Physical Layer Authentication for Massive MIMO Systems with Hardware Impairments

Pinchang Zhang, *Student Member, IEEE*, Tarik Taleb, *Senior Member, IEEE*,

Xiaohong Jiang, *Senior Member, IEEE*, and Bin Wu

***Abstract*—We study transmitter authentication in massive multiple-input multiple-output (MIMO) systems with non-ideal hardware for the fifth generation (5G) and beyond networks. A new channel-based authentication scheme is proposed by taking hardware impairments into account. Based on signal processing theory, we first 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 then 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.**

***Index Terms*—Physical layer authentication, wireless security, massive multiple-input multiple-output (MIMO), hardware impairments, the fifth generation (5G) and beyond networks.**

I. Introduction

The emerging massive MIMO system, with tens or hundreds of base station antennas serving a large number of mobile terminals, has recently attracted substantial interests from both academic and industry societies [1], [2]. It is deemed as one of the most promising techniques to support the increasing demand for wireless services in the fifth generation (5G) and beyond networks [3]. Compared to currently deploying MIMO systems with a small number of antennas, massive MIMO systems provide tremendous performance gains in terms of energy and/or spectral efficiency by utilizing large-scale antenna arrays with largely enhanced spatial resolution and array gain [4], [5]. Also, channel random impairments such as propagation losses and thermal noise can be easily mitigated by coherent beamforming/combining [6], [7]. Therefore, there is no doubt that massive MIMO will be integrated in the upcoming 5G and beyond standards.

Authentication serves as an unrivalled security service by verifying entity identity to achieve secure communications [8], [9]. Thus, providing flexible and cost-effective authentication paradigms verifying the claimed identity of a legitimate transmitter and refusing an adversarial impersonation is becoming an increasingly urgent demand for massive MIMO systems. This is because the open nature of wireless medium makes wireless networks more vulnerable to impersonation attacks [10]. Moreover, the increasingly dynamic mobile environments and the randomness of devices, joining in or leaving the network at anytime, lead to a highly challenging issue on cryptographic key distribution and management in decentralized networks [11]. In other words, cryptographic-based authentication is hard to be implemented in massive MIMO systems. Although there has been important efforts on massive MIMO research, authentication paradigms detecting impersonation attacks.

Recently, physical layer authentication techniques, which exploit the intrinsic and unique features of physical layer for authentication, have received significant attention to enhance the conventional cryptography-based authentication solutions [12]–[23]. So far, extensive research efforts have been devoted to the study of effective physical layer authentication methods for non-massive MIMO systems. Those existing works can be broadly divided into three categories: fingerprinting authentication, watermarking authentication, and channel-based authentication. The basic principle of fingerprinting authentication is that the radio frequency-distinct native attribute (RF-DNA) can hardly be mimicked after being manufactured, such that RF-DNA can be used to uniquely identify devices. The authors in [12] explore the analysis of distortion signals resulting from hardware impairments to identify wireless devices. Carrier frequency offset caused by hardware impairments is investigated to identify wireless transmitters in [13]. The authors in [14] examine the reliability and differentiability of fingerprinting authentication via theoretical modeling as well as experiment validation. In [15]–[17], watermarking authentication is based on a secret tag that is superimposed to the modulated signals to be transmitted. A watermarking-based authentication scheme, which relies on a cryptography secure low-power authentication tag hidden in the modulated signals for authentication, is investigated in [15]. The authors in [16] further conduct tag-based authentication experiments in software defined radio systems. An extension of conventional

This work was supported in part by China NSFC Grants No.61702068, 61972308, 61571352, U1736216 and U1536202, in part by Japan JSPS Grant No.18H03235, and in part by Anhui NSF Grant 1808085MF165. This work was also supported in part by the European Union's Horizon 2020 research and innovation programme under the INSPIRE-5Gplus project (Grant No. 871808). (Corresponding author: Tarik Taleb.)

P. Zhang is with the School of Computer and Information Engineering, Chuzhou University, Anhui, 239000, China. He is also with the School of Systems Information Science, Future University Hakodate, 116-2 Kamedanakano-cho, Hakodate, Hokkaido, 041-8655, Japan (e-mail: zpcap0505238@163.com).

T. Taleb is with the Department of Communications and Networking, Aalto University, Espoo 02150, Finland. He is also with Sejong University, seoul, Korea (e-mail: talebtarik@ieee.org).

X. Jiang is with the School of Systems Information Science, Future University Hakodate, 116-2 Kamedanakano-cho, Hakodate, Hokkaido, 041-8655, Japan (e-mail: jiang@fun.ac.jp).

B. Wu is with the School of Computer Science and Technology, Tianjin University, Tianjin, 300072 P. R. China (e-mail: binwu.tju@gmail.com).

watermarking methods is proposed in [17] to authenticate wireless devices by jointly utilizing side-channel information and tag selection.

