &\triangleq \mathbf{Pr}(O_\theta^{[m,n]} = 1, \Lambda\theta_{A,A}^{[m,n]} > \delta_\theta \mid H_0) + \mathbf{Pr}(O_\theta^{[m,n]} = 1, \Lambda\theta_{A,A}^{[m,n]} \le \delta_\theta \mid H_0) \\ &= (1 - \mathbf{Pr}(\Lambda\theta_{A,A}^{[m,n]} \le \delta_\theta))(1-P_{e_\theta}) + \mathbf{Pr}(\Lambda\theta_{A,A}^{[m,n]} \le \delta_\theta)P_{e_\theta}. \end{aligned} \tag{4.24}$$

Combining (4.23) *and* (4.24)*, we can obtain* (4.22)*.*

We now present the following lemma regarding the probability that D defined in (5.19) equals to a fixed integer value. It is of great importance for the modeling of P_F and P_M.

$$\mathbf{Pr}(D=z)=\begin{cases}\sum\limits_{z_1=0}^{z}\binom{N}{z_1}\binom{N}{z-z_1}(P_h^{H_0})^{z_1}(1-P_h^{H_0})^{N-z_1}(P_\theta^{H_0})^{z-z_1}(1-P_\theta^{H_0})^{N-z+z_1}, \\ z\in[0,N], \qquad (4.25a)\\ \sum\limits_{z_1=z-N}^{N}\binom{N}{z-z_1}\binom{N}{z_1}(P_h^{H_0})^{(z-z_1)}(1-P_h^{H_0})^{N-z+z_1}(P_\theta^{H_0})^{z_1}(1-P_\theta^{H_0})^{N-z_1}, \\ z\in(N,2N]. \qquad (4.25b)\end{cases}$$

Lemma 8 *On H_0, the probability that $D = D_h + D_\theta$ equals a fixed integer value $z \in [0, 2N]$ can be given in* (4.25)*, where D_h, $D_\theta \in [0, N]$.*

Proof 9 *The proof of Lemma 8 is straightforward, and please refer to [14, Appendix].*

Based on lemmas 7 and 8, we can establish the following theorem on P_F.

Theorem IV.1 *P_F for the proposed physical layer authentication scheme jointly utilizing the characteristics of location-specific channel gains and transmitter-specific phase noise in MIMO systems, can be determined as* (4.27).

Proof 10 *Based on* (4.15)*, P_F can be written as*

$$P_F=\mathbf{Pr}(D>Z|H_0)=\mathbf{Pr}(D=Z+1,Z+2,\cdots,2N|H_0)=\sum_{z=Z+1}^{2N}\mathbf{Pr}(D=z|H_0)$$
$$=\sum_{z=Z+1}^{N}\mathbf{Pr}(D=z|H_0)+\sum_{z=N+1}^{2N}\mathbf{Pr}(D=z|H_0). \qquad (4.26)$$

According to (4.25)*, P_F can be modeled by considering two cases: when $Z \in [0, N]$, substituting* (4.25a) *into* (4.26) *yields* (4.27a)*; when $Z \in (N, 2N]$, substituting* (4.25b) *into* (4.26) *yields* (4.27b).

If the proposed scheme utilizes channel gains or phase noise separately to discriminate transmitters, we can get the following corollary regarding P_F.

$$P_F = \begin{cases} \sum\limits_{z=Z+1}^{N} \sum\limits_{z_1=0}^{z} \binom{N}{z_1} (P_h^{H_0})^{z_1} (1-P_h^{H_0})^{N-z_1} \binom{N+1}{z-z_1} (P_\theta^{H_0})^{z-z_1} (1-P_\theta^{H_0})^{N-z+z_1} \\ + \sum\limits_{z=N}^{2N} \sum\limits_{z_1=z-N}^{N} \binom{N}{z-z_1} \binom{N}{z_1} (P_h^{H_0})^{z-z_1} (1-P_h^{H_0})^{N-z+z_1} (P_\theta^{H_0})^{z_1} (1-P_\theta^{H_0})^{N-z_1}, \\ Z \in [0, N], \quad (4.27a) \\ \sum\limits_{z=Z+1}^{2N} \sum\limits_{z_1=z-N}^{N} \binom{N}{z-z_1} \binom{N}{z_1} (P_h^{H_0})^{z-z_1} (1-P_h^{H_0})^{N-z+z_1} (P_\theta^{H_0})^{z_1} (1-P_\theta^{H_0})^{N-z_1}, \\ Z \in (N, 2N]. \quad (4.27b) \end{cases}$$

Corollary 1 *P_F for the proposed scheme separately utilizing channel gains or phase noise, can be evaluated as*

$$P_F = \sum_{z=Z+1}^{N} \binom{N}{z} (P_\imath^{H_0})^z (1-P_\imath^{H_0})^{N-z}, \ \imath = \{h, \theta\}. \quad (4.28)$$

Proof 11 *When only using channel gains, we have $D = D_h \in [0, N]$ on H_0. Based on* (4.15) *and* (4.18)*, P_F is determined as*

$$P_F = \mathbf{Pr}(D > Z | H_0) = \sum_{z=Z+1}^{N} \binom{N}{z} (P_h^{H_0})^z (1-P_h^{H_0})^{N-z}. \quad (4.29)$$

When only using phase noise, we have $D = D_h \in [0, N]$ on H_0. Based on (4.15) *and* (4.22)*, P_F is determined as*

$$P_F = \mathbf{Pr}(D > Z | H_0) = \sum_{z=Z+1}^{N} \binom{N}{z} (P_\theta^{H_0})^z (1-P_\theta^{H_0})^{N-z}. \quad (4.30)$$

Hence, (4.29) *and* (4.30) *can be summarized as* (4.28)*.*

4.3.2 Missed Detection Probability

Similarly, to model P_M, we need to explore exact expressions for $P_\theta^{H_1}$ and $P_h^{H_1}$.

$$P_M = \begin{cases} \sum_{z=0}^{Z} \sum_{z_1=0}^{z} \binom{N}{z_1}\binom{N}{z-z_1}(P_h^{H_1})^{z_1}(1-P_h^{H_1})^{N-z_1}(P_\theta^{H_1})^{z-z_1}(1-P_\theta^{H_1})^{N-z+z_1}, Z \in [0,N], & (4.33a) \\ \sum_{z=0}^{N} \sum_{z_1=0}^{z} \binom{N}{z_1}\binom{N}{z-z_1}(P_h^{H_1})^{z_1}(1-P_h^{H_1})^{N-z_1}(P_\theta^{H_1})^{z-z_1}(1-P_\theta^{H_1})^{N-z+z_1} \\ + \sum_{z=N+1}^{Z-1} \sum_{z_1=z-N}^{N} \binom{N}{z-z_1}\binom{N}{z_1}(P_h^{H_1})^{z-z_1}(1-P_h^{H_1})^{N-z+z_1}(P_\theta^{H_1})^{z_1}(1-P_\theta^{H_1})^{N-z_1}, \\ Z \in (N,2N]. & (4.33b) \end{cases}$$

Lemma 9 *For a given δ_h, $P_h^{H_1}$ can be evaluated as*

$$P_h^{H_1} = P_{e_h} + (1-2P_{e_h}) \exp\left(-\frac{\delta_h}{\sigma_{h_A}^{2\ [m,n]} + \sigma_{h_E}^{2\ [m,n]} + 2\sigma_w^2/(pL_t)} \right), \tag{4.31}$$

Proof 12 *Following a similar proof as that of* (4.22) *yields* (4.31)*, we omit it here.*

Lemma 10 *For a given phase noise threshold δ_θ, $P_\theta^{H_1}$ can be evaluated as*

$$P_\theta^{H_1} = P_{e_\theta} + (1-2P_{e_\theta}) \mathbf{\Gamma}_{\chi^2_{L_p}} \left(\frac{\delta_\theta}{\sigma^2_{\Delta_A^{[m]}} + \sigma^2_{\Delta_E^{[m]}} + 2\sigma^2_{\Delta_B^{[n]}} + 4\sigma^2_{w_\theta}} \right). \tag{4.32}$$

Proof 13 *Following a similar proof as that of Lemma 7 yields* (4.32)*, so we omit it here.*

Based on Lemma 8, 9, and 10, P_M can given by Theorem IV.2.

Theorem IV.2 *P_M for the proposed scheme can be evaluated as* (4.33)*.*

Proof 14 *According to* (4.15)*, P_M can be written as*

$$P_M = \mathbf{Pr}(D \le Z|H_1) = \sum_{z=0}^{N} \mathbf{Pr}(D=z|H_1) + \sum_{z=N+1}^{Z} \mathbf{Pr}(D=z|H_1). \tag{4.34}$$

Similarly, using $P_h^{H_1}$ *and* $P_\theta^{H_1}$ *to replace* $P_h^{H_0}$ *and* $P_\theta^{H_0}$ *in* (4.25)*, respectively, we can obtain* P_M *for two cases.*

We also give the following corollary regarding P_M.

Corollary 2 *P_M for the proposed scheme utilizing channel gains or phase noise, can be evaluated as*

$$P_M = 1 - \sum_{z=Z+1}^{N} \binom{N}{z} (P_\imath^{H_1})^z (1 - P_\imath^{H_1})^{N-z}, \ \imath = \{h, \theta\}. \tag{4.35}$$

Proof 15 *When only using channel gains, we have* $D = D_h \in [0, N]$ *on* H_1*. Based on* (4.15) *and* (4.31)*,* P_M *is determined as*

$$P_M = \mathbf{Pr}(D \leq Z|H_1) = 1 - \mathbf{Pr}(D > Z|H_1) = 1 - \sum_{z=Z+1}^{N} \binom{N}{z} (P_h^{H_1})^z (1 - P_h^{H_1})^{N-z}. \tag{4.36}$$

When only using phase noise, we have $D = D_\theta \in [0, N]$ *on* H_1*. Based on* (4.15) *and* (4.32)*,* P_M *is determined as*

$$P_M = \mathbf{Pr}(D \leq Z|H_1) = 1 - \mathbf{Pr}(D > Z|H_1) = 1 - \sum_{z=Z+1}^{N} \binom{N}{z} (P_\theta^{H_1})^z (1 - P_\theta^{H_1})^{N-z}. \tag{4.37}$$

Similarly, (4.36) *and* (4.37) *can be summarized as* (4.35)*.*

4.4 Simulation Results

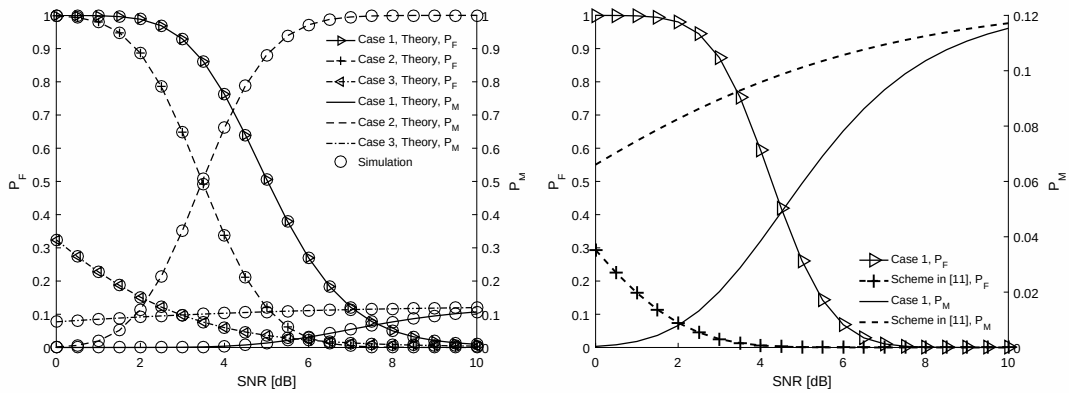
4.4.1 System Parameters and Simulation Settings

To comprehensively investigate the impact of channel fading on authentication performance, we consider small-scale fading due to mobility (e.g., scatters moving) as

well as large-scale fading due to path loss and shadow fading as a function of distance [62]. In the former, the effect of Doppler shift associated with moving entities or scatters is considered. Each α corresponds to a normalized Doppler frequency value representing a channel status. In the latter, the effect of path loss as a function of distance is investigated. Since the distance between transmitter and receiver is much larger than antenna separation, we get the approximation $d_X^{[1,1]} \approx \cdots \approx d_X^{[N_t,N_r]} = d_X$. From (4.1), we have $\sigma_{h_X}^{2\ [m,n]} \approx \cdots \approx \sigma_{h_X}^{2\ [N_t,N_r]} = \sigma_{h_X}^2$ and let $\kappa_h = \frac{\sigma_{h_E}^2}{\sigma_{h_A}^2}$ denote the ratio of the locally average channel gains for Alice and Eve. To investigate the impact of the location of Eve on the performance, we fix the location of Alice by setting $\sigma_{h_A}^2 = 1$ and adjust $\sigma_{h_E}^2$ to achieve a specified κ_h. To analyze the spatially-averaged performance under the large-scale fading, we assume that both Alice and Eve are randomly deployed at arbitrary positions in a circular area centered on Bob for outdoor, and there is no shadow fading in that area. By using (4.1), κ_h by dB value without shadow fading can be written as $\kappa_h = 10\beta \log\left(\frac{d_A}{d_E}\right)$. Hence, κ_h completely depends on the ratio of d_A and d_E. According to [12], the probability density function of κ_h by dB value is a double-sided exponential given by

$$f_{\kappa_h}(x) = \frac{\ln(10)}{10\beta} 10^{\frac{-|x|}{5\beta}}. \tag{4.38}$$

Wiener phase noise is generated for each entity with the following assumption $\sigma_{\Delta_A^{[1]}}^2 = \cdots = \sigma_{\Delta_A^{[N_t]}}^2 = \sigma_{\Delta_A}^2$, $\sigma_{\Delta_B^{[1]}}^2 = \cdots = \sigma_{\Delta_B^{[N_r]}}^2 = \sigma_{\Delta_B}^2$, and $\sigma_{\Delta_E^{[1]}}^2 = \cdots = \sigma_{\Delta_E^{[N_t]}}^2 = \sigma_{\Delta_E}^2$. Moreover, we set $\sigma_{\Delta_A}^2 = \sigma_{\Delta_B}^2 = 10^{-4}$ rad^2, and then a specified κ_Δ can be obtained by adjusting parameter $\sigma_{\Delta_E}^2 \in \{10^{-3}, 10^{-4}, 10^{-5}\}$ rad^2 [64]. We consider two frames of 300 symbols. The first frame originated from Alice is authenticated by Bob, and 500 μs later, the second one is received. If the second one still originates from Alice, we generate time-varying MIMO channel matrix spanning two frames with specified parameters in each run. If it originates from Eve, we generate mutually



(a) Theoretical and simulated performance. (b) Performance comparison

Figure 4.2: (P_F, P_M) vs. SNR with the settings ($Z = 3$, $\kappa_h = \kappa_\Delta = 0$ dB, $\delta_h = 0.5$, $\delta_\theta = 0.0815$, $L_t = 3$, $L_p = 6$, $\alpha = \rho = 0.9$, $P_{e_h} = P_{e_\theta} = 0$).

independent channel matrices instead. Let SNR=$\frac{p\sigma_{h_A}^2}{\sigma_w^2}$ be signal-to-noise ratio.We set $N_t = 3$ and $N_r = 2$. Due to the assumption of spatially independent fading channels, we can use the simple mutually orthogonal training sequences, e.g., $[\sqrt{p}\ 0\ 0]^T$, $[0\ \sqrt{p}\ 0]^T$, and $[0\ 0\ \sqrt{p}]^T$ [23]. To evaluate the average authentication performance, we carry out 10^5 independent Monte-Carlo trials.

4.4.2 Model Validation and Authentication Performance Comparison

Extensive simulations have been conducted to verify theoretical results in terms of P_F and P_M. For the fixed setting of ($Z = 3$, $\kappa_h = \kappa_\Delta = 0$ dB, $\delta_h = 0.5$, $\delta_\theta = 0.0815$, $L_t = 3$, $L_p = 6$, $\alpha = 0.9$, $P_{e_h} = P_{e_\theta} = 0$), we provide plots of the theoretical and simulated performance for the proposed scheme in Fig. 4.2(a), and we consider three cases (case 1: jointly utilizing channel gains and phase noise; case 2: only utilizing phase noise; and case 3: only utilizing channel gains). As shown in Fig. 4.2(a), simulation results agree well with the theoretical ones for three cases. This indicates that our theoretical models is highly accurate in depicting P_F and P_M for the proposed authentication scheme in MIMO systems.

As observed from Fig. 4.2(a) that when SNR increases, P_F decreases and P_M rises. In the whole SNR region, the proposed scheme under case 1 achieves the lowest P_M

and the highest P_F. The scheme under case 2 has the highest P_M and a lower P_F than that under case 1. This is because phase noise increment variance is always small, and minor changes of phase noise are difficult to detect. At low SNR, the proposed scheme under case 3 has the lowest P_F. At high SNR, this scheme has almost the same P_F (which approaches 0) under all three cases. The proposed scheme under case 1 can decrease P_M by 9% over that under case 3 and 85% under case 2. It indicates that the proposed scheme can reap performance benefits by utilizing channel gains and phase noise for high SNR. In addition, the above results imply that there exists a trade-off between reliability and security in terms of P_F and P_M.

According to [74], multipath delay is generally at the microsecond level and time interval [14] and phase noise increment are assumed to have identical variances. In comparison with [14], we derive variances of practical estimation errors in terms of channel gains and phase noise. Since the dimension of phase noise is more than that of the time interval by 1, P_M in our proposed scheme is smaller than that in [14].

To fairly compare our proposed scheme with that in [14], the number of multipath channels is set to 6 and thus that of time interval for multipath delays is 5. Assume there is an error-free quantization for each quantizer. If we use $\boldsymbol{\tau}_A^{[i]}$ (resp. $\boldsymbol{\tau}_E^{[i]}$) ($i = 1, 2, ..., 5$) to denote the i^{th} time interval for Alice (resp. Eve), then $\boldsymbol{\tau}_A^{[i]}$ (resp. $\boldsymbol{\tau}_E^{[i]}$) is exponentially distributed random variable with parameter $\lambda = \frac{1}{\sigma^2_{\tau_A}}$ (resp. $\lambda = \frac{1}{\sigma^2_{\tau_E}}$) (please refer to [14] for details). Let $\kappa_\tau = \frac{\sigma^2_{\tau_E}}{\sigma^2_{\tau_A}}$ be the ratio of average time interval for Alice and Eve, and ρ be correlation coefficient of time interval. Fig. 4.2(b) illustrates the performance comparison between our proposed scheme under case 1 and that in [14], given that $Z = 3$, $\kappa_h = \kappa_\Delta = \kappa_\tau = 0$ dB, $\delta_h = 0.5$, $\delta_\theta = \delta_\tau = 0.0815$, $L_t = 3$, $L_p = 6$, $\alpha = \rho = 0.9$, and $P_{e_h} = P_{e_\theta} = 0$. Specifically, the former significantly outperforms the latter in terms of P_M in the whole SNR region, and both works exhibit the same P_F (which approaches 0) in a relatively high SNR regime. This is because increasing SNR can improve estimation accuracies for channel gains and

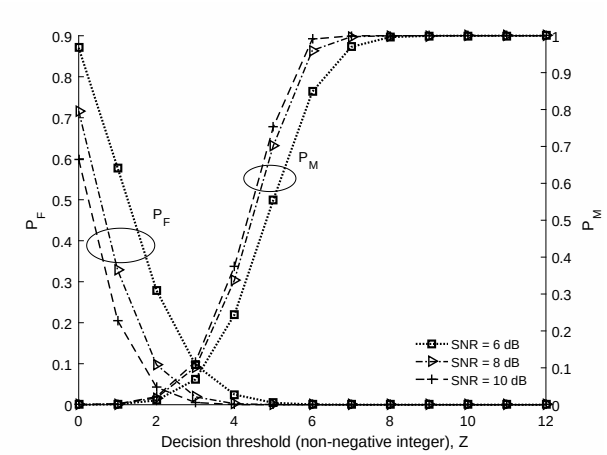
phase noise, and thus reduce P_F. As compared with [14], the proposed scheme under case 1 has benefited immensely from a relatively high SNR regime.

It is notable, however, that the power level of wireless networks is generally required to be below a certain level because of energy constraint and interference among simultaneous transmissions. Therefore, it is necessary to find the optimal setting of parameters (e.g., Z, δ_h, and δ_θ) to achieve a desired authentication performance (P_F, P_M) constraint under a given power limitation.

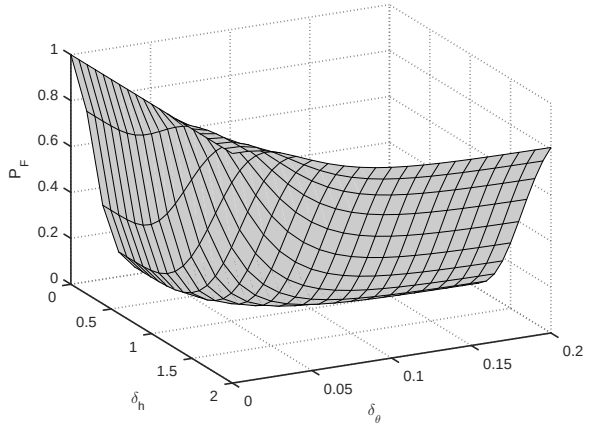
4.4.3 Control of P_F and P_M

With the help of our theoretical results for P_F and P_M, we explore in Fig. 4.3 how the proposed scheme under case 1 enables performance (P_F, P_M) to be flexibly controlled by thresholds (i.e, Z, δ_h, and δ_θ) in a large region. We first summarize in Fig. 4.3(a) that under different SNR scenarios, how (P_F, P_M) varies with Z, given that $\kappa_h = \kappa_\Delta = 0$ dB, $L_t = 3$, $L_p = 6$, $\alpha = 0.9$, and $P_{e_h} = P_{e_\theta} = 0$. For a particular SNR, as Z increases, P_F declines monotonously and even approaches 0, while P_M increases monotonously and then approaches 1, i.e., a larger Z leads to a lower P_F and higher P_M. We can also see from Fig. 4.3(a) that for a fixed Z, when SNR increases, P_F decreases while P_M increases. To ensure secure communications, both P_F and P_M are commonly required to be below 0.1 [12, 31]. Therefore, for a given Z and SNR (i.e., $Z = 3$ and SNR = 8 dB), we need to examine how to adjust other thresholds (e.g., δ_h and δ_θ) to achieve the required constraints (e.g., $P_F, P_M < 0.1$).

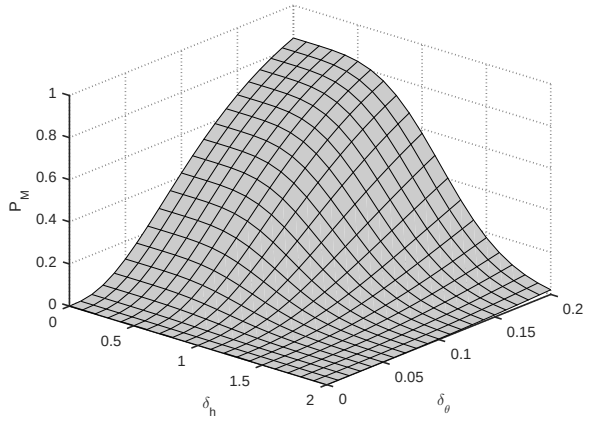
Fig. 4.3(b) and Fig. 4.3(c) demonstrate how P_F and P_M vary with parameters (δ_h, δ_θ), respectively. As observed in Fig. 4.3(b) (resp. Fig. 4.3(c)) that for a specified constraint p_f of P_F (resp. p_m of P_M), we can accordingly set a specified constraint plane intersecting with z-axis orthogonally at the point (1, 0.1, p_f) (resp. (1, 0.07, p_m)), and then can determine a set of (δ_h, δ_θ)-pairs corresponding to the surface below the defined constraint plane. Finding the intersection of these two sets of (δ_h,



(a) (P_F, P_M) vs. Z with $\delta_h = 0.5$, and $\delta_\theta = 0.0615$.



(b) P_F vs. $(\delta_h, \delta_\theta)$ with $Z = 3$ and SNR = 8 dB.



(c) P_M vs. $(\delta_h, \delta_\theta)$ with $Z = 3$ and SNR = 8 dB.

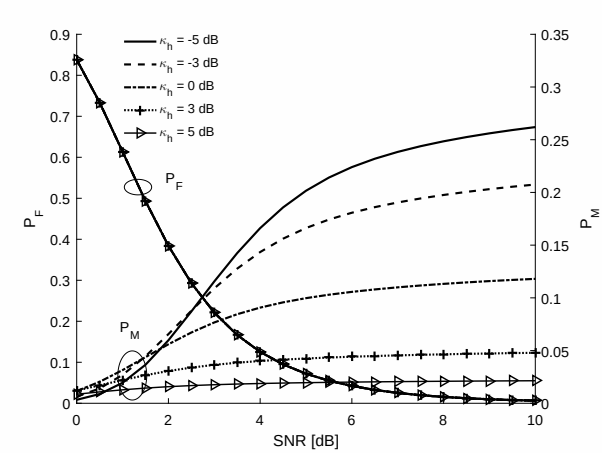
Figure 4.3: (P_F, P_M) vs. $(Z,\ \delta_h,\ \delta_\theta)$ with the settings ($\kappa_h = \kappa_\Delta = 0$ dB, $L_t = 3$, $L_p = 6$, $\alpha = 0.9$, $P_{e_h} = P_{e_\theta} = 0$).

δ_θ)-pairs yields the region of (δ_h, δ_θ)-pairs to achieve p_f and p_m. E.g., for $\delta_h \in [0, 1]$ and $\delta_\theta \in [0.005, 0.08]$, one can easily see from Fig. 4.3(b) and Fig. 4.3(c) that the constraint $P_F \leq 0.1$ can be achieved in the region of ($\delta_h \in [0, 1]$, $\delta_\theta \in [0.005, 0.08]$), while constraint $P_M \leq 0.1$ is achieved in the region of ($\delta_h \in [0, 2]$, $\delta_\theta \in [0.005, 0.015]$). Thus, the requirement of (P_F, $P_M \leq 0.1$) for the concerned network scenario is achieved under $\delta_h \in [0, 1]$ and $\delta_\theta \in [0.005, 0.08]$. From Fig. 4.3, we can find that the proposed authentication scheme under case 1 is flexible and general. (P_F, P_M) can be flexibly controlled by adjusting decision threshold Z, channel gain threshold δ_h, and phase noise threshold δ_θ. In addition, a trade-off between reliability and security can be achieved by an appropriate setting of (Z, δ_h, δ_θ). Based on this background, we need to further explore authentication efficiency of the proposed scheme.

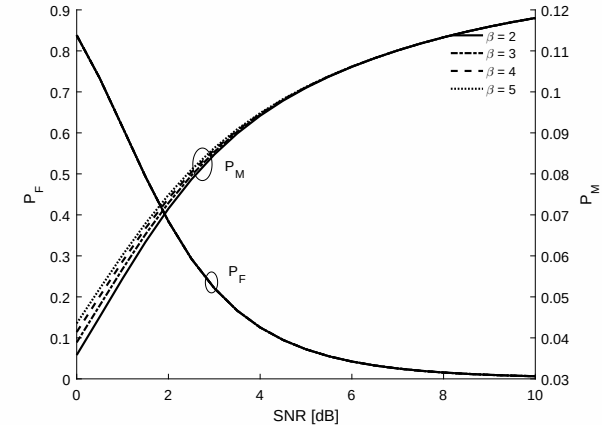
4.4.4 Authentication Efficiency Analysis

To present authentication efficiency, we now study the authentication performance of the proposed scheme under case 1 for various network scenarios (κ_h, κ_Δ L_t, L_p, α, $P_{e_h}(P_{e_\theta})$) in Fig. 4.4. Fig. 4.4(a) shows how the location of Eve under small-scale variations of channels, characterized by κ_h, can impact (P_F, P_M) with the settings ($\kappa_\Delta = 0$ dB, $Z = 3$, $\delta_h = 0.5$, $\delta_\theta = 0.0615$, $L_t = 3$, $L_p = 10$, $\alpha = 0.9$, $P_{e_h} = P_{e_\theta} = 0$). As observed from Fig. 4.4(a) that for a given κ_h, P_F is not affected by κ_h and P_M decreases when κ_h increases. At low SNR, different κ_h yield nearly indistinguishable P_M. At high SNR, κ_h has a significant impact on P_M, especially when $\kappa_h < 0$ dB, P_M will exceed 0.1. This indicates that for a fixed location of Alice, Eve might search a "good" location in which she has a high probability to impersonate Alice successfully.

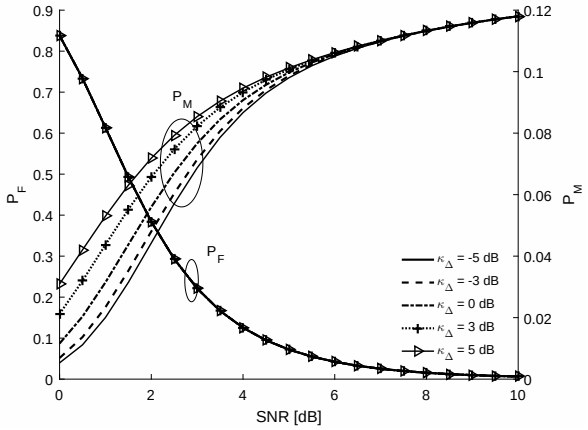
Now, we analyze the spatially-averaged performance under a distance dependent large-scale fading, i.e., κ_h is a random variable over all possible joint locations of Alice and Eve. SNR for both Alice and Eve can be maintained the same by using power



(a) Impact of κ_h on (P_F, P_M) with $\kappa_\Delta = 0$ dB.



(b) Impact of β on (P_F, P_M) with $\kappa_\Delta = 0$ dB.

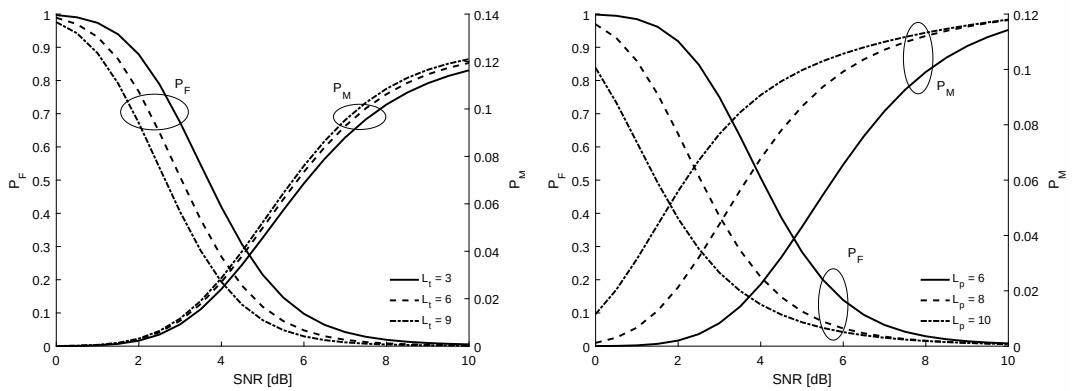


(c) Impact of κ_Δ on (P_F, P_M) with $\kappa_h = 0$ dB.

Figure 4.4: (P_F, P_M) vs. SNR with the settings ($Z = 3$, $\delta_h = 0.5$, $\delta_\theta = 0.0615$, $L_t = 3$, $L_p = 10$, $\alpha = 0.9$, $P_{e_h} = P_{e_\theta} = 0$).

control. Combining (5.30) and (4.38), we can obtain a new $P_h^{H_1}$ and then substitute it into (5.32) to acquire P_M under any κ_h through the complex integration in Matlab. Fig. 4.4(b) illustrates how path loss exponent, β covering a wide practical range can impact (P_F, P_M). As illustrated in Fig. 4.4(b), P_F is irrelevant to β, and P_M slightly rises (a small range: 0-0.02) at low SNR as β increases. While at high SNR, P_M is not affected by β. This shows that the proposed scheme is reasonably robust to large-scale fading and random location of Eve.

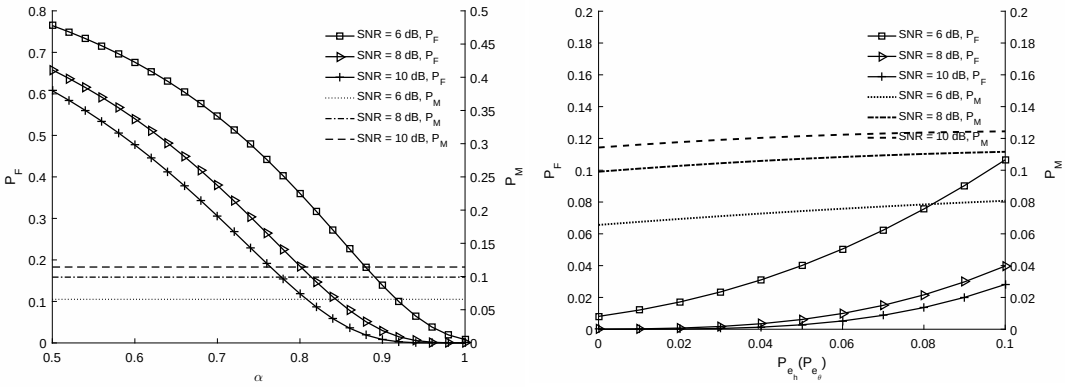
Fig. 4.4(c) shows how κ_Δ can impact (P_F, P_M) with the settings ($\kappa_h = 0$ dB, $Z = 3$, $\delta_h = 0.5$, $\delta_\theta = 0.0615$, $L_t = 3$, $L_p = 10$, $\alpha = 0.9$, $P_{e_h} = P_{e_\theta} = 0$). As shown in Fig. 4.4(c), for a given SNR, P_F remains unchanged and P_M rises when κ_Δ increases. This indicates that using a high-quality hardware has a low phase noise increment variance, and thus identifying phase noise faces a large challenge. Although Eve can choose a "good" position and/or a higher-quality hardware, the proposed scheme under case 1 can effectively resist against impersonation and/or replay attacks by setting parameters (e.g., Z, δ_h, and δ_θ) properly.



(a) Impact of L_t on (P_F, P_M) with $L_p = 6$. (b) Impact of L_p on (P_F, P_M) with $L_t = 3$.

Figure 4.5: Impact of (L_t, L_p) on (P_F, P_M) with the settings ($Z = 3$, $\kappa_h = \kappa_\Delta = 0$ dB, $\delta_h = 0.5$, $\delta_\theta = 0.0615$, $\alpha = 0.9$, $P_{e_h} = P_{e_\theta} = 0$).

Fig. 4.5 shows that how L_t and L_p can impact (P_F, P_M) with the settings ($Z = 3$, $\kappa_h = \kappa_\Delta = 0$ dB, $\delta_h = 0.5$, $\delta_\theta = 0.0615$, $\alpha = 0.9$, $P_{e_h} = P_{e_\theta} = 0$). For a given SNR,



(a) (P_F, P_M) vs. α with $P_{e_h} = P_{e_\theta} = 0$. (b) (P_F, P_M) vs. $P_{e_h}(P_{e_\theta})$ with $\alpha = 0.9$.

Figure 4.6: (P_F, P_M) vs. $(\alpha,\ P_{e_h},\ P_{e_\theta})$ with the settings ($Z = 3$, $\kappa_h = \kappa_\Delta = 0$ dB, $\delta_h = 0.5$, $\delta_\theta = 0.0615$, $L_t = 3$, $L_p = 6$).

as L_t (resp. L_p) increases, P_F increases and P_M decreases. L_p has a more significant impact on $(P_F,\ P_M)$ by comparing Fig. 4.5(a) and Fig. 4.5(b). The reason is that for a larger L_t and L_p, Bob can estimate channel gains and phase noise more accurately, and thus improve $(P_F,\ P_M)$. It reveals that our proposed scheme is more sensitive to the length of pilot symbols. Under a specified constraint of P_F and L_t we need to properly set L_p so that the proposed scheme can be efficient to achieve secure communications in MIMO systems.