Channel-based authentication [20]–[23] mainly utilizes the inherent characteristics of wireless channels (e.g., location-specific property) to identify wireless devices. The authors in [20] present a channel-based authentication scheme by exploiting spatial variability of channel frequency response over time-varying channels in a rich scattering environment. The authors in [22] propose a novel scheme jointly using location-specific properties of both channel amplitude and multipath delay of channels to identify transmitters. The authors in [23] propose a logistic regression-based authentication by exploiting channel state information (CSI) and multiple landmarks to improve the spoofing detection accuracy.

It is notable that the above mentioned works mainly focus on authenticating in non-massive MIMO systems. To simplify performance analysis, the above studies assume that the transceivers involved in communications are ideal hardware. However, practical transceivers are commonly non-ideal, especially when a large number of cheap low-power hardware is deployed in massive MIMO systems. It is proved that hardware impairments such as power amplifiers non-linearities, amplitude/phase imbalance in I/Q (in-phase and quadrature) mixers, multiuser interference due to I/Q mismatch [24], and mutual coupling hardware mismatches induced channel non-reciprocity [25], are attributed to the use of inexpensive low-power hardware [1], [5]–[7]. Although the impact of hardware impairments can be substantially reduced by analog or digital compensation algorithms, these approaches cannot completely eliminate impairments. This is because physical transceiver implementations are comprised of multiple modules such as converters, amplifiers, and oscillators [26], and thus each module distorts signals in its own way [6]. It is difficult (if not impossible) to accurately estimate and analytically model the characteristics of some impairments [27]. As a result, hardware imperfections are unavoidable.

It is demonstrated that the presence of hardware impairments not only limits capacity but also deteriorates channel estimation accuracy in the high-power regime [6], [7]. 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 impairment which is characterized by independent additive distortion noise at base station and user terminals [6], [27]–[29]. Based on this background, this paper focuses on physical layer authentication for massive MIMO systems with aggregate residual hardware impairments.

The main contributions of this paper are summarized as follows:

- By utilizing location-specific property of wireless channels and considering hardware impairments to authenticate transmitters, we 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 the probabilities of false alarm and detection under different channel covariance matrix models. Simulation results are also provided to validate theoretical modeling of the two probabilities.
- Through theoretical models, we further examine how different levels of hardware impairments can impact on authentication performance, and also determine how to set antennas correlation pattern and the number of base station antennas to achieve a required authentication performance.

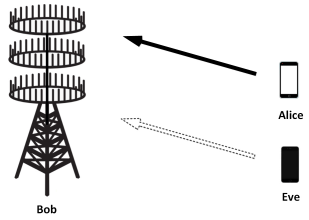


Fig. 1. System model.

The remainder of this paper is organized as follows. Section II depicts the envisioned system model. Section III presents the proposed scheme. Formulations of the probabilities for false alarm and detection is provided in Section IV, and numerical results are shown in Section V. Finally, Section VI concludes this paper.

Notation: Let $(\cdot)^*$, $(\cdot)^T$, and $(\cdot)^H$ denote conjugate, transpose, and conjugate transpose operators, respectively. $|\cdot|$ denotes absolute value operator. $\|\mathbf{x}\|$ denotes L_2-norm of a vector $\mathbf{x}$. $\mathbb{C}^{M\times K}$ represents the set of complex-valued $M \times K$ matrices. $\mathrm{Cov}(\cdot)$ denotes covariance operator. $\det(\cdot)$ denotes determinant operator. A circularly symmetric complex Gaussian random vector $\mathbf{x}$ with zero mean and covariance matrix $\mathbf{R}$ can be denoted by $\mathbf{x} \sim \mathcal{CN}(\mathbf{0}, \mathbf{R})$. Identity matrix is denoted by $\mathbf{I}$. $\mathbb{E}(\cdot)$ and $\mathbf{Pr}(\cdot)$ represent expectation and probability operators, respectively. $\triangleq$, and $\mathrm{tr}(\cdot)$ represent definitions and matrix trace function, respectively. $\mathrm{diag}[\lambda_1, ..., \lambda_M]$ represents a diagonal matrix with $\lambda_1, ..., \lambda_M$ on the main diagonal. $\exp(\cdot)$ denotes exponential function. $\mathbf{\Gamma}_{\chi_i^2}(\cdot)$ denotes the right-tail probability function for a χ_i^2 random variable with i degrees of freedom.

II. SYSTEM MODEL

A. Network Model

As illustrated in Fig. 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 [21], [22]. 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 as being from Alice by using a standard higher-layer protocol [21], and Bob stores the channel information connecting Alice with him. The other one, received by Bob at time k, is either from Alice or Eve. Therefore, Bob is supposed 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 the channel coherence time [20].