Herein, we use channel correction coefficient α to characterize temporal channel variations when the last one of continuous multiframes arrives at Bob. Bob just needs to estimate the channel gain and phase noise corresponding to the last frame. Fig. 4.6 demonstrates that (P_F, P_M) vs. $(\alpha,\ P_{e_h},\ P_{e_\theta})$ with the settings ($Z = 3$, $\kappa_h = \kappa_\Delta = 0$ dB, $\delta_h = 0.5$, $\delta_\theta = 0.0615$, $L_t = 3$, $L_p = 6$). Fig. 4.6(a) illustrates that how α can impact $(P_F,\ P_M)$. As shown in Fig. 4.6(a), for a given SNR, when α rises P_F monotonically decreases and even approaches 0 when $\alpha \to 1$ while P_M maintains unchanged. This indicates that excessive false alarm events may happen in a higher mobile scenario and fewer false alarm events may occur in a low mobility scenario. In other words, the proposed scheme enables Bob to authenticate many frames from the same transmitter without sacrificing authentication performance in low mobility

scenarios, because $P_F \rightarrow 0$ as α approaches 1 and P_M maintains unchanged in a relatively stable environment.

Finally, we examine in Fig. 4.6(b) how (P_F, P_M) varies with quantization error probabilities, P_{e_h} and P_{e_θ}. For simplicity, we assume that $P_{e_h} = P_{e_\theta}$. As illustrated in Fig. 4.6(b), for a given SNR, decreasing P_{e_h} (P_{e_θ}) leads to different declines in the shape of P_F and P_M, and P_F is more sensitive to P_{e_h} (P_{e_θ}) than P_M. For a given P_{e_h} (P_{e_θ}), increasing SNRs can contribute to achieving a low P_F but incurring a high P_M. Generally, quantization error is closely related to SNR. In order to control P_{e_h} (P_{e_θ}) at a low level, it is necessary to improve SNR. However, a high SNR may cause a large P_M. Therefore, we need to adjust system parameters (e.g., SNR) accordingly to achieve a desired authentication performance.

4.5 Summary

Distinguished from [14] focusing on wireless channel feature (e.g., channel impulse response in the dimensions of amplitude and path delay), this chapter attempted to jointly take both wireless channel and hardware features into account for authentication. To this end, we proposed a new physical layer authentication scheme jointly utilizing channel gains and phase noise in heterogeneous MIMO systems. We also determined variances of estimation errors in terms of channel gains and phase noise, and then derived closed-form expressions for false alarm and missed detection probabilities while taking quantization errors into account. We further demonstrated that the proposed scheme enables flexible performance control by adjusting thresholds (for channel gain, phase noise, and decision, respectively). This indicates that the proposed scheme has the capability of satisfying different performance requirements for future emerging heterogeneous MIMO systems. The results in this chapter enable us to find the graceful tradeoff between reliability and security requirements, and we expect the methodology developed in this chapter to be valuable for devising new

physical layer authentication schemes in other types of networks.

CHAPTER V

End-to-End Physical Layer Authentication for Dual-Hop Wireless Networks

End-to-end (E2E) physical layer authentication for multi-hop wireless networks is still not well-explored by now. As one step forward in this direction, this chapter focuses on the E2E physical layer authentication in a dual-hop wireless network with an untrusted relay and proposes a corresponding physical layer authentication scheme. The scheme fully utilizes the location-specific features of both channel gain and delay interval of cascaded channels, and also adopts the artificial jamming technique, so that it is not only resistant to the impersonate attack from an unauthorized transmitter but also resilient to the replay attack from the untrusted relay. Theoretical analysis is further conducted to derive the expressions for the probabilities of false alarm and missed detection, which serves as two fundamental metrics of authentication performance. Finally, numerical and simulation results are provided to illustrate both the efficiency of these theoretical results and the E2E authentication performance of dual-hop wireless networks under the proposed scheme.

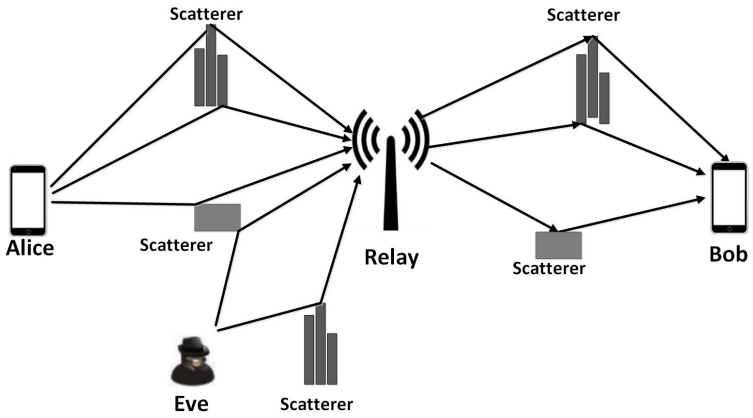


Figure 5.1: System model. The transmitter Alice (A) communicates with the receiver Bob (B) with the help of an AF untrusted relay (R), and Eve (E) serves as the adversary who impersonates A. The transmissions between A (E) and R, R and B experience different multipath effects.

5.1 System Model

5.1.1 Network Model

As depicted in Fig. 5.1, we consider a dual-hop wireless network scenario consisting of four entities: one legitimate transmitter Alice (A), one untrusted relay (R) adopting the amplify-and-forward (AF) strategy, one intended receiver Bob (B) and one adversary Eve (E). Each entity is equipped with an omnidirectional antenna and operates in half-duplex mode. These entities are spatially located at separated positions in a rich-scattering environment (e.g., urban areas). We assume that the spatial separation between any two entities is more than a distance of half a wavelength, making fading paths independent of each other according to the well-known Jakes uniform scattering model [30]. This is widely adopted in [12, 14, 31, 51] and reasonable because if two entities are spaced within a distance of half a wavelength, they will fail to work due to mutual strong interference.

The direct link between A (E) and B suffers from the deep fading so that they can only communicate with the help of R. R can record all the received signals and

then initiate replay attacks with the aggressive signals, that is, R serves as a potential adversary. The adversary E monitors the network and attempts to inject aggressive signals into the network in the hope of impersonating A. B not only can receive signals but also can transmit signals (e.g., jamming signals) to R. We assume that adversaries have the knowledge of the modulation method employed in the network, the channel estimation technique employed by B and authentication scheme adopted in the network [2, 31]. Similar to that of previous studies [12, 31], we further assume that E cannot arrive at A's previous location before a new signal arrives at B.

Suppose that B receives two messages (i.e., frames) at times $k-1$ and k (time interval is much less than the channels coherence time). The first one is validated by B from Alice using a standard higher-layer protocol [75]. Based on this authentication, B measures and stores CA and DI at time $k-1$. The objective of authentication at B is that utilizing multipath channel estimation for R-B and A-R-B in terms of gain and delay based on previous message originated from A, he needs to decide whether the second message received at time k is still from A. The message to be authenticated is not required to transmit continuously but it is necessary to ensure the continuity of the authentication process by probing the channel at time intervals smaller than the channels coherence time [1, 12, 14, 31, 33].

We assume that the channels are reciprocal and remain correlated within the total processing time, which mainly includes propagation delay, transmitting time and operation (e.g., AF operation) delay at each entity. This is due to the fact that the total processing time is much less than channel coherent time T_C, For example, for a typical $f = 2.4$ GHz radio frequency carrier, relative motion speed $v = 2$ m/s, T_C can be calculated as $T_C = \frac{9c}{16\pi vf}$ =11.2 ms (c is light speed). The propagation delay might be 30 μ s for a rich scattering environment (will be at least 10 μs) if the distance is 3 km. Transmitting time is 0.5 μs for a frame consisting of 20 symbols when the symbol sampling rate is 40 MHz. In general, the operation delay has the

same order of transmitting time.

5.1.2 Channel Model

Since all entities are in a rich scattering and reflecting environment, channel impulse response of each hop in the concerned dual-hop network is modeled as a sum of paths with time-varying CA and propagation delay caused by the changes in the propagation environment (e.g., relative motion between entities and/or movement of scatterers/reflectors). Considering N_{ij} multipath channels between two entities i and j, channel impulse response measured at time k under the delay spread index denoted by d can be expressed by

$$h_{ij}(k,d) = \sum_{n_{ij}=1}^{N_{ij}} h_{n_{ij}}(k)\delta(k - d_{n_{ij}}(k)),\ n_{ij} = 1, ..., N_{ij}. \tag{5.1}$$

where $h_{n_{ij}}(k)$ and $d_{n_{ij}}(k)$ $(d_{1_{ij}}(k) < d_{2_{ij}}(k) < \cdots < d_{N_{ij}}(k))$ are the time-varying CA and propagation delay associated with the ${n_{ij}}^{\text{th}}$ multipath component, respectively, and $\delta(\cdot)$ is the Dirac pulse function. Subscript ij can be AR, ER or RB. Hence, the CA and propagation delay of the ${n_{AR}}^{\text{th}}$, ${n_{RB}}^{\text{th}}$ and ${n_{ER}}^{\text{th}}$ multipath components for A-R, R-B and E-R are denoted by $h_{n_{AR}}(k)$, $h_{n_{RB}}(k)$ and $h_{n_{ER}}(k)$, and $d_{n_{AR}}(k)$, $d_{n_{RB}}(k)$ and $d_{n_{ER}}(k)$, respectively.

Each multipath component is assumed to be suffer from statistically independent Rayleigh fading and CA and propagation delay of it are assumed to remain constant over a frame (message transmission is organized by frame-by-frame) but to vary independently and continuously from one frame to the next. Such temporal variations in terms of CA and delay are highly correlated [58, 76]. Thus, $h_{n_{ij}}$ follows a zero-mean complex Gaussian distribution. In general, for a specific multipath channel, their channel gains might not be identical variance (e.g., in the case of an exponential power delay profile). We assume that channel gains have an identical variance, i.e.,

$h_{n_{ij}} \sim \mathcal{CN}(0, \sigma^2_{h_{ij}})$ [12, 14, 31].

To explore the temporal variation of channel gain, we need to investigate the auto-correlation function of channel gain. According to the Jakes' model [30], channel variation is affected by the Doppler frequency. Similar to that of previous studies [14, 31, 50], identical maximum Doppler frequency is considered in multipath channels. Then, the auto-correlation function of $h_{n_{ij}}(k)$ under arbitrary time lag t_s, can be given by

$$\varphi_{ij}(k_s) = \mathbf{E}\{h_{n_{ij}}(k)h^*_{n_{ij}}(k+k_s)\} = \sigma^2_{h_{ij}} J_0(2\pi f_{ij} k_s), \tag{5.2}$$

where $J_0(\cdot)$ is the zero$^{\text{th}}$ order Bessel function of the first kind. f_{ij} is maximum normalized Doppler frequency. Based on the above results, we employ the auto-regressive model of order 1 (AR-1) [14, 31, 50] to describe time-varying channel gain, and then we have

$$h_{n_{ij}}(k) = \alpha_{ij} h_{n_{ij}}(k-1) + \sqrt{1-\alpha^2_{ij}}\, u_{n_{ij}}(k), \tag{5.3}$$

where AR coefficient α_{ij} is denoted by $\varphi_{ij}(1)/\sigma^2_{h_{ij}}$; $u_{n_{ij}} \sim \mathcal{CN}(0, \sigma^2_{h_{ij}})$ is independent of $h_{n_{ij}}(k-1)$.

Moreover, the propagation delay of multipath components can be modeled by a Poisson process [77]. Therefore, delay interval (DI) between two delays of adjacent multipath components at time k is an exponentially distributed random variable, which is defined by

$$\tau_{k_{ij}}(k) \triangleq d_{k_{ij}}(k) - d_{k-1_{ij}}(k),\ k = 1, 2, ..., N_{ij} - 1. \tag{5.4}$$

One can easily see that $\tau_{k_{ij}}(k)$ is also time-varying. Similar to the simple assumption for $h_{n_{ij}}$, we also assume that $\tau_{k_{ij}}(k)$ is statistically independent and identically

distributed random variable [14].

We adopt correlated Gaussian random variables to characterize the correlation between $\tau_{k_{ij}}(k-1)$ and $\tau_{k_{ij}}(k)$. This is reasonable due to the fact that an exponentially distributed random variable can be decomposed as the sum of the squares of two independent Gaussian distributed random variables. Hence, $\tau_{k_{ij}}(k-1)$ and $\tau_{k_{ij}}(k)$ can be decomposed, respectively, as

$$\tau_{k_{ij}}(k-1) = \tau^{(1)}_{k_{ij}}(k-1) + \tau^{(2)}_{k_{ij}}(k-1), \tag{5.5}$$

$$\tau_{k_{ij}}(k) = \tau^{(1)}_{k_{ij}}(k) + \tau^{(2)}_{k_{ij}}(k), \tag{5.6}$$

where $\tau^{(1)}_{k_{ij}}, \tau^{(2)}_{k_{ij}}$ are mutually independent Gaussian distributed random variables with zero mean and variance $\sigma^2_{\tau_{ij}}$. Similar to previous work [14], we also use AR-1 to model the temporal processes of $\tau^{(1)}_{k_{ij}}$ and $\tau^{(2)}_{k_{ij}}$, and then we have

$$\tau^{(\ell)}_{k_{ij}}(k) = \beta_{ij}\tau^{(\ell)}_{k_{ij}}(k-1) + \sqrt{(1-\beta^2_{ij})\sigma^2_{\tau_{ij}}}u^{(\ell)}_{k_{ij}}(k),\ \ell = 1, 2, \tag{5.7}$$

where AR coefficient β_{ij} has the similar definition with α_{ij} given in (5.3); $u^{(\ell)}_{ij,k}(k) \sim N(0,1)$ is independent of $\tau^{(\ell)}_{ij,k}(k-1)$.

5.1.3 Communication Model

Frame-by-frame transmission is considered in this paper. A transmission frame consists of deterministic pilot symbols used for channel estimation and stochastic data symbols. When A transmits a signal $s(k)$ to B at time k with the aid of R, the total transmission is accomplished by the following two phases. For Phase I, A transmits the signal $s(k)$ at the average transmitted power P to R and the signal received at R

$$y_B(k) = \xi\sqrt{P}\sum_{n_{AR}=1}^{N_{AR}}\sum_{n_{RB}=1}^{N_{RB}} h_{n_{AR}}(k)h_{n_{RB}}(k)s(k-d_{n_{AR}}(k)-d_{n_{RB}}(k)) + \xi\sqrt{P}\sum_{n_{RB}=1}^{N_{RB}} h_{n_{RB}}(k)w_R(k-d_{n_{RB}}(k)) + w_B(k). \quad (5.9)$$

---

is

$$y_R(k) = \sqrt{P}\sum_{n_{AR}=1}^{N_{AR}} h_{n_{AR}}(k)s(k-d_{n_{AR}}(k)) + w_R(k), \quad (5.8)$$

where $w_R \sim \mathcal{CN}(0, \sigma_w^2)$ is AWGN.

For Phase II, $y_R(k)$ is then multiplied by an amplification factor ξ and retransmitted to B at power P. The amplification factor commonly used in the literature is $\xi = \sqrt{\frac{P}{P\sigma_{h_{AR}}^2+\sigma_w^2}}$.

Since the transmitting time and operation delay are much less than the operation delay (see Section 5.1.1), so we can neglect the transmitting time and operation delay. Therefore, the AWGN at B is denoted by $w_B \sim \mathcal{CN}(0, \sigma_w^2)$, and then the corresponding received signal at B is given in (5.9),

From (5.9), we know that the channel impulse response from the transmitter to the receiver via the relay is a cascade of the multipath channels in two hops, i.e., it is a cascaded multipath channel. According to [76], the double (cascaded) Rayleigh fading model can be used to characterize such cascaded multipath components. The CA of each cascaded multipath component is the product of CAs of two multipath components in two hops over this cascaded path, and the corresponding delay is the sum of delays of two multipath components in two hops over this cascaded path. The CA of the n_{AB}^{th} cascaded multipath components for *A*-*R*-*B* is denoted by

$$h_{n_{AB}}(k) = h_{n_{AR}}(k)h_{n_{RB}}(k), \quad (5.10)$$

where $n_{AB} = 1, 2, ..., N_{AR}N_{RB}$, $n_{AR} = 1, 2, ..., N_{AR}$ and $n_{RB} = 1, 2, ..., N_{RB}$.

Then, the corresponding DI is denoted by

$$\tau_{k_{AB}}(k) = \tau_{k_{AR}}(k) + \tau_{k_{RB}}(k), \tag{5.11}$$

where $k_{AB} = 1, 2, ..., N_{AR}N_{RB} - 1$, $k_{AR} = 1, 2, ..., N_{AR} - 1$ and $k_{RB} = 1, 2, ..., N_{RB} - 1$. When E transmits signals to B, it has the same the transmission process.

Using methods in [78–82], B can estimate cascaded multipath channel parameters from the received signal in (5.9) In addition, channel and delay estimation are corrupted by estimation error (additive noise), and such estimation error is much less than the temporal variations of channel and delay [14]. Therefore, both CA and delay of channels can be utilized to differentiate between the legitimate transmitter A and illegitimate transmitter E.

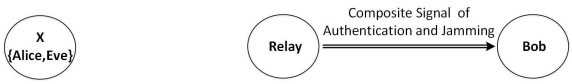
5.2 Proposed E2E Authentication Scheme

The basic principle for the proposed E2E scheme is that the channels are location-specific, which has been widely adopted to complement and improve traditional security approaches [12, 14, 31, 51]. Most importantly, this is demonstrated by the well-known Jakes model [30], that is, the spatial separation of one to two wavelengths results in independent fading channels. It is difficult (if not impossible) for an attacker to generate or accurately model the channel being used by the transmitter-receiver pair. In other words, the channels between different geographic locations decorrelate rapidly in space due to path loss and channel fading [30, 31]. As a result, the channel of A-R-B is independent of that of E-R-B. Meanwhile, the channel for the identical transmitter-receiver pair is highly corrected over time. Hence, CA and DI of multipath channels can be jointly exploited to authenticate transmitters.

The main procedures of the proposed E2E scheme are illustrated in Fig. 5.2,

Figure 5.2: The main procedures of the proposed E2E authentication scheme.

(a) First phase transmission.



(b) Second phase transmission.

Figure 5.3: Transmissions of authentication and jamming signals.

where a typical challenge-response procedure is first conducted to initiate the authentication process. The transmissions of authentication and jamming signals is then implemented between entities involved in communication. Finally, B carries out a verification procedure to verify whether the current frame is from A or not.

5.2.1 Challenge-response Procedure

In the available authentication schemes developed for wireless systems with direct link between an unknown transmitter X (i.e., $X = \{A, E\}$) and the receiver B [12, 14, 31, 51], the transmitter can directly send authentication signals to the receiver B anytime without pursing the synchronization with the receiver in advance. For the dual-hop wireless system with an untrusted relay R concerned in this paper, however, to deal with the possible replay attack from R based on our proposed scheme, the synchronization between the transmitter and receiver B is required before the authentication process. For this purpose, the transmitter first sends an authentication request frame (i.e., a challenge), which contains only the synchronization signal indicating the time that the authentication signal is to be transmitted, such that the synchronization between the transmitter and B can be achieved in the next transmission process of authentication and jamming signals.

$$y_R(k) = \sqrt{P_A} \sum_{n_{AR}=1}^{N_{AR}} h_{n_{AR}}(k) s(k - d_{n_{AR}}(k)) + \sqrt{P_B} \sum_{n_{RB}=1}^{N_{RB}} h_{n_{RB}}(k) v(k - d_{n_{RB}}(k)) + w_R(k). \quad (5.12)$$

After receiving the authentication request from A, B sends back a response frame which only includes one symbol to confirm the requested time for synchronization.

5.2.2 Transmissions of Authentication and Jamming Signals

As illustrated in Fig. 5.3, the transmissions of authentication and jamming signals at time t includes two phases. In the first phase transmission (Fig. 5.3(a)), transmitter X sends a frame including the signal to be authenticated to R. This signal consists of both pilot and data symbols. Concurrently, receiver B sends a jamming frame with the jamming signal to R in a cooperative manner, making the untrusted R to receive a composite signal of authentication and jamming signals. Here, the jamming signals can be generated by utilizing a pseudo-random noise generator such as the one in [83] and is then stored at B. Without loss of generality, we use $v(k)$ and P_B to denote jamming signal and the average symbol power transmitted by B. When the transmitter is A (i.e., $X = A$), (5.8) becomes (5.12). As shown in Fig. 5.3(b), the composite signal given in (5.12) is then multiplied by an amplification factor at R and retransmitted to B in the second phase transmission [58].

5.2.3 Verification Procedure

5.2.3.1 CA/DI Estimation

After receiving the composite signal of jamming and authentication signals, B first removes the jamming signal through the well-developed self-interference cancellation

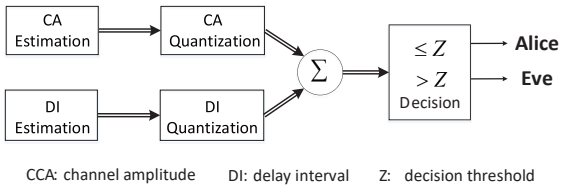


Figure 5.4: Illustration of CA/DI estimation/quantization and decision.

techniques [84, 85], and thus extracts the desired signal (i.e., consisting of pilot and data symbols) from the unknown transmitter X (when $X = A$, the received signal at B has the same expression as (5.9)). Since the generated jamming signal is known to B, it can be eliminated up to a certain extent through efficient techniques of interference cancellation [84, 85]. To present the optimal authentication performance of the proposed scheme, we follow the ideal assumption here that perfect cancellation of self-interference is achievable [86]. As shown in Fig. 5.4, the estimation of CA/DI are first performed. Based on the estimation results of CA/DI, quantization is to qualify the temporal variations of CA/DI. Finally, B decides whether the current frame is from A or not under a simple binary hypothesis test.

Note that CA and DI of the cascaded channel can be estimated by using the deterministic pilot symbols, but only a noisy version of cascaded channel is available at B due to the presence of AWGN. The estimation error caused by such AWGN is random and independent of the channel, and can be modeled as AWGN random variable with zero mean.

CA Estimation: Let $\hat{h}_{n_{AB}}(k)$ (resp. $\hat{h}_{EB,n_{EB}}(k)$) denote the estimation of $h_{n_{AB}}(k)$ (resp. $h_{EB,n_{EB}}(k)$) at time k, and then each of them can be modeled as a sum of its real value and a complex Gaussian noise (i.e., estimation error) $w_h(k) \sim \mathcal{CN}(0, \sigma_w^2)$. Then, we have

$$\begin{aligned}\hat{h}_{XB,l_{XB}}(k) &\triangleq h_{XB,l_{XB}}(k) + w_h(k) \\ &= h_{XR,l_{XR}}(k)h_{n_{RB}}(k) + w_h(k),\ X = \{A, E\}. \end{aligned} \tag{5.13}$$

where σ_w^2 is defined as $\sigma_w^2 = P_w/P$ with P_w representing the average noise power at the receiver, respectively [12].

DI Estimation: Actually, the DI of each adjacent paths pair in a cascaded channel can be considered to be the superposition of DIs of each adjacent multipath components pair in two hops over that paths pair. This is because the delays of multipath channels are spatially uncorrelated, and the processing time at each entity is sufficiently small and thus can be neglected [2].

Let $\hat{\tau}_{k_{AR}}^{(1)}(k)$ and $\hat{\tau}_{k_{AR}}^{(2)}(k)$ represent the estimations of $\tau_{k_{AR}}^{(1)}(k)$ and $\tau_{k_{AR}}^{(2)}(k)$, respectively. Similar to the estimation of CA, each of them can also be considered as a sum of its real value and an AWGN (estimation error) $w_\tau(k)$, they can be written, respectively, as

$$\hat{\tau}_{k_{AR}}^{(\ell)}(k) \triangleq \tau_{k_{AR}}^{(\ell)}(k) + w_\tau^{(\ell)}(k),\ \ell = 1, 2. \tag{5.14}$$

where $w_\tau^{(\ell)}(k) \sim \mathcal{N}(0, \sigma_w^2/2)$. One can easily see that $\hat{\tau}_{k_{AR}}^{(\ell)}(k) \sim \mathcal{N}(0, \sigma_{\tau_{AR}}^2 + \sigma_w^2/2)$. Let $\hat{\tau}_{k_{AR}}(k)$ denote the estimation of $\tau_{k_{AR}}(k)$, and then we have

$$\hat{\tau}_{k_{AR}}(k) = (\hat{\tau}_{k_{AR}}^{(1)}(k))^2 + (\hat{\tau}_{k_{AR}}^{(2)}(k))^2. \tag{5.15}$$

It can be seen from (5.15) that $\hat{\tau}_{k_{AR}}(k)$ also follows exponential distribution with parameter $\lambda = \frac{1}{2\sigma_{\tau_{AR}}^2 + \sigma_w^2}$. By using the similar derivation, we can see that both $\hat{\tau}_{k_{RB}}(k)$ and $\hat{\tau}_{k_{ER}}(k)$ are also exponentially distributed random variables with parameters $\lambda = \frac{1}{2\sigma_{\tau_{RB}}^2 + \sigma_w^2}$ and $\lambda = \frac{1}{2\sigma_{\tau_{ER}}^2 + \sigma_w^2}$, respectively. The total DI of the k_{XB}^{th} cascaded multipath component can be given by

$$\tau_{k_{XB}}(k) = \hat{\tau}_{k_{XR}}(k) + \hat{\tau}_{k_{RB}}(k). \tag{5.16}$$

5.2.3.2 CA/DI Quantization

CA Quantization: To quantify the temporal variation of CA, we use a CA quantizer which compares the square of absolute value of the CA difference between the current and previous CA estimations of the same path at adjacent time with a specified CA threshold. In particular, when the difference is not larger than the specified CA threshold, the output of CA quantizer is 0; otherwise, the output of that is 1. We use Q_h and $O_{h,n}$ to denote CA quantizer and the n^{th} output of CA quantizer, respectively, where $n \in \{1, 2, ..., N\}$ and $N = \min\{N_{XR}N_{RB}, N_{AR}N_{RB}\}$. Then, CA quantization can be formulated as

$$\begin{aligned} O_{h,n} &\triangleq Q_h[|\hat{h}_{n_{AB}}(k) - \hat{h}_{n_{AB}}(k-1)|^2] \\ &= \begin{cases} 0, \ |\hat{h}_{n_{AB}}(k) - \hat{h}_{n_{AB}}(k-1)|^2 \leq \delta_h, \\ 1, \ \text{otherwise.} \end{cases} \end{aligned} \tag{5.17}$$

where δ_h represents the specified CA threshold.

DI Quantization: To quantify the temporal variation of DI, a DI quantizer is employed by comparing the absolute value of the DI difference between the current and previous DI estimations of the same paths pair at adjacent time with a specified time threshold. Specifically, when the difference is not larger than the specified time threshold, the output of DI quantizer is 0; otherwise, the output of that is 1. Let Q_τ and $O_{\tau,k}$ denote DI quantizer and the k^{th} output of the quantizer, respectively, where $k \in \{1, 2, ..., N-1\}$. Then, DI quantization can be formulated as

$$\begin{aligned} O_{\tau,k} &\triangleq Q_\tau[|\hat{\tau}_{k_{XB}}(k) - \hat{\tau}_{k_{AB}}(k-1)|] \\ &= \begin{cases} 0, \ |\hat{\tau}_{k_{XB}}(k) - \hat{\tau}_{k_{AB}}(k-1)| \leq \delta_\tau, \\ 1, \ \text{otherwise.} \end{cases} \end{aligned} \tag{5.18}$$

where δ_τ is the specified time threshold.

Let D_h and D_τ represent the sum of $O_{h,n}$ and $O_{\tau,k}$, respectively, and then we have $D_h \triangleq \sum_{n=1}^{N} O_{h,n}$ and $D_\tau \triangleq \sum_{k=1}^{N-1} O_{\tau,k}$. It is easy to see that we have $D_h \in \{0, 1, ..., N\}$ and $D_\tau \in \{0, 1, ..., N-1\}$.

5.2.3.3 Decision Criterion

Based on the above quantization results, we establish a decision criterion under a binary hypothesis test to differentiate between the legitimate frame from A and illegitimate frame from E. For simplicity, we denote by D the sum of D_h and D_τ, and then the binary hypothesis test can be formulated as

$$\begin{aligned} H_0 :&\ D \triangleq D_h + D_\tau \leq Z \\ H_1 :&\ D \triangleq D_h + D_\tau > Z, \end{aligned} \tag{5.19}$$

where Z represents a non-negative integer decision threshold between 0 and $2N-1$. Under H_0, the newly received frame at B is still from legitimate transmitter A. Under H_1, it is from adversary E.

5.2.4 Security Analysis

The location-specific characteristics of CA and DI make the proposed scheme immune to impersonate attacks from external attackers. Meanwhile, jamming signals can confuse the untrusted R and thus avoid simple replay attack from the possible internal attacker R. Furthermore, due to the lack of pilots and any pre-known reference symbols in the authentication request signal, the attackers cannot probe channel in advance. All of these properties ensure the security of this scheme, as analyzed in the followings.

(1) A simple attacker E: When attacker E is located near a legitimate transmitter A, it will fail to impersonate A by injecting aggressive signals due to that both CA and DI are location-specific.

(2) A smart attacker E: Due to the presence of artificial jamming, it becomes much more difficult (if not impossible) to estimate multipath channels A-E and E-R via the authentication signal transmitted by A, even the attacker E is close to R. Therefore, the smart attacker E will fail to construct multipath channel A-R and E-R and impersonate A by modifying its signal.

(3) Untrusted relay R: Although the untrusted relay R or other active attacker E are able to replay signals based on what they obtained, the proposed scheme can be immune to such attacks. This is mainly due to the fact that artificial jamming is randomly generated by B and different artificial jamming sequences that are unknown to attackers are generated at different time sots.

Remark 5 *In fact, the system model in terms of network topology is symmetrical. By setting work pattern of entities (e.g., Alice has the ability to estimate, qualify channel and make a decision), E2E mutual authentication between Alice and Bob can be achieved.*

5.3 Modeling of FA and MD Probabilities

In this section, we first derive some basic results regarding the probabilities that the outputs of two quantizers under two hypotheses are 1, respectively, and then use the results to derive the expressions for P_F and P_M. For simplicity, we use $P_\imath^{H_\jmath}$ to denote the probability that $Q_\imath$ outputs 1 on $H_\jmath$, for $\imath = \{h, \tau\}$ and $\jmath = \{0, 1\}$.

5.3.1 FA Probability

Lemma 11 *For a given CA threshold δ_h and $h_{n_{RB}}(k-1)$, $P_h^{H_0}$ can be evaluated as*

$$P_h^{H_0} = \exp\left(-\frac{\delta_h}{2(1-\alpha)\sigma_{h_{AR}}^2|h_{n_{RB}}(k-1)|^2 + 2\sigma_w^2}\right). \tag{5.20}$$

where $\alpha = \alpha_{AR}\alpha_{RB} \leq 1$ is the equivalent auto-correlation coefficient of $h_{n_{AB}}$.

Proof 16 *Under H_0, the newly received signal at B is regarded to be from A, i.e., $X = A$. Based on (5.3), we first explore the CA evolution of cascaded multipath channels with the time-series model [87]. Time-series model of $h_{n_{AB}}$ can be written as*

$$h_{n_{AB}}(k) = \alpha h_{n_{AB}}(k-1) + \sqrt{1-\alpha^2} h_{n_{RB}}(k-1) u_{n_{AR}}(k). \tag{5.21}$$

Let $\Delta_{h_{A,A}}$ denote the difference between $\hat{h}_{n_{AB}}(k)$ and $\hat{h}_{n_{AB}}(k-1)$,

$$\begin{aligned} \Delta h_{A,A} &\triangleq \hat{h}_{n_{AB}}(k) - \hat{h}_{n_{AB}}(k-1) \\ &= (\alpha - 1) h_{n_{AR}}(k-1) h_{n_{RB}}(k-1) + \sqrt{1-\alpha^2} h_{n_{RB}}(k-1) u_{n_{AR}}(k) \\ &\quad + w_h(k) - w_h(k-1). \end{aligned} \tag{5.22}$$

Since the $h_{n_{AR}}(k-1) h_{n_{RB}}(k-1)$ term in (5.22) follows complex double Gaussian distribution under Rayleigh fading model [88], it is difficult (if not possible) to analytically model $\Delta h_{A,A}$. It is interesting to see that for a given $h_{n_{RB}}(k-1)$, $h_{n_{AR}}(k-1) h_{n_{RB}}(k-1)$ follows complex Gaussian distribution. In the concerned network scenario, B extracts $h_{n_{RB}}(k-1)$ by exploiting the blind channel estimation techniques widely adapted in previous studies [89]. Therefore, for a given $h_{n_{RB}}(k-1)$, $\Delta h_{A,A}$ is also a complex Gaussian distributed random variable with zero mean and variance $\sigma^2_{\Delta h_{A,A}} = 2(1-\alpha)\sigma^2_{h_{AR}} |h_{n_{RB}}(k-1)|^2 + 2\sigma^2_w$. We can see that $|\Delta h_{A,A}|^2$ follows exponential distribution and its cumulative distribution function (CDF) is given by

$$F_{|\Delta h_{A,A}|^2}(x) = 1 - \exp\left(-\frac{x}{2(1-\alpha)\sigma^2_{h_{AR}} |h_{n_{RB}}(k-1)|^2 + 2\sigma^2_w} \right). \tag{5.23}$$

According to (5.17), $P_h^{H_0}$ *can be determined as*

$$P_h^{H_0} \triangleq \mathbf{Pr}(Q_h[|\hat{h}_{n_{AB}}(k) - \hat{h}_{n_{AB}}(k-1)|^2] = 1 \mid H_0) = 1 - \mathbf{Pr}(|\Delta_{h_{A,A}}|^2 \leq \delta_h). \quad (5.24)$$

Substituting (5.23) *into* (5.24), *we can obtain* (5.20).

Lemma 12 *For a given time threshold* δ_τ, $P_\tau^{H_0}$ *can be evaluated as*

$$P_\tau^{H_0} = \frac{1}{1 - \frac{\eta_{AR}^2}{\eta_{RB}^2}} \left(\frac{\eta_{AR}^2}{\eta_{RB}^2} \exp\left(-\frac{\delta_\tau}{\eta_{AR}} \right) - \exp\left(-\frac{\delta_\tau}{\eta_{RB}} \right) \right), \quad (5.25)$$

where

$$\eta_{AR} = \sqrt{4\sigma_{\tau_{AR}}^4 (1 - \beta_{AR}^2) + 4\sigma_{\tau_{AR}}^2 \sigma_w^2 + \sigma_w^4}, \quad (5.26a)$$

$$\eta_{RB} = \sqrt{4\sigma_{\tau_{RB}}^4 (1 - \beta_{RB}^2) + 4\sigma_{\tau_{RB}}^2 \sigma_w^2 + \sigma_w^4}. \quad (5.26b)$$

Proof 17 *See Appendix B.1 for the proof.*

We use P_F to represent the probability of FA. Then, we can establish the following main results on P_F based on Lemma 11 and 12.

Theorem V.1 P_F *of the proposed scheme in a dual-hop wireless network with an untrusted relay is given in* (5.27).

Proof 18 *Based on* (5.19) *and the law of total probability formula,* P_F *can be given*

$$P_F = \begin{cases} \sum\limits_{z=Z+1}^{N-1} \sum\limits_{z_1=0}^{z} \binom{N}{z_1} (P_h^{H_0})^{z_1} (1-P_h^{H_0})^{N-z_1} \binom{N-1}{z-z_1} (P_\tau^{H_0})^{z-z_1} (1-P_\tau^{H_0})^{N-1-z+z_1} \\ + \sum\limits_{z=N}^{2N-1} \sum\limits_{z_1=z-N}^{N-1} \binom{N}{z-z_1} (P_h^{H_0})^{z-z_1} (1-P_h^{H_0})^{N-z+z_1} \binom{N-1}{z_1} (P_\tau^{H_0})^{z_1} (1-P_\tau^{H_0})^{N-1-z_1}, \\ Z \in [0, N-1], \end{cases} \tag{5.27a}$$

$$\begin{cases} \sum\limits_{z=Z+1}^{2N-1} \sum\limits_{z_1=z-N}^{N-1} \binom{N}{z-z_1} (P_h^{H_0})^{z-z_1} (1-P_h^{H_0})^{N-z+z_1} \binom{N-1}{z_1} (P_\tau^{H_0})^{z_1} (1-P_\tau^{H_0})^{N-1-z_1}, \\ Z \in [N, 2N-1]. \end{cases} \tag{5.27b}$$

by

$$\begin{aligned} P_F &= \mathbf{Pr}(D > Z | H_0) \\ &= \mathbf{Pr}(D = Z+1, Z+2, \cdots, 2N-1 | H_0) \\ &= \sum_{z=Z+1}^{2N-1} \mathbf{Pr}(D=z) = \sum_{z=Z+1}^{N-1} \mathbf{Pr}(D=z) + \sum_{z=N}^{2N-1} \mathbf{Pr}(D=z). \end{aligned} \tag{5.28}$$

Therefore, P_F can be derived based on the following two cases: When $Z \in [0, N-1]$, substituting (4.25a) *into* (5.28)*, P_F can be given in* (5.27a)*; when $Z \in [N, 2N-1]$, substituting* (4.25b) *into* (5.28)*, we can obtain* (5.27b)*.*

Corollary 3 *In a dual-hop wireless network with an untrusted relay, P_F of the proposed scheme utilizing the location-specific of CA separately to discriminate transmitters, can be given by*

$$P_F = \mathbf{Pr}(D > Z | H_0) = \sum_{z=Z+1}^{N} \mathbf{Pr}(D=z) = \sum_{z=Z+1}^{N} \binom{N}{z} (P_h^{H_0})^z (1-P_h^{H_0})^{N-z}. \tag{5.29}$$

5.3.2 MD Probability

Under H_1, the current transmitter is regarded as E, i.e., $X = E$. The following lemmas are dedicated to regarding the probability denoted by $P_{\imath}^{H_1}$, $\imath = \{h, \tau\}$.