B. 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 works in [18], [20], [22], we assume that channels from the same transmitter-receiver pair are temporally correlated and follow Rayleigh fading. The temporally correlated channel may be spatially independent or correlated. We use the subscripts A, B and E to denote the terms associated with Alice, Bob and Eve, respectively. Then, let subscript t denote an unknown transmitter, i.e., $t = \{A, E\}$. The channel vector between t and Bob at time k is denoted by $\mathbf{h}_t[k] = [h_{t,1}[k] \cdots h_{t,M}[k]]^T \in \mathbb{C}^{M\times 1}$ with $h_{t,m}[k]$ representing the m^{th} component of $\mathbf{h}_t[k]$, and then we have $\mathbf{h}_t[k] \sim \mathcal{CN}(\mathbf{0}, \mathbf{R}_t)$ where $\mathbf{R}_t = \mathbb{E}\{\mathbf{h}_t[k]\mathbf{h}_t^H[k]\} \in \mathbb{C}^{M\times M}$ is a symmetric positive semi-definite matrix. Following existing related literature [21], it is assumed that the statistical information of channel can be available at Bob. This assumption is generic and has been adopted in [17].

Here, we exemplify temporal channel variations. We first focus on the time-autocorrelation of channels, which is caused by the Doppler rate. Similar to [20], 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

$$\begin{aligned}\mathbf{\Psi}_A[k_s] &= \mathbb{E}\{\mathbf{h}_A[k]\mathbf{h}_A^*[k+k_s]\} \\ &= \mathbf{R}_A J_0(2\pi f k_s),\end{aligned} \tag{1}$$

where $J_0(\cdot)$ is the zero^{th} order Bessel function of the first kind.

Similar to [18], a first-order Gauss-Markov process is employed to model the fluctuation of channel state. According to [18], correlation coefficient matrix of $\mathbf{h}_A[k]$ can be defined as $\mathbf{\Psi}_A(k_s)\mathbf{R}_A^{-1}$. Thus, we have

$$\mathbf{h}_A[k] = \alpha \mathbf{h}_A[k-1] + \sqrt{1-\alpha^2}\mathbf{e}_A[k], \tag{2}$$

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

C. 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; and 2) the received signal is distorted during reception processing. Such impairments are treated as the inclusion of additional distortion noise sources which are in general relevant to signal power and channel gain. Various sources of impairments (e.g., I/Q imbalance and phase noise) result in distortion noise [6].

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 [6]. For the authentication performance investigation, this is reasonable because considering the aggregate effect of all residual hardware impairments is more significant than considering them 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 t 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 t 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

$$\mathbf{y}_{Bt}[k] = \mathbf{h}_t[k](s[k] + \eta_t[k]) + \boldsymbol{\eta}_B[k] + \boldsymbol{\nu}[k], \tag{3}$$

where $\eta_t[k] \in \mathbb{C}$ denotes the independent additional distortion noise at t and $\boldsymbol{\eta}_B[k] \in \mathbb{C}^{M\times 1}$ denotes that at Bob. According to [6], ergodic stochastic processes can model aggregate residual impairments. Note that distortion noise caused by hardware impairments is irrelevant to $s[k]$, but statistically depends on channel realizations. Also, this distortion noise follows a complex Gaussian distribution for a given channel

TABLE I
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

realization, which is verified experimentally and supported by several theoretical results [6]. Specifically, under a given $\mathbf{h}_t[k]$ the conditional distributions are $\eta_t \sim \mathcal{CN}(0, \varsigma_t)$ and $\boldsymbol{\eta}_B \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{\Upsilon}_B)$, respectively, wherein ς_t and $\boldsymbol{\Upsilon}_B$ can be modeled as

$$\varsigma_t = \kappa_t p, \tag{4}$$

$$\boldsymbol{\Upsilon}_B = \kappa_B p \mathrm{diag}[|h_{t_1}[k]|^2, ..., |h_{t_M}[k]|^2], \tag{5}$$

where both κ_t, $\kappa_B \geq 0$ characterize levels of hardware impairments at t 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 is illustrated by an example: EVM at t can be formulated as

$$\mathbf{EVM}_t = \sqrt{\frac{\mathbb{E}\{|\eta_t[k]|^2\}}{\mathbb{E}\{|s[k]|^2\}}} = \sqrt{\kappa_t}. \tag{6}$$

Remark 1. *A small EVM result is required in the transmitter and receiver for correct demodulation when modulation density increases. Table I 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 paper 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.*

III. Proposed Physical Layer Authentication Scheme

The basic principle for the proposed scheme is that channels are location-specific, which has been widely adopted to authenticate transmitters to complement and improve traditional security approaches [20]–[22], [31]. More 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 [20], [30], [32]. 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 [21]). 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 channel estimation with hardware impairments and decision criterion.