Lemma 13 *For a given CA threshold δ_h and $h_{n_{RB}}(k-1)$, $P_h^{H_1}$ can be evaluated as*

$$P_h^{H_1} = \exp\left(-\frac{\delta_h}{\sigma^2_{\Delta h_{E,A}}}\right), \tag{5.30}$$

where $\sigma^2_{\Delta h_{E,A}} = (\alpha^2_{RB}\sigma^2_{h_{ER}} + \sigma^2_{h_{AR}})|h_{n_{RB}}(k-1)|^2 + (1-\alpha^2_{RB})\sigma^2_{h_{ER}}\sigma^2_{h_{RB}} + 2\sigma^2_w$.

Proof 19 *See Appendix B.2 for the proof.*

Lemma 14 *For a given time threshold δ_τ, $P_\tau^{H_1}$ can be evaluated as*

$$P_\tau^{H_1} = \frac{1}{\left(1 - \frac{\eta^2_{EA}}{\eta^2_{RB}}\right)}\left(\frac{\eta^2_{EA}}{\eta^2_{RB}}e^{-\frac{\delta_\tau}{\eta_{EA}}} - e^{-\frac{\delta_\tau}{\eta_{RB}}}\right), \tag{5.31}$$

where $\eta_{EA} = \sigma^2_{\tau_{ER}} + \sigma^2_{\tau_{AR}} + \sigma^2_w$.

Proof 20 *See Appendix B.3 for the proof.*

We use P_M to represent the probability of missed detection. Then, we can establish the following main results on P_M based on Lemma 13 and 14.

Theorem V.2 *P_M of the proposed scheme in a dual-hop wireless network with an untrusted relay is given in* (5.32).

Proof 21 *Based on* (5.19), *P_M is expressed as*

$$P_M = \mathbf{Pr}(D \leq Z|H_1) = \sum_{z=0}^{Z}\mathbf{Pr}(D=z|H_1) = \sum_{z=0}^{N-1}\mathbf{Pr}(D=z) + \sum_{z=N}^{Z}\mathbf{Pr}(D=z). \tag{5.33}$$

$$P_M = \begin{cases} \sum_{z=0}^{Z}\sum_{z_1=0}^{z} \binom{N}{z_1}(P_h^{H_1})^{z_1}(1-P_h^{H_1})^{N-z_1}\binom{N-1}{z-z_1}(P_\tau^{H_1})^{z-z_1}(1-P_\tau^{H_1})^{N-1-z+z_1}, \\ Z \in [0, N-1], \end{cases} \tag{5.32a}$$

$$\begin{cases} \sum_{z=0}^{N-1}\sum_{z_1=0}^{z} \binom{N}{z_1}(P_h^{H_1})^{z_1}(1-P_h^{H_1})^{N-z_1}\binom{N-1}{z-z_1}(P_\tau^{H_1})^{z-z_1}(1-P_\tau^{H_1})^{N-1-z+z_1} \\ +\sum_{z=N}^{Z-1}\sum_{z_1=z-N}^{N-1} \binom{N}{z-z_1}(P_h^{H_1})^{z-z_1}(1-P_h^{H_1})^{N-z+z_1}\binom{N-1}{z_1}(P_\tau^{H_1})^{z_1}(1-P_\tau^{H_1})^{N-1-z_1}, \\ Z \in [N, 2N-1]. \end{cases} \tag{5.32b}$$

Under H_1, P_M can be derived based on the following two cases: When $Z \in [0, N-1]$, using $P_h^{H_1}$ and $P_\tau^{H_1}$ to replace $P_h^{H_0}$ and $P_\tau^{H_0}$ in (4.25a)*, respectively, and we substitute the new resulting into* (5.33)*. Then, P_M can be given by* (5.32a)*; when $Z \in [N, 2N-1]$, following a similar process, P_M can be given by* (5.32b)*.*

Corollary 4 *In a dual-hop wireless network with an untrusted relay, P_M of the proposed scheme utilizing the location-specific of CA separately to discriminate transmitters, can be given by*

$$P_M = \mathbf{Pr}(D \le Z|H_1) = \sum_{z=0}^{N} \mathbf{Pr}(D=z) = \sum_{z=0}^{N} \binom{N}{z}(P_h^{H_1})^z(1-P_h^{H_1})^{N-1-z}. \tag{5.34}$$

5.4 Numerical Results

5.4.1 System Parameters and Simulation Settings

We list in Table 5.1 the main system parameters that determine the authentication performance in terms of P_F and P_M. We use $\bar{\gamma}$ to denote the average signal-to-noise ratio (SNR) per hop and use κ_h to denote the ratio of average CA gains for A-R and E-R. Since $\sigma_w^2 = P_w/P$, the variance of noise σ_w^2 in the equation of $\bar{\gamma}$ is inversely proportional to the average transmit power P for a given P_w. In our simulation, we

Table 5.1: Main system parameters affecting performance

Parameter	**Description**
$\bar{\gamma}_{AR} = \frac{\sigma^2_{h_{AR}}}{\sigma^2_w}$	The average SNR of the first hop
$\bar{\gamma}_{RB} = \frac{\sigma^2_{h_{RB}}}{\sigma^2_w}$	The average SNR of the second hop
$\kappa_h = \frac{\sigma^2_{h_{ER}}}{\sigma^2_{h_{AR}}}$	The ratio of averaged channel gains for A-R and E-R
α_{AR}, α_{RB}	Channel correlation coefficients
β_{AR}, β_{RB}	Delay interval correlation coefficients
δ_h, δ_τ, Z	CA threshold, time threshold, decision threshold

set $\sigma^2_{h_{AR}} = \sigma^2_{h_{RB}} = \sigma^2_{\tau_{AR}} = \sigma^2_{\tau_{RB}} = 1$, and then adjust the parameters $\sigma^2_{h_{ER}}$ and $\sigma^2_{\tau_{ER}}$ to achieve a specified κ_h.

Table 5.2: Three fading scenarios

	Channel status	$f_{AR} = f_{RB} = f$
Case I	Slow-fading	.001
Case II	Fast-fading	.10
Case III	Faster-fading	.15

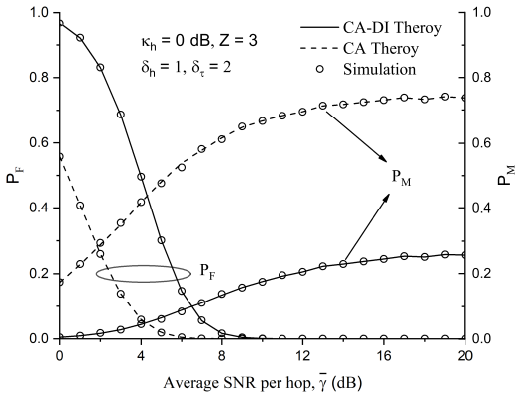


Figure 5.5: The authentication performance (P_F, P_M) for the proposed scheme based on CA-DI or CA vs. average SNR per hop ($\bar{\gamma}_{AR} = \bar{\gamma}_{RB} = \bar{\gamma}$) under slow-fading channels.

To validate the theoretical modeling, a dedicated simulator in MATLAB is developed, which is now available at [90]. The time-varying channels are generated by

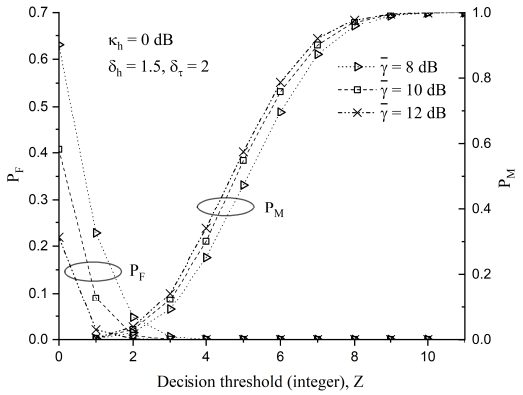
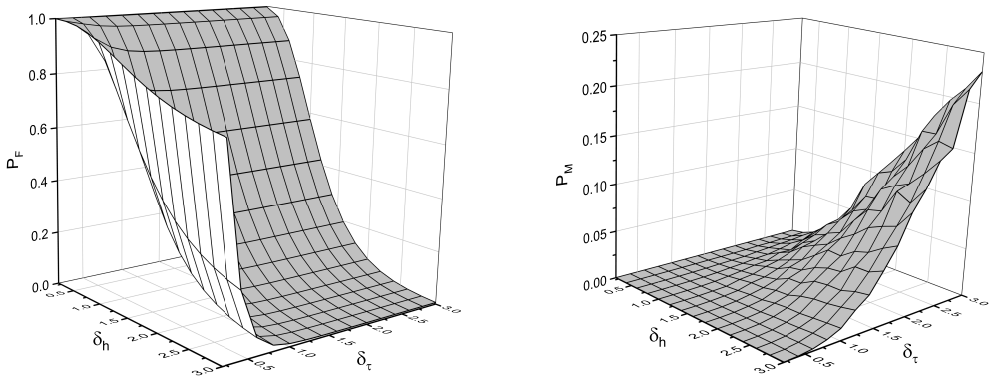


Figure 5.6: Effect of average SNR per hop ($\bar{\gamma}_{AR} = \bar{\gamma}_{RB} = \bar{\gamma}$) on the authentication performance (P_F, P_M) vs. decision threshold Z under slow-fading channels.

using the method introduced in [91]. The temporal correlation of the time-varying channels is a function of the normalized Doppler frequency that is affected by the mobility, carrier frequency and symbol duration. Based on the normalized Doppler frequencies illustrated in Table 5.2, we consider three fading channels associated with correlation coefficients (i.e., α_{AR}, α_{RB}, $\alpha = \alpha_{AR}\alpha_{RB}$, β_{AR}, β_{RB}). Different normalized Doppler frequency values in Table 5.2 correspond to different moving velocity of entity under a given carrier frequency and symbol duration[58]. For a fair comparison, the number of multipath components for A-R and E-R are both fixed to be 3, while that of multipath components for R-B is fixed to be 2. For Monte-Carlo experiments, 10^5 independent trails are conducted to obtain the average results.

5.4.2 Model Validation

To verify the theoretical results, we summarize in Fig. 5.5 both the simulation and theoretical models of P_F and P_M for the proposed scheme based on CA-DI or CA, where the slow-fading channels and settings of ($Z = 3$, $\kappa_h = 0$ dB, $\delta_h = 1$, $\delta_\tau = 2$) are assumed. Fig. 5.5 shows clearly that the simulation results agree well with the theoretical ones, confirming that our theoretical models can be used to nicely characterize P_F and P_M. It can also be seen from Fig. 5.5 that as the average SNR per



(a) P_F vs. (δ_h, δ_τ) (b) P_M vs. (δ_h, δ_τ)

Figure 5.7: P_F and P_M vs. (δ_h, δ_τ) when $Z = 1$, $\bar{\gamma} = 10$ dB and $\kappa_h = 0$ dB under slow-fading channels.

hop $\bar{\gamma}$ increases, P_F for the proposed scheme based on CA-DI or CA declines rapidly, whereas the corresponding P_M increases slowly. This reveals that a trade-off exists between reliability and security in terms of P_F and P_M. Moreover, Fig. 5.5 shows that when $\bar{\gamma}$ is small, the proposed scheme jointly using CA and DI outperforms that only using CA in terms of P_M. While for P_F, the result is reverse. This is due to the fact that by jointly utilizing the location-specific properties of cascaded channel in terms of CA and DI, the proposed scheme can effectively detect impersonation attacks at the cost of incurring more false alarm events. In addition, it can be observed from Fig. 5.5 that when $\bar{\gamma}$ is large (e.g., $\bar{\gamma} \geq 12$dB), P_F for the proposed scheme based on either CA-DI or CA approaches 0, but the scheme jointly using CA and DI still outperforms that only using CA in terms of P_M.

5.4.3 Control of FA and MD Probabilities

To demonstrate that the proposed scheme enables the authentication performance to be flexibly controlled in a large region, we now explore how P_F and P_M vary with parameters Z, δ_h and δ_τ. Fig. 5.6 shows the effect of $\bar{\gamma}$ on the P_F and P_M versus Z, where the slow-fading channels and settings of ($\kappa_h = 0$ dB, $\delta_h = 1.5$, $\delta_\tau = 2$) are assumed. As observed from Fig. 5.6, P_M increases rapidly as Z increases,

while P_F declines quickly with Z. It is interesting to notice that P_F is extremely sensitive to the variations of Z. For example, when $Z \geq 4$, P_F approaches 0 under $\bar{\gamma} = \{8 \text{ dB}, 10 \text{ dB}, 12 \text{ dB}\}$. We can see from Fig. 5.6 that for a fixed Z, P_F decreases with $\bar{\gamma}$ while P_M increases with $\bar{\gamma}$. Noticed that for an authentication system, both P_F and P_M are in general required to be below 0.1 [12, 14, 31, 32]. Therefore, for the specified P_F and P_M constraints (e.g., P_F, $P_M \leq 0.1$), we can increase the transmit power and find an optimal setting of Z. For example, we can set $Z = 1$ and $\bar{\gamma} \geq 10$ dB to ensure P_F, $P_M \leq 0.1$. It is notable, however, that for general wireless networks applications, the total power of system is limited to a certain level due to the energy constraint and interference requirement among simultaneous transmissions, so it is of great significance to find the optimal setting of parameters δ_h and δ_τ to satisfy the specified P_F and P_M constraints under a given Z and power limitation.

Fig. 5.7 shows how P_F and P_M vary with parameters (δ_h, δ_τ) under the slow-fading channels and settings of ($Z = 1$, $\bar{\gamma}$= 10 dB, $\kappa_h = 0$ dB). As shown in Fig. 5.7(a) (resp. Fig. 5.7(b)) that for a specified constraint p_f of P_F (resp. a specified constraint p_m of P_M), we can accordingly set a specified constraint plane intersecting the z-axis orthogonally at the point (2, 2, p_f) (resp. at the point (2, 2, p_m)), and then can determine a set of (δ_h, δ_τ)-pairs corresponding to the surface below the defined constraint plane. By finding the intersection of these two sets of (δ_h, δ_τ)-pairs, we can obtain the region of (δ_h, δ_τ)-pairs to achieve the p_f and p_m constraints in terms of P_F and P_M. For example, when δ_h, $\delta_\tau \in [0.2, 3]$, one can observe from Fig. 5.7 that the requirement of $P_F \leq 0.1$ can be achieved in the regions of ($\delta_h \in [1, 3]$, $\delta_\tau \in [1.2, 3]$), while the requirement of $P_M \leq 0.1$ is achieved in the regions of ($\delta_h \in [0.2, 2.5]$, $\delta_\tau \in [0.2, 2.2]$). Thus, the constraints of (P_F, $P_M \leq 0.1$) under the considered network scenario are achieved when $\delta_h \in [1, 2.5]$ and $\delta_\tau \in [1.2, 2.2]$.

Fig. 5.6 and Fig. 5.7 indicate that the proposed scheme is flexible and general, since P_F and P_M can be flexibly controlled by a proper setting of the decision threshold

Z, CA threshold δ_h, and time threshold δ_τ. Also, a trade-off between reliability and security can be controlled by an appropriate setting of δ_h and δ_τ.

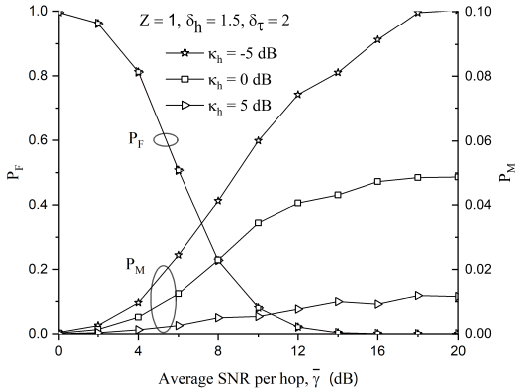


Figure 5.8: Effect of κ_h on the authentication performance (P_F, P_M) vs. average SNR per hop ($\bar{\gamma}_{AR} = \bar{\gamma}_{RB} = \bar{\gamma}$) under slow-fading channels.

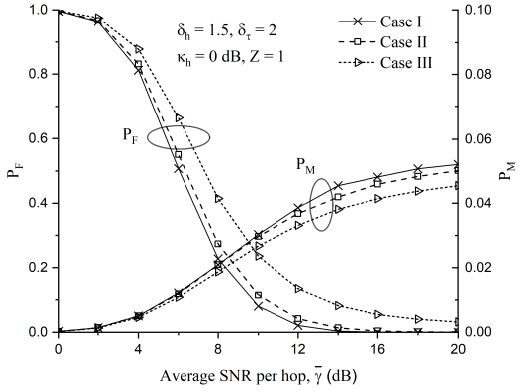


Figure 5.9: Effect of channel status on the authentication performance (P_F, P_M) vs. average SNR per hop ($\bar{\gamma}_{AR} = \bar{\gamma}_{RB} = \bar{\gamma}$).