A. Channel Estimation

If $\mathbf{R}_{t,\mathrm{diag}} = \mathrm{diag}[r_{11}, ..., r_{MM}]$ consists of diagonal elements of $\mathbf{R}_t$, the covariance matrix of $\mathbf{y}_{Bt}[k]$ according to (3) is denoted as

$$\begin{aligned}\mathbf{R}_{\mathbf{y}_{Bt}} &= \mathbb{E}\{\mathbf{y}_{Bt}[k]\mathbf{y}_{Bt}^H[k]\} \\ &= p(1+\kappa_t)\mathbf{R}_t + p\kappa_B\mathbf{R}_{t,\mathrm{diag}} + (\sigma_I^2 + \sigma_N^2)\mathbf{I}.\end{aligned} \tag{7}$$

Let $\hat{\mathbf{h}}_t[k]$ denote the estimation of $\mathbf{h}_t[k]$ and then by using linear minimum mean square error estimator [6] we have

$$\hat{\mathbf{h}}_t[k] = s^*[k]\mathbf{R}_t\mathbf{R}_{\mathbf{y}_{Bt}}^{-1}\mathbf{y}_{Bt}[k]. \tag{8}$$

Based on (7) and (8), we can establish the following lemma on channel estimation with hardware impairments. The proof is straightforward, and a similar one can be found in [6].

Lemma 1. *$\hat{\mathbf{h}}_t[k]$ can be decomposed as*

$$\hat{\mathbf{h}}_t[k] = \mathbf{h}_t[k] - \boldsymbol{\epsilon}_t[k]. \tag{9}$$

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

$$\mathbf{R}_{\boldsymbol{\epsilon}_t} = \mathbb{E}\{\boldsymbol{\epsilon}_t[k]\boldsymbol{\epsilon}_t^H[k]\} = \mathbf{R}_t - p\mathbf{R}_t\mathbf{R}_{\mathbf{y}_{Bt}}^{-1}\mathbf{R}_t. \tag{10}$$

As observed from (7) and (10), 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 results from AWGN and interference.

B. 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 previous ones at time $k-1$. We use "$\simeq$" to denote similarity and "$\neg$" denotes negation. Therefore, the hypothesis test can be formulated as

$$\begin{aligned}\mathcal{H}_0 &: \ \hat{\mathbf{h}}_t[k] \simeq \hat{\mathbf{h}}_A[k-1], \\ \mathcal{H}_1 &: \ \hat{\mathbf{h}}_t[k] \neg\simeq \hat{\mathbf{h}}_A[k-1],\end{aligned} \tag{11}$$

where the null hypothesis $\mathcal{H}_0$ represents that the current transmitter is still Alice, i.e., $t = A$. In contrast, the alternative hypothesis $\mathcal{H}_1$ represents that the current transmitter is adversary Eve, i.e., $t = E$. To evaluate performance analysis, we

use P_F to denote the probability of false alarm (i.e., a signal transmitted by legitimate transmitter (Alice) is mistakenly regarded as unauthentic). We use P_D to denote the probability of detection (i.e., a signal originated from illegitimate transmitter (Eve) is successfully judged as unauthentic).

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 paper considers that Bob receives two messages (i.e., frames) at time $k-1$ and time k. The first one 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 authentications it enables a receiver to quickly differentiate between legitimate and illegitimate transmitters without complete higher-layer processing. In this paper, 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 $\hat{\mathbf{h}}_t[k]$ 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 the previous channel estimations with x_m representing the m^{th} component, i.e., $\mathbf{x} = \hat{\mathbf{h}}_t[k] - \hat{\mathbf{h}}_A[k-1]$, 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. Based on Lemma 1, we can explore the distribution of $\mathbf{x}$ on the two hypotheses. Using (2) and (9) on $\mathcal{H}_0$, we have

$$\begin{aligned}\mathbf{x} &= \mathbf{h}_A[k] - \mathbf{h}_A[k-1] + \boldsymbol{\epsilon}_A[k-1] - \boldsymbol{\epsilon}_A[k] \\ &= (\alpha-1)\mathbf{h}_A[k-1] + \sqrt{1-\alpha^2}\mathbf{e}_A[k] + \boldsymbol{\epsilon}_A[k-1] - \boldsymbol{\epsilon}_A[k].\end{aligned} \tag{12}$$

From (12), we can see that $\mathbf{x}$ is a zero-mean complex Gaussian random vector with covariance matrix $2(1-\alpha)\mathbf{R}_A + 2(\mathbf{R}_A - p\mathbf{R}_A\mathbf{R}_{\mathbf{y}_{B,A}}^{-1}\mathbf{R}_A)$. This is because $\mathbf{h}_A$, $\mathbf{e}_A$ and $\boldsymbol{\epsilon}_A$ are mutually independent zero-mean complex Gaussian random vectors.

Similarly, $\mathbf{x}$ on $\mathcal{H}_1$ can