5.4.4 Authentication Efficiency Analysis

To illustrate the authentication efficiency of the proposed scheme, we further explore the authentication performance of the scheme under diverse network scenarios with different positions of E and different channels. Fig. 5.8 shows the effect of the position of E on the authentication performance (P_F, P_M) versus $\bar{\gamma}$ under slow-fading channels, where Z=1, $\delta_h = 1.5$, $\delta_\tau = 2$, and κ_h varies from -5 dB to 5 dB. We can

see from Fig. 5.8 that P_F is not affected by the variations of κ_h. This is because the position of E is not related to FA events. It is noticed that for a fixed $\bar{\gamma}$, P_M increases as κ_h reduces. We also find that when $\bar{\gamma}$ is small, the setting of κ_h has a little impact on P_M; when $\bar{\gamma}$ is large, the setting of κ_h has a significant impact on P_M, especially when κ_h=-5 dB, P_M approaches 0.1. It implies that a "smart" intruder E would seek a position where it could have a higher probability to impersonate attacks successfully. Fortunately, the proposed scheme can adjust dynamically the optimal setting of (Z, δ_h, δ_τ) to resist against such attacks. This indicates that for various positions of E, the proposed scheme is efficient to discriminate transmitters through the proper settings of (Z, δ_h, δ_τ).

Finally, Fig. 5.9 demonstrates the effect of channel status on the authentication performance (P_F, P_M) versus $\bar{\gamma}$, given that $Z = 1$, $\kappa_h = 0$ dB, $\delta_h = 1.5$, $\delta_\tau = 2$. Notice also that in terms of P_F, the proposed scheme under case I (slow-fading channels) outperforms that under the others (case II and case III), and the scheme under case III (faster-fading channels) leads to the highest P_F. This indicates that excessive false alarm events will happen when channels are faster-fading (e.g. in a higher mobile scenario). However, in terms of P_M, the proposed scheme under faster-fading channels has the lowest P_M, which means that the faster-fading channels are beneficial for effectively detecting impersonation attacks. Fig. 5.9 confirms that under various channels, the proposed scheme is also efficient in identifying transmitters by a proper setting of the decision threshold Z, CA threshold δ_h, time threshold δ_τ and average SNR per hop $\bar{\gamma}$.

5.5 Summary

This chapter represents an attempt to explore the E2E authentication issue for dual-hop wireless networks by exploiting the intrinsic properties of cascaded multi-path channels. We showed that the proposed E2E authentication scheme is not only

efficient in discrimination between the legitimate and illegitimate transmitters, but also resistant to impersonating attacks with the attacker near the legitimate transmitter and resistant to replay attacks with aggressive signals from the untrusted relay. We also proved that the proposed scheme is flexible and general in the sense that this scheme makes it possible for us to flexibly control FA and MD probabilities in a large region through the proper settings of parameters. This is an important property for wireless networks to support various applications with different authentication performance requirements. In addition, it is expected that the proposed authentication scheme and related theoretical models will be useful for providing a guideline to devise the coping strategies under various attacks, as well as for understanding the fundamental E2E authentication performance of multi-hop wireless networks.

CHAPTER VI

Conclusion

In this thesis, we studied physical layer authentication for wireless communications, where intrinsic and unique features of physical layer are explored to authenticate transmitters for wireless communications. We first explored channel-based authentication solution for massive MIMO systems with hardware impairments, and then investigated authentication solution which exploits two physical layer features in terms of location-specific wireless channels and transmitter-specific hardware impairments. Finally, we examined E2E channel-based authentication issue in dual-hop wireless networks with untrusted relays.

For channel-based authentication solution with the consideration of hardware impairments, we studied in Chapter III an uplink massive MIMO system consisting of three different entities with hardware impairments. We proposed a channel-based authentication scheme for massive MIMO systems with different levels of hardware impairments. We theoretically analyzed FA and SD probabilities. Our results show that the performance is clearly deteriorated by hardware impairments, with a non-trivial impact from the choice of antenna patterns. Notice that multiple hardware impairments (such as I/Q imbalance and phase noise) can be effectively utilized to authenticate transmitters, which is demonstrated in the literature. While in this chapter their effects have been taken into account by using κ-parameters.

For authentication solution exploiting two physical layer features in terms of location-specific wireless channels and transmitter-specific hardware impairments, we investigated in Chapter IV a MIMO system consisting of three different entities: Alice and Eve with N_t antennas and Bob with N_r antennas. We proposed a physical layer authentication scheme jointly utilizing channel gains and phase noise in heterogeneous MIMO systems. We also determined the variances of estimation errors in terms of channel gains and phase noise, and then derived closed-form expressions for false alarm and missed detection probabilities with the consideration of quantization errors. We further demonstrated that the proposed scheme enables flexible performance control by adjusting thresholds (for channel gain, phase noise, and decision, respectively). This indicates that the proposed scheme has the capability of satisfying different performance requirements for future emerging heterogeneous MIMO systems.

In Chapter V, we addressed E2E authentication issue in dual-hop wireless networks with untrusted relays. We first proposed an E2E channel-based authentication scheme. We showed that the proposed scheme is efficient in discrimination between the legitimate and illegitimate transmitters as well as resistant to impersonating attacks with the attacker near the legitimate transmitter and resistant to replay attacks with aggressive signals from the untrusted relay. We also proved that the proposed scheme is flexible and general in the sense that this scheme makes it possible for us to flexibly control FA and MD probabilities in a large region through the proper settings of parameters. This is an important characteristic for wireless networks to support various applications with different authentication performance requirements. In addition, it is expected that the proposed authentication scheme and related theoretical models will be useful for providing a guideline to devise the coping strategies under various attacks, as well as for understanding the fundamental E2E authentication performance of multi-hop wireless networks.

It is notable that, this thesis considers relatively simple physical layer features such as wireless channels (e.g., channel gains and multi-path delay) and hardware impairments (e.g., phase noise). In practice, however, considering multi-dimensional features (e.g., channel impulse response, channel frequency response, I/Q imbalance, carrier frequency offset, and phase noise) for authentication is an interesting research topic for our future work to further explore how these physical layer features can be jointly used to improve the security of wireless communications.

In addition, for 5G and IoT networks with low latency requirements, a more effective design for fast authentication approach should be investigated. This issue offers us an interesting future research direction and diverse further research in our future work.

APPENDICES

APPENDIX A

Proofs in Chapter IV

A.1 Proof of Lemma 2

Based on Lemma 1, we can explore the distribution of $\mathbf{x}$ on the two hypotheses. Using (3.1) on H_0, we have

$$\begin{aligned}\mathbf{x} &= \mathbf{h}_A[k] - \mathbf{h}_A[k-1] + \boldsymbol{\epsilon}_A[k-1] - \boldsymbol{\epsilon}_A[k] \\ &= (\alpha-1)\mathbf{h}_A[k-1] + \sqrt{1-\alpha^2}\mathbf{e}_A[k] + \boldsymbol{\epsilon}_A[k-1] - \boldsymbol{\epsilon}_A[k]. \end{aligned} \tag{A.1}$$

From (A.1), we can see that $\mathbf{x}$ is a zero-mean complex Gaussian random vector. This is because $\mathbf{h}_A$, $\mathbf{e}_A$ and $\boldsymbol{\epsilon}_A$ are mutually independent zero-mean complex Gaussian random vectors. Using (3.6), $\mathbf{C}_0$ is determined as (3.10a). Similarly, $\mathbf{x}$ on H_1 can be written as

$$\mathbf{x} = \mathbf{h}_E[k] - \mathbf{h}_A[k-1] + \boldsymbol{\epsilon}_A[k-1] - \boldsymbol{\epsilon}_E[k]. \tag{A.2}$$

Since $\mathbf{h}_E$, $\mathbf{h}_A$, $\boldsymbol{\epsilon}_A$ and $\boldsymbol{\epsilon}_E$ are mutually independent zero-mean complex Gaussian random vectors, $\mathbf{x}$ on H_1 is also a zero-mean complex Gaussian random vector. Based

on (3.6), $\mathbf{C}_1$ is determined as (3.10b). We can see from (3.10a) and (3.10b) that $\mathbf{C}_1$ can be decomposed as

$$\mathbf{C}_1 = \mathbf{C}_0 + \mathbf{K}. \tag{A.3}$$

where $\mathbf{K}$ is given by (3.11).

For simplicity, we define the inverse of $\mathbf{C}_i$ as $\mathbf{Q}_i$, that is, $\mathbf{Q}_i \triangleq \mathbf{C}_i^{-1}$. Note that both $\mathbf{R}_A$ and $\mathbf{R}_E$ are nonsingular due to the assumption of complex Gaussian random channel vector, so $\mathbf{C}_0$, $\mathbf{C}_1$, and $\mathbf{K}$ are also nonsingular. Therefore, we can always find $\mathbf{Q}_0$ and $\mathbf{Q}_1$. Let $\Delta\mathbf{Q} = \mathbf{Q}_0 - \mathbf{Q}_1$, which can be further written by applying matrix inversion lemma stated in [55, Lemma 2.3] as $\Delta\mathbf{Q} = \mathbf{C}_0^{-1}\mathbf{K}\mathbf{C}_1^{-1}$.

Using (3.10a) and (3.10b), the probability density functions (PDFs) of $\mathbf{x}$ on the two hypotheses can be written as

$$f(\mathbf{x}|\mathcal{H}_i) = \frac{1}{\pi^M \det(\mathbf{C}_i)} \exp(-\mathbf{x}^H \mathbf{Q}_i^{-1} \mathbf{x}),\ i = 0, 1. \tag{A.4}$$

We use $\mathcal{L}_0(\mathbf{x})$ to denote a LRT and δ_0 a threshold. Neyman–Pearson Criterion in [59] leads us to a LRT, which can be written as

$$\mathcal{L}_0(\mathbf{x}) \triangleq \frac{f(\mathbf{x}|H_1)}{f(\mathbf{x}|H_0)} = \frac{\det(\mathbf{C}_0)}{\det(\mathbf{C}_1)} \frac{\exp(-\mathbf{x}^H \mathbf{Q}_1 \mathbf{x})}{\exp(-\mathbf{x}^H \mathbf{Q}_0 \mathbf{x})} \underset{H_0}{\overset{H_1}{\gtrless}} \delta_0. \tag{A.5}$$

Taking logarithms and retaining only data-dependent terms, we can obtain logarithmic LRT as (3.8).

A.2 Proof of Theorem III.1

Proof of Theorem III.1 for IID: Combining (3.9) and (3.10), we can obtain the LRT in (3.8) under IID case as

$$\mathcal{L}(\mathbf{x}) \triangleq \frac{\lambda_{\mathbf{D}}}{\lambda_{\mathbf{C}_0}(\lambda_{\mathbf{C}_0}+\lambda_{\mathbf{D}})} \sum_{m=1}^{M} |x_m|^2 \underset{H_0}{\overset{H_1}{\gtrless}} \delta. \tag{A.6}$$

Based on the above results, we now derive expressions for P_F and P_D under IID case. Since $x_m/\sqrt{\lambda_{\mathbf{C}_0}}$ on H_0 is independent zero mean complex Gaussian variable with variance 1, $\sum_{m=1}^{M} |x_m/\sqrt{\lambda_{\mathbf{C}_0}}|^2$ is a chi-square random variable with $2M$ degrees of freedom, that is, $\sum_{m=1}^{M} |x_m/\sqrt{\lambda_{\mathbf{C}_0}}|^2 \sim \chi^2_{2M}$. P_F under IID can be given by

$$P_F = \mathbf{Pr}(\mathcal{L}(\mathbf{x}) > \delta | H_0) = \mathbf{Pr}\left(\sum_{m=1}^{M} \left| \frac{x_m}{\sqrt{\lambda_{\mathbf{C}_0}}} \right|^2 > \left(\frac{\lambda_{\mathbf{C}_0}}{\lambda_{\mathbf{D}}} + 1 \right) \delta | H_0 \right). \tag{A.7}$$

Substituting the right-tail probability function of chi-square random variable into (A.7) yields (3.19a) under IID case.

Following the same steps, we can obtain P_D under IID case as

$$P_D = \mathbf{Pr}(\mathcal{L}(\mathbf{x}) > \delta | H_1). \tag{A.8}$$

Substituting the right-tail probability function of chi-square random variable into (A.8) yields (3.20a) under IID case.

Proof of Theorem III.1 for IUV: Combining (3.9) and (3.17), we can obtain the LRT in (3.8) under IUV case as

$$\mathcal{L}(\mathbf{x}) \triangleq \frac{\lambda_{\mathbf{D}_m}}{\lambda_{\mathbf{C}_0,m}(\lambda_{\mathbf{C}_0,m}+\lambda_{\mathbf{D}_m})} \sum_{m=1}^{M} |x_m|^2 \underset{H_0}{\overset{H_1}{\gtrless}} \delta. \tag{A.9}$$

Under H_0, the characteristic function of $|x_m|^2$ is

$$\begin{aligned} M_{|x_m|^2|H_0}(j\omega) &= \mathbb{E}\{\exp(j\omega|x_m|^2)|H_0\} \\ &= \int_{-\infty}^{\infty} \frac{\exp\left((j\omega - \frac{1}{\lambda_{\mathbf{C}_0,m}})|x_m|^2\right)}{\pi\lambda_{\mathbf{C}_0,m}} dx_m \\ &= (1 - j\omega\lambda_{\mathbf{C}_0,m})^{-1}. \end{aligned} \tag{A.10}$$

Let $a_m = \frac{\lambda_{\mathbf{K},m}}{\lambda_{\mathbf{C}_0,m}+\lambda_{\mathbf{K},m}}$. Thus, we can obtain the characteristic function of $\mathcal{L}(\mathbf{x})$ on H_0 as

$$M_{\mathcal{L}(\mathbf{x})|H_0}(j\omega) = \prod_{m=1}^{M} (1 - j\omega a_m)^{-1}. \tag{A.11}$$

We use a partial fraction expansion [92] of (A.11) to obtain

$$M_{\mathcal{L}(\mathbf{x})|H_0}(j\omega) = \sum_{m=1}^{M} b_m (1 - j\omega a_m)^{-1}, \tag{A.12}$$

where

$$b_m = \prod_{\substack{i=1 \\ i\neq m}}^{M} \frac{a_m}{a_m - a_i}. \tag{A.13}$$

As observed from (A.12), the characteristic function of $\mathcal{L}(\mathbf{x})$ is a weighted superposition of exponentially distributed characteristic functions. After taking inverse Fourier transform for (A.12), we can see that the PDF of $\mathcal{L}(\mathbf{x})$ is also a weighted superposition of exponentially distributed PDFs [59], that is,

$$f(\mathcal{L}(\mathbf{x})|H_0) = \sum_{m=1}^{M} \frac{b_m}{a_m} \exp\left(-\frac{\mathcal{L}}{a_m}\right). \tag{A.14}$$

P_F under IUV case can be obtained by integrating the following formula as

$$P_F = \mathbf{Pr}(\mathcal{L}(\mathbf{x}) > \delta | H_0) = \int_{\delta}^{\infty} f(\mathcal{L}(\mathbf{x})|H_0) d\mathcal{L}(\mathbf{x}). \tag{A.15}$$

Then, substituting (A.12) into (A.15) yields (3.19b) under IUV.

Similarly, the characteristic function of $\mathcal{L}(\mathbf{x})$ on H_1 is

$$M_{\mathcal{L}(\mathbf{x})|H_1}(j\omega) = \prod_{m=1}^{M} (1 - j\omega c_m)^{-1}. \tag{A.16}$$

Following the same steps, we can obtain P_D (3.20b) by integrating the following formula (A.17).

$$P_D = \mathbf{Pr}(\mathcal{L}(\mathbf{x}) > \delta | H_1) = \int_{\delta}^{\infty} f(\mathcal{L}(\mathbf{x})|H_1) d\mathcal{L}(\mathbf{x}). \tag{A.17}$$

A.3 Proof of Theorem III.2

When $|\rho_A| = |\rho_E| = 1$, the characteristic functions of $\mathcal{L}(\mathbf{x})$ on the two hypotheses can be obtained by using (3.25) as

$$M_{\mathcal{L}(\mathbf{x})|H_0}(j\omega) = \left(1 - j\omega \frac{\lambda_{2_1}}{\lambda_{0_1} + \lambda_{2_1}}\right)^{-1}, \tag{A.18}$$

$$M_{\mathcal{L}(\mathbf{x})|H_1}(j\omega) = \left(1 - j\omega \frac{\lambda_{2_1}}{\lambda_{0_1}}\right)^{-1}. \tag{A.19}$$

$\mathcal{L}(\mathbf{x})$ is an exponentially distributed random variable. Similar to the derivations of (3.19b) and (3.20b), we can obtain (3.28a) and (3.29a).

When $0 < |\rho_{\mathrm{t}}| < 1$, using the modal matrix $\mathbf{u}_{\mathbf{Dw}}^H$ in (3.27), we transform from $\mathbf{x_w}$ to $\mathbf{x_{wu}} = [x_{\mathbf{wu},1} \cdots x_{\mathbf{wu},M}]^T$, which is denoted by $\mathbf{x_{wu}} = \mathbf{u}_{\mathbf{Dw}}^H \mathbf{x_w} = \mathbf{u}_{\mathbf{Dw}}^H \mathbf{w}^H \mathbf{x}$. Now, we explore the covariance matrices of $\mathbf{x_{wu}}$ on the two hypotheses. On H_0, its covariance

matrix is given by

$$\begin{aligned}\mathrm{Cov}(\mathbf{x_{wu}}|H_0) &= \mathbb{E}\{\mathbf{x_{wu}}\mathbf{x}_{\mathbf{wu}}^H|H_0\} = \mathbb{E}\{\mathbf{u}_{\mathbf{Dw}}^H\mathbf{x_w}\mathbf{x}_{\mathbf{w}}^H\mathbf{u_{Dw}}|H_0\} \\ &= \mathbf{u}_{\mathbf{Dw}}^H\mathbf{I}\mathbf{u_{Dw}} = \mathbf{I}. \end{aligned} \tag{A.20}$$

Similarly, on H_1 the covariance matrix of $\mathbf{x_{wu}}$ is given by

$$\begin{aligned}\mathrm{Cov}(\mathbf{x_{wu}}|H_1) &= \mathbb{E}\{\mathbf{x_{wu}}\mathbf{x}_{\mathbf{wu}}^H|H_1\} = \mathbf{u}_{\mathbf{Dw}}^H\mathbf{R}_{1\mathbf{w}}\mathbf{u_{Dw}} \\ &= \mathbf{u}_{\mathbf{Dw}}^H\mathbf{u_{Dw}}[\Lambda_{\mathbf{Dw}} + \mathbf{I}]\mathbf{u}_{\mathbf{Dw}}^H\mathbf{u_{Dw}} = \Lambda_{\mathbf{Dw}} + \mathbf{I}. \end{aligned} \tag{A.21}$$

We define a diagonal matrix $\mathbf{Q_{wu}}$ as

$$\mathbf{Q_{wu}} \triangleq \mathrm{diag}\left[\frac{\lambda_{\mathbf{Dw},1}}{\lambda_{\mathbf{Dw},1}+1}, ..., \frac{\lambda_{\mathbf{Dw},M}}{\lambda_{\mathbf{Dw},M}+1}\right] = \mathrm{diag}\left[\lambda_{\mathbf{wu},1}, ..., \lambda_{\mathbf{wu},M}\right]. \tag{A.22}$$

Applying some derivations similar to that in [59, Chapter 3], under spatially correlated channel component case, the LRT $\mathcal{L}(\mathbf{x})$ in (3.8) becomes

$$\mathcal{L}(\mathbf{x}) \triangleq \mathcal{L}(\mathbf{x_{wu}}) = \mathbf{x}_{\mathbf{wu}}^H\mathbf{Q_{wu}}\mathbf{x_{wu}} = \sum_{m=1}^{M}\frac{\lambda_{\mathbf{Dw},m}}{\lambda_{\mathbf{Dw},m}+1}|x_{\mathbf{wu},m}|^2 = \sum_{m=1}^{M}\lambda_{\mathbf{wu},m}|x_{\mathbf{wu},m}|^2 \underset{H_0}{\overset{H_1}{\gtrless}} \delta. \tag{A.23}$$

Note that $\mathbf{x}$ is linearly transformed to $\mathbf{x_{wu}}$, and the effect of this transform is to decorrelate $\mathbf{x}$. Therefore, $\mathbf{x}_{\mathbf{wu},m}$ also follows zero-mean complex Gaussian distribution and thus $|x_{\mathbf{wu},m}|^2$ follows exponential distribution. When $0 < |\rho_{\mathrm{t}}| < 1$, P_F and P_D can be evaluated as

$$P_F = \mathbf{Pr}(\mathcal{L}(\mathbf{x}) > \delta|H_0) = \int_{\delta}^{\infty} f(\mathcal{L}(\mathbf{x})|H_0)d\mathcal{L}(\mathbf{x}) \triangleq \int_{\delta}^{\infty} f(\mathcal{L}(\mathbf{x_{wu}})|H_0)d\mathcal{L}(\mathbf{x_{wu}}), \tag{A.24}$$

$$P_D = \mathbf{Pr}(\mathcal{L}(\mathbf{x}) > \delta | H_1) = \int_{\delta}^{\infty} f(\mathcal{L}(\mathbf{x})|H_1) d\mathcal{L}(\mathbf{x}) \triangleq \int_{\delta}^{\infty} f(\mathcal{L}(\mathbf{x_{wu}})|H_1) d\mathcal{L}(\mathbf{x_{wu}}). \tag{A.25}$$

Following a similar method as that of in Section 3.3.1, we can obtain P_F and P_D as (3.28b) and (3.29b) for $0 < |\rho_{\mathrm{A}}|, |\rho_{\mathrm{E}}| < 1$.

APPENDIX B

Proofs in Chapter V

B.1 Proof of Lemma 12

Let $\Delta\tau_{AR} = \hat{\tau}_{k_{AR}}(t) - \hat{\tau}_{k_{AR}}(t-1)$, $\Delta\tau_{RB} = \hat{\tau}_{k_{RB}}(t) - \hat{\tau}_{k_{RB}}(t-1)$, and $\Delta\tau_{A,A} = \hat{\tau}_{k_{AB}}(t) - \hat{\tau}_{k_{AB}}(t-1)$, then $\Delta\tau_{AB}$ can be further written as

$$\begin{aligned}\Delta\tau_{A,A} &\triangleq \hat{\tau}_{k_{AB}}(t) - \hat{\tau}_{k_{AB}}(t-1) \\ &= \hat{\tau}_{k_{AR}}(t) - \hat{\tau}_{k_{AR}}(t-1) + \hat{\tau}_{k_{RB}}(t) - \hat{\tau}_{k_{RB}}(t-1) \\ &= \Delta\tau_{AR} + \Delta\tau_{RB}. \end{aligned} \tag{B.1}$$

To explore the distribution of random variable $\Delta\tau_{A,A}$, we first examine that of $\Delta\tau_{AR}$. Based on (5.15), $\Delta\tau_{AR}$ can be written as (B.2). Combining (5.7), (5.14) and (5.15), the random variable C_1 defined in (B.2) can be given by

$$\begin{aligned} C_1 &= \hat{\tau}^{(1)}_{k_{AR}}(t) - \hat{\tau}^{(1)}_{k_{AR}}(t-1) \\ &= (\beta_{AR}-1)\tau^{(1)}_{k_{AR}}(t-1) + \sqrt{(1-\beta_{AR}^2)\sigma^2_{\tau_{AR}}}u^{(1)}_{k_{AR}}(t) + w^{(1)}_{\tau}(t) - w^{(1)}_{\tau}(t-1). \end{aligned} \tag{B.3}$$

Since $\tau^{(1)}_{k_{AR}}$, $u^{(1)}_{k_{AR}}$ and $w^{(1)}_{\tau}$ are independent Gaussian distributed random variables

$$\begin{aligned}\Delta\tau_{AR} &= (\hat{\tau}^{(1)}_{k_{AR}}(t))^2 + (\hat{\tau}^{(2)}_{k_{AR}}(t))^2 - (\hat{\tau}^{(1)}_{k_{AR}}(t-1))^2 - (\hat{\tau}^{(2)}_{k_{AR}}(t-1))^2 \\ &= \underbrace{(\hat{\tau}^{(1)}_{k_{AR}}(t) - \hat{\tau}^{(1)}_{k_{AR}}(t-1))}_{C_1}\underbrace{(\hat{\tau}^{(1)}_{k_{AR}}(t) + \hat{\tau}^{(1)}_{k_{AR}}(t-1))}_{C_2} \\ &\quad + \underbrace{(\hat{\tau}^{(2)}_{k_{AR}}(t) - \hat{\tau}^{(2)}_{k_{AR}}(t-1))}_{C_3}\underbrace{(\hat{\tau}^{(2)}_{k_{AR}}(t) + \hat{\tau}^{(2)}_{k_{AR}}(t-1))}_{C_4}. \end{aligned} \tag{B.2}$$

---

with zero mean, C_1 is also a Gaussian distributed random variable with zero mean and variance $\sigma^2_{C_1} = 2\sigma^2_{\tau_{AR}}(1-\beta_{AR}) + \sigma^2_w$. After a similar derivation, C_2, C_3, C_4 are also Gaussian distributed random variables with zero means and variances with $\sigma^2_{C_3} = \sigma^2_{C_1}$, and $\sigma^2_{C_2} = \sigma^2_{C_4} = 2\sigma^2_{\tau_{AR}}(1+\beta_{AR}) + \sigma^2_w$. It is easy to see that C_1, C_2, C_3 and C_4 are independent of each other, and we have

$$\begin{aligned}\Delta\tau_{AR} &= C_1C_2 + C_3C_4 \\ &= \underbrace{\sqrt{4\sigma^4_{\tau_{AR}}(1-\beta^2_{AR}) + 4\sigma^2_{\tau_{AR}}\sigma^2_w + \sigma^4_w}}_{\eta_{AR}}\left(\frac{C_1}{\sigma_{C_1}}\frac{C_2}{\sigma_{C_2}} + \frac{C_3}{\sigma_{C_3}}\frac{C_4}{\sigma_{C_4}}\right)\end{aligned} \tag{B.4}$$

According to [93, Eq.(2.2.13)], $\Delta\tau_{AR}$ follows the Laplace distribution, that is, $\Delta\tau_{AR} \sim Laplace(0, \eta_{AR})$. After a similar derivation, we know that $\Delta\tau_{RB}$ also follows the Laplace distribution $\Delta\tau_{RB} \sim Laplace(0, \eta_{RB})$.

According to [93, Eq.(2.3.23)], the probability density function (PDF) of the sum of two independent Laplace distributed random variables $\Delta\tau_{AR}$ and $\Delta\tau_{RB}$ can be determined as

$$f_{\Delta\tau_{AB}}(x) = 1 + \frac{1}{1-\frac{\eta^2_{AR}}{\eta^2_{RB}}}\left(\frac{\eta^2_{AR}}{\eta^2_{RB}}e^{-\frac{x}{\eta_{AR}}} - e^{-\frac{x}{\eta_{RB}}}\right). \tag{B.5}$$

After a simple mathematical derivation, we can obtain the CDF of $|\Delta\tau_{A,A}|$ as

$$F_{|\Delta\tau_{A,A}|}(x) = 1 + \frac{1}{1-(\frac{\eta_{AR}}{\eta_{RB}})^2}\left(\frac{\eta_{AR}^2}{\eta_{RB}^2}e^{-\frac{x}{\eta_{AR}}} - e^{-\frac{x}{\eta_{RB}}}\right). \tag{B.6}$$

Using (5.18), the probability that Q_τ outputs 1 under H_0, can be determined as

$$\begin{aligned} P_{\tau,H_0} &\triangleq \mathbf{Pr}(Q_\tau[|\hat{\tau}_{k_{AB}}(t) - \hat{\tau}_{k_{AB}}(t-1)|] = 1 \mid H_0) \\ &= \mathbf{Pr}(|\Delta\tau_{A,A}| > \delta_\tau) = 1 - \mathbf{Pr}(|\Delta\tau_{A,A}| \leq \delta_\tau). \end{aligned} \tag{B.7}$$

Substituting (B.6) into (B.7), we can obtain (5.25).

B.2 Proof of Lemma 13

Let $\Delta_{h_{E,A}}$ represent the difference between $\hat{h}_{n_{EB}}(t)$ and $\hat{h}_{n_{AB}}(t-1)$. Combining (5.13) and (5.21), $\Delta_{h_{E,A}}$ is determined as

$$\begin{aligned} \Delta_{h_{E,A}} &\triangleq \hat{h}_{n_{EB}}(t) - \hat{h}_{n_{AB}}(t-1) \\ &= \alpha_{RB}h_{n_{ER}}(t)h_{n_{RB}}(t-1) \\ &\quad + h_{n_{ER}}(t)\sqrt{1-\alpha_{RB}^2}u_{n_{RB}}(t-1) - h_{n_{AR}}(t-1)h_{n_{RB}}(t-1) + w_h(t) - w_h(t-1). \end{aligned} \tag{B.8}$$

For a given $h_{n_{RB}}(t-1)$, $\Delta_{h_{E,A}}$ is a complex Gaussian distributed random variable with zero mean and variance $\sigma^2_{\Delta h_{E,A}} = (\alpha_{RB}^2\sigma^2_{h_{ER}} + \sigma^2_{h_{AR}})|h_{n_{RB}}(t-1)|^2 + (1-\alpha_{RB}^2)\sigma^2_{h_{ER}}\sigma^2_{h_{RB}} + 2\sigma_w^2$. Then, the CDF of $|\Delta h_{E,A}|^2$ can be derived as

$$F_{|\Delta h_{E,A}|^2}(x) = 1 - \exp\left(-\frac{x}{\sigma^2_{\Delta h_{E,A}}}\right). \tag{B.9}$$

Based on (5.17), P_{h,H_1} can be evaluated as

$$P_{h,H_1} \triangleq \mathbf{Pr}(Q_h[|\hat{h}_{n_{EB}}(t) - \hat{h}_{n_{AB}}(t-1)|^2] = 1 \mid H_1) = 1 - \mathbf{Pr}(|\Delta h_{E,A}|^2 \leq \delta_h). \tag{B.10}$$

Substituting (B.9) into (B.10), we can get (5.30).

B.3 Proof of Lemma 14

Let $\Delta\tau_{E,A} = \hat{\tau}_{k_{EB}}(t) - \hat{\tau}_{k_{AB}}(t-1)$ and $\Delta\tau_{EA} = \hat{\tau}_{k_{ER}}(t) - \hat{\tau}_{k_{AR}}(t-1)$, and then we have $\Delta\tau_{E,A} \triangleq \Delta\tau_{EA} + \Delta\tau_{RB}$. Based on (5.15), $\Delta\tau_{EA}$ can be further written as

$$\Delta\tau_{EA} \triangleq (\hat{\tau}^{(1)}_{k_{ER}}(t))^2 + (\hat{\tau}^{(2)}_{k_{ER}}(t))^2 - (\hat{\tau}^{(1)}_{k_{AR}}(t-1))^2 - (\hat{\tau}^{(2)}_{k_{AR}}(t-1))^2. \tag{B.11}$$

We know that $\Delta\tau_{EA}$ is also a Laplace distributed random variable [93], that is,

$$\Delta\tau_{EA} \sim \mathit{Laplace}(0, \eta_{EA}). \tag{B.12}$$

Similarly, we can derive the PDF of the sum of two independent Laplace distributed random variables $\Delta\tau_{EA}$ and $\Delta\tau_{RB}$ by applying [93, Eq.(2.3.23)]. Thus, we have

$$f_{\Delta\tau_{E,A}}(x) = 1 + \frac{1}{\left(1 - \frac{\eta_{EA}^2}{\eta_{RB}^2}\right)} \left(\frac{\eta_{EA}^2}{\eta_{RB}^2} e^{-\frac{x}{\eta_{EA}}} - e^{-\frac{x}{\eta_{RB}}} \right). \tag{B.13}$$

The CDF of $|\Delta\tau_{E,A}|$ can be given by

$$F_{|\Delta\tau_{E,A}|}(x) = 1 + \frac{1}{\left(1 - \frac{\eta_{EA}^2}{\eta_{RB}^2}\right)} \left(\frac{\eta_{EA}^2}{\eta_{RB}^2} e^{-\frac{x}{\eta_{EA}}} - e^{-\frac{x}{\eta_{RB}}} \right). \tag{B.14}$$

According to (5.18), the probability that Q_τ outputs 1 under H_1 can be given by

$$P_{\tau,H_1} \triangleq \mathbf{Pr}(Q_\tau[|\hat{\tau}_{k_{EB}}(t) - \hat{\tau}_{k_{AB}}(t-1)|] = 1 \mid H_1) = 1 - \mathbf{Pr}(|\Delta\tau_{EB}| \leq \delta_\tau). \quad \text{(B.15)}$$

Finally, substituting (B.14) into (B.15) yields (5.31).

BIBLIOGRAPHY

BIBLIOGRAPHY

[1] Y. Zhang, Y. Shen, X. Jiang, and S. Kasahara, "Mode selection and spectrum partition for D2D inband communications: A physical layer security perspective," *IEEE Trans. Commun.*, vol. 67, no. 1, pp. 623–638, Jan. 2019.

[2] D. Shan, K. Zeng, W. Xiang, P. Richardson, and Y. Dong, "PHY-CRAM: Physical layer challenge-response authentication mechanism for wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 9, pp. 1817–1827, Sep. 2013.

[3] T. O. Olwal, K. Djouani, and A. M. Kurien, "A survey of resource management toward 5G radio access networks," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 3, pp. 1656–1686, thirdquarter 2016.

[4] W. Diffie and M. E. Hellman, "New directions in cryptography," *IEEE Trans. Inf. Theory*, vol. 22, no. 6, pp. 644–654, Nov. 1976.

[5] S. V. Kartalopoulos, "A primer on cryptography in communications," *IEEE Commun. Mag.*, vol. 44, no. 4, pp. 146–151, Apr. 2006.

[6] P. Christof, J. Pelzl, and B. Preneel, *Understanding Cryptography: A Textbook for Students and Practitioners.* Springer, 2010.

[7] U. Raza, P. Kulkarni, and M. Sooriyabandara, "Low power wide area networks: An overview," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 2, pp. 855–873, Jan. 2017.

[8] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. K. Soong, and J. C. Zhang, "What will 5G be?" *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, Jun. 2014.

[9] X. B. Wang, P. Hao, and L. Hanzo, "Physical-layer authentication for wireless security enhancement: current challenges and future developments," *IEEE Commun. Mag.*, vol. 54, no. 6, pp. 152–158, Jun. 2016.

[10] N. Yang, L. Wang, G. Geraci, M. Elkashlan, J. Yuan, and M. D. Renzo, "Safeguarding 5G wireless communication networks using physical layer security," *IEEE Commun. Mag.*, vol. 53, no. 4, pp. 20–27, Apr. 2015.

[11] P. Gope and B. Sikdar, "Lightweight and privacy-preserving two-factor authentication scheme for IoT devices," *IEEE Internet Things J.*, vol. 6, no. 1, pp. 580–589, Feb. 2019.

[12] L. Xiao, L. J. Greenstein, N. B. Mandayam, and W. Trappe, "Channel-based spoofing detection in frequency-selective Rayleigh channels," *IEEE Trans. Wireless Commun.*, vol. 8, no. 12, pp. 5948–5956, Dec. 2009.

[13] X. W. L. Xiao and Z. Han, "PHY-layer authentication with multiple landmarks with reduced overhead," *IEEE Trans. Wireless Commun.*, vol. 17, no. 3, pp. 1676–1687, Dec. 2017.

[14] J. Z. Liu and X. B. Wang, "Physical layer authentication enhancement using two-dimensional channel quantization," *IEEE Trans. Wireless Commun.*, vol. 15, no. 6, pp. 4171–4182, Feb. 2016.

[15] V. Brik, S. Banerjee, M. Gruteser, and S. Oh, "Paradis: Wireless device identification with radiometric signatures," in *Proc. ACM MobiCom*, Sep. 2008, pp. 116–127.

[16] W. Wang, Z. Sun, S. Piao, B. Zhu, and K. Ren, "Wireless physical-layer identification: Modeling and validation," *IEEE Trans. Inf. Forensics Security*, vol. 11, no. 9, pp. 2091–2106, Sep. 2016.

[17] E. Björnson, J. Hoydis, M. Kountouris, and M. Debbah, "Massive MIMO systems with non-ideal hardware: Energy efficiency, estimation, and capacity limits," *IEEE Trans. Inf. Theory*, vol. 60, no. 11, pp. 7112–7139, Nov. 2014.

[18] E. Björnson, M. Matthaiou, and M. Debbah, "Massive MIMO with non-ideal arbitrary arrays: Hardware scaling laws and circuit-aware design," *IEEE Tran. Wireless Commun.*, vol. 14, no. 8, pp. 4353–4368, Aug. 2015.

[19] C. Studer, M. Wenk, and A. Burg, "MIMO transmission with residual transmit-RF impairments," in *Proc. ITG/IEEE Workshop on Smart Antennas (WSA)*, 2010.

[20] T. Schenk, *RF Imperfections in High-Rate Wireless Systems: Impact and Digital Compensation*. Springer, 2008.

[21] M. Wenk, "MIMO-OFDM-Testbed: Challenges, implementations, and measurement results, ser. series in microelectronics," Ph.D. dissertation, ETH Zurich, Hartung-Gorre, 2010.

[22] and R. A. Pacheco, , and D. Hatzinakos, "Joint estimation of channel response, frequency offset, and phase noise in OFDM," *IEEE Trans. Signal Process.*, vol. 54, no. 9, pp. 3542–3554, Sep. 2006.

[23] H. Mehrpouyan, A. A. Nasir, S. D. Blostein, T. Eriksson, G. K. Karagiannidis, and T. Svensson, "Joint estimation of channel and oscillator phase noise in MIMO systems," *IEEE Trans. Signal Process.*, vol. 60, no. 9, pp. 4790–4807, Sep. 2012.

[24] O. H. Salim, A. A. Nasir, H. Mehrpouyan, W. Xiang, S. Durrani, and R. A. Kennedy, "Channel, phase noise, and frequency offset in OFDM systems: Joint estimation, data detection, and hybrid Cramér-Rao lower bound," *IEEE Trans. Commun.*, vol. 62, no. 9, pp. 3311–3325, Sep. 2014.

[25] F. Septier, Y. Delignon, A. Menhaj-Rivenq, and C. Garnier, "Monte Carlo methods for channel, phase noise, and frequency offset estimation with unknown noise variances in OFDM systems," *IEEE Trans. Signal Process.*, vol. 56, no. 8, pp. 3613–3626, Aug. 2008.

[26] X. He and A. Yener, "End-to-end secure multi-hop communication with untrusted relays," *IEEE Trans. Wireless Commun.*, vol. 12, no. 1, pp. 1–11, Jan. 2013.

[27] S. Vatedka, N. Kashyap, and A. Thangaraj, "Secure compute-and-forward in a bidirectional relay," *IEEE Trans. Inf. Theory*, vol. 61, no. 5, pp. 2531–2556, May 2015.

[28] Q. Liu and G. Gong, "Physical layer secure information exchange protocol for mimo ad hoc networks against passive attacks," in *Proc. 2016 IEEE global communications conference (GLOBECOM)*, Dec. 2016, pp. 1–6.

[29] P. Zhang and X. Jiang, "Channel-based authentication for dual-hop wireless networks," in *Proc. 2018 International Conference on Networking and Network Applications (NaNA)*, Oct. 2018, pp. 42–46.

[30] W. C. Jakes and D. C. Cox, *Microwave mobile communications.* Wiley, 1994.

[31] L. Xiao, G. L. J, N. B. Mandayam, and W. Trappe, "Using the physical layer for wireless authentication in time-variant channels," *IEEE Trans. Wireless Commun.*, vol. 7, no. 7, pp. 2571–2579, Jul. 2008.

[32] P. Baracca, N. Laurenti, and S. Tomasin, "Physical layer authentication over MIMO fading wiretap channels," *IEEE Trans. Wireless Commun.*, vol. 11, no. 7, pp. 2564–2573, Jul. 2012.

[33] J. K. Tugnait and H. Kim, "A channel-based hypothesis testing approach to enhance user authentication in wireless networks," in *Proc. IEEE COMSNETS*, Jan. 2010, pp. 1–9.

[34] J. K. Tugnait, "Wireless user authentication via comparison of power spectral densities," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 9, pp. 1791–1802, Sep. 2013.

[35] A. Ferrante, N. Laurenti, C. Masiero, M. Pavon, and S. Tomasin, "On the error region for channel estimation-based physical layer authentication over Rayleigh fading," *IEEE Trans. Inf. Forensics Security*, vol. 10, no. 5, pp. 941–952, May 2015.

[36] K. Zeng, K. Govindan, and P. Mohapatra, "Non-cryptographic authentication and identification in wireless networks," *IEEE Wireless Commun.*, vol. 17, no. 5, pp. 56–62, Oct. 2010.

[37] A. C. Polak, S. Dolatshahi, and D. L. Goeckel, "Identifying wireless users via transmitter imperfections," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 7, pp. 1469–1479, Aug. 2011.

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

[39] P. Hao, X. Wang, and A. Behnad, "Relay authentication by exploiting i/q imbalance in amplify-and-forward system," in *Proc. IEEE Conf. on Global Commun. (GLOBECOM)*, Dec. 2014, pp. 613–618.

[40] W. Hou, X. Wang, J. Chouinard, and A. Refaey, "Physical layer authentication for mobile systems with time-varying carrier frequency offsets," *IEEE Trans. Commun.*, vol. 62, no. 5, pp. 1658–1667, May 2014.

[41] W. E. Cobb, E. D. Laspe, R. O. Baldwin, M. A. Temple, and Y. C. Kim, "Intrinsic physical-layer authentication of integrated circuits," *IEEE Trans. Inf. Forensics Security*, vol. 7, no. 1, pp. 14–24, Feb. 2011.

[42] T. J. Bihl, K. W. Bauer, and M. A. Temple, "Feature selection for RF fingerprinting with multiple discriminant analysis and using zigbee device emissions," *IEEE Trans. Inf. Forensics Security*, vol. 11, no. 8, pp. 1862–1874, Aug. 2016.

[43] S. V. Radhakrishnan, A. S. Uluagac, and R. Beyah, "Gtid: A technique for physical deviceanddevice type fingerprinting," *IEEE Trans. Depend. Sec. Comput.*, vol. 12, no. 5, pp. 519–532, Sep. 2014.

[44] P. L. Yu, J. S. Baras, and B. M. Sadler, "Physical-layer authentication," *IEEE Trans. Inf. Forensics Security*, vol. 3, no. 1, pp. 38–51, Feb. 2008.

[45] P. L. Yu and B. M. Sadler, "MIMO authentication via deliberate fingerprinting at the physical layer," *IEEE Trans. Inf. Forensics Security*, vol. 6, no. 3, pp. 606–615, Sep. 2011.

[46] G. Verma, P. Yu, and B. M. Sadler, "Physical layer authentication via fingerprint embedding using software-defined radios," *IEEE Access*, vol. 3, pp. 81–88, Jan. 2015.

[47] N. Xie and S. Zhang, "Blind authentication at the physical layer under time-varying fading channels," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 7, pp. 1465–1479, Jul. 2018.

[48] N. Xie and C. Chen, "Slope authentication at the physical layer," *IEEE Trans. Inf. Forensics Security*, vol. 13, no. 6, pp. 1579–1594, Jun. 2018.

[49] J. B. Perazzone, P. L. Yu, B. M. Sadler, and R. S. Blum, "Cryptographic side-channel signaling and authentication via fingerprint embedding," *IEEE Trans. Inf. Forensics Security*, vol. 13, no. 9, pp. 2216–2225, Sep. 2018.

[50] C. Komninakis, C. Fragouli, A. H. Sayed, and R. D. Wesel, "Multi-input multi-output fading channel tracking and equalization using kalman estimation," *IEEE Trans. Signal Process.*, vol. 50, no. 5, pp. 1065–1076, May 2002.

[51] H. Liu, Y. Wang, J. Liu, J. Yang, Y. Chen, and H. V. Poor, "Authenticating users through fine-grained channel information," *IEEE Trans. Mobile Comput.*, vol. 17, no. 2, pp. 251–264, Feb. 2018.

[52] E. Björnson. Channel hardening makes fading channels behave as deterministic. [Online]. Available: https://ma-mimo.ellintech.se/2017/01/25/channel-hardening-makes-fading-channels-behave-as-deterministic/

[53] S. L. Loyka, "Channel capacity of MIMO architecture using the exponential correlation matrix," *IEEE Commun. Lett.*, vol. 5, no. 9, pp. 369–371, Sep. 2001.

[54] E. Bjornson, D. Hammarwall, and B. Ottersten, "Exploiting quantized channel norm feedback through conditional statistics in arbitrarily correlated MIMO systems," *IEEE Trans. Signal Process.*, vol. 57, no. 10, pp. 4027–4041, Oct. 2009.

[55] A. Hjørungnes, *Complex-valued Matrix Derivatives: With Applications in Signal Processing and communications*. Cambridge University, 2011.

[56] S. M. Kay, *Fundamentals of Statistical Signal Processing, Volumne II: Detection Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1998.

[57] S. Wang, A. Abdi, J. Salo, H. M. El-Sallabi, J. W. Wallace, P. Vainikainen, and M. A. Jensen, "Time-varying MIMO channels: Parametric statistical modeling and experimental results," *IEEE Trans. Veh. Technol.*, vol. 56, no. 4, pp. 1949–1963, Jul. 2007.

[58] M. R. Avendi and H. H. Nguyen, "Performance of selection combining for differential amplify-and-forward relaying over time-varying channels," *IEEE Trans. Wireless Commun.*, vol. 13, no. 8, pp. 4156–4166, Apr. 2014.

[59] H. L. V. Trees, K. L. Bell, and Z. Tian, *Detection, Estimation, and Modulation Theory, Part I: Detection, Estimation, and Filtering Theory*, 2nd ed. Wiley, 2014.

[60] A. Sabharwal, A. Khoshnevis, and E. Knightly, "Opportunistic spectral usage: Bounds and a multi-band CSMA/CA protocol," *IEEE/ACM Trans. Netw.*, vol. 15, no. 3, pp. 533–545, Jun. 2007.

[61] D. Tse and P. Viswanath, *Fundamentals of wireless communication*. Cambridge university press, 2005.

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

[63] A. Pitarokoilis, E. Björnson, and E. G. Larsson, "ML detection in phase noise impaired simo channels with uplink training," *IEEE Trans. Commun.*, vol. 64, no. 1, pp. 223–235, Jan. 2016.

[64] A. Pitarokoilis, S. K. Mohammed, and E. G. Larsson, "Uplink performance of time-reversal MRC in massive MIMO systems subject to phase noise," *IEEE Tran. Wireless Commun.*, vol. 14, no. 2, pp. 711–723, Feb. 2015.

[65] O. H. Salim, A. A. Nasir, H. Mehrpouyan, and W. Xiang, "Multi-relay communications in the presence of phase noise and carrier frequency offsets," *IEEE Tran. on Commun.*, vol. 65, no. 1, pp. 79–94, Jan. 2017.

[66] Y. Wang and J. Lee, "A simple phase noise suppression scheme for massive MIMO uplink systems," *IEEE Trans. Veh. Technol.*, vol. 66, no. 6, pp. 4769–4780, Jun. 2017.

[67] K. J. Kim, R. A. Iltis, and H. V. Poor, "Frequency offset and channel estimation in cooperative relay networks," *IEEE Trans. Veh. Technol.*, vol. 60, no. 7, pp. 3142–3155, Sep. 2011.

[68] A. A. Nasir, H. Mehrpouyan, R. Schober, and Y. Hua, "Phase noise in MIMO systems: Bayesian Cramér-Rao bounds and soft-input estimation," *IEEE Trans. Signal Process.*, vol. 61, no. 10, pp. 2675–2692, May 2013.

[69] C. Zhao, M. Huang, L. Huang, X. Du, and M. Guizani, "A robust authentication scheme based on physical-layer phase noise fingerprint for emerging wireless networks," *Computer Networks*, vol. 128, no. 9, pp. 164–171, May 2017.

[70] O. Besson and P. Stoica, "On parameter estimation of MIMO flat-fading channels with frequency offsets," *IEEE Trans. Signal Process.*, vol. 51, no. 3, pp. 602–613, Mar. 2003.

[71] Z. Wang, P. Babu, and D. P. Palomar, "Effective low-complexity optimization methods for joint phase noise and channel estimation in OFDM," *IEEE Trans. Signal Process.*, vol. 65, no. 12, pp. 3247–3260, Jun. 2017.

[72] K. J. Kim, M. Pun, and R. A. Iltis, "Joint carrier frequency offset and channel estimation for uplink MIMO-OFDMA systems using parallel Schmidt Rao-Blackwellized particle filters," *IEEE Trans. Commun.*, vol. 58, no. 9, pp. 2697–2708, Sep. 2010.

[73] P. Tichavsky, C. H. Muravchik, and A. Nehorai, "Posterior Cramér-Rao bounds for discrete-time nonlinear filtering," *IEEE Trans. Signal Process.*, vol. 46, no. 5, pp. 1386–1396, May 1998.

[74] D. Cox and R. Leck, "Distributions of multipath delay spread and average excess delay for 910-MHz urban mobile radio paths," *IEEE Trans. Antennas Propag.*, vol. 23, no. 2, pp. 206–213, Mar. 1975.

[75] U. M. Maurer, "Authentication theory and hypothesis testing," *IEEE Trans. Inf. Theory*, vol. 46, no. 4, pp. 1350–1356, Jul. 2000.

[76] M. Uysal, "Diversity analysis of space-time coding in cascaded Rayleigh fading channels," *IEEE Commun. Lett.*, vol. 10, no. 3, pp. 165–167, Mar. 2006.

[77] Y. Cho, J. Kim, W. Yang, and C. Kang, *MIMO-OFDM wireless communications with MATLAB.* John Wiley & Sons, 2010.

[78] T. G. Manickam, R. J. Vaccaro, and D. W. Tufts, "A least-squares algorithm for multipath time-delay estimation," *IEEE Trans. Signal Process.*, vol. 42, no. 11, pp. 3229–3233, Nov. 1994.

[79] M. C. Vanderveen, A. . V. der Veen, and A. Paulraj, "Estimation of multipath parameters in wireless communications," *IEEE Trans. Signal Process.*, vol. 46, no. 3, pp. 682–690, Mar. 1998.

[80] F. Ge, D. Shen, Y. Peng, and V. O. K. Li, "Super-resolution time delay estimation in multipath environments," *IEEE Comput. Sci. Eng. Mag.*, vol. 54, no. 9, pp. 1977–1986, Sep. 2007.

[81] C. S. Patel and G. L. Stuber, "Channel estimation for amplify and forward relay based cooperation diversity systems," *IEEE Trans. Wireless Commun.*, vol. 6, no. 6, pp. 2348–2356, Jun. 2007.

[82] F. Gao, T. Cui, and A. Nallanathan, "On channel estimation and optimal training design for amplify and forward relay networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 5, pp. 1907–1916, May 2008.

[83] C. Wang, H. M. Wang, and X. G. Xia, "Hybrid opportunistic relaying and jamming with power allocation for secure cooperative networks," *IEEE Trans. Wireless Commun.*, vol. 14, no. 2, pp. 589–605, Feb. 2015.

[84] E. Everett, A. Sahai, and A. Sabharwal, "Passive self-interference suppression for full-duplex infrastructure nodes," *IEEE Tran. Wireless Commun.*, vol. 13, no. 2, pp. 680–694, Feb. 2014.

[85] F. Zhu, F. Gao, T. Zhang, K. Sun, and M. Yao, "Physical-layer security for full duplex communications with self-interference mitigation." *IEEE Trans. Wireless Commun.*, vol. 15, no. 1, pp. 329–340, Jan. 2016.

[86] J. B. Kim, J. Lim, and J. M. Cioffi, "Capacity scaling and diversity order for secure cooperative relaying with untrustworthy relays," *IEEE Trans. Wireless Commun.*, vol. 14, no. 7, pp. 3866–3876, Jul. 2015.

[87] M. R. Avendi and H. H. Nguyen, "Performance of differential amplify-and-forward relaying in multinode wireless communications," *IEEE Trans. Veh. Technol.*, vol. 62, no. 8, pp. 3603–3613, Oct. 2013.

[88] N. O'Donoughue and J. M. F. Moura, "On the product of independent complex gaussians," *IEEE Trans. Signal Process.*, vol. 60, no. 3, pp. 1050–1063, Mar. 2012.

[89] C. Wang, H. M. Wang, and X. G. Xia, "Hybrid opportunistic relaying and jamming with power allocation for secure cooperative networks," *IEEE Trans. Wireless Commun.*, vol. 14, no. 2, pp. 589–605, Feb. 2015.

[90] Baesd-MATLAB simulator for E2E PLA for dual-hop wireless networks. [Online]. Available: https://github.com/zpcanson/E2E-PLA-simulation.

[91] Y. Zheng and C. Xiao, "Simulation models with correct statistical properties for Rayleigh fading channels," *IEEE Transactions on communications*, vol. 51, no. 6, pp. 920–928, Jun. 2003.

[92] K. R. Rao and N. Ahmed, "Recursive techniques for obtaining the partial fraction expansion of a rational function," *IEEE Trans. Educ.*, vol. 11, no. 2, pp. 152–154, Jun. 1968.

[93] S. Kotz, T. Kozubowski, and K. Podgorski, *The Laplace distribution and generalizations: a revisit with applications to communications, economics, engineering, and finance*. Springer Science & Business Media, 2012.

Publications

Jounal Articles

[1] Pinchang Zhang, Jinxiao Zhu, Yin Chen, and Xiaohong Jiang. "End-to-end physical layer authentication for dual-hop wireless networks," *IEEE Access,* vol.7, pp. 38322-38336, Mar. 2019.

[2] Pinchang Zhang, Tarik Taleb, Xiaohong Jiang, and Bin Wu. "Physical layer authentication for massive MIMO systems with hardware impairments," *IEEE Transactions on Wireless Communications*, vol. 19, no. 3, pp. 349-360, Mar. 2020.

[3] Pinchang Zhang, Jun Liu, Yulong Shen, Hewu Li, and Xiaohong Jiang. "Lightweight tag-based PHY-layer authentication for IoT devices in smart cities," *IEEE Internet of Things Journal*, DOI: 10.1109/JIOT.2019.2958079.

[4] Pinchang Zhang, Yulong Shen, Xiaohong Jiang, and Bin Wu. "Physical layer authentication jointly utilizing channel and phase noise in MIMO systems," *IEEE Transactions on Communications*, DOI: 10.1109/TCOMM.2020.2967393.

[5] Pinchang Zhang, Jun Liu, Yulong Shen, and Xiaohong Jiang. "Exploiting channel gain and phase noise for PHY-layer authentication in massive MIMO systems," *IEEE Transactions on Information Forensics and Security.* (Response to major revision, submitted)

Conference Papers

[6] Pinchang Zhang and Xiaohong Jiang. Channel-based authentication for dual-hop wireless networks. 2018 International Conference on Networking and Network Applications (NaNA), Xi'an, China, Oct., pp. 42-46, 2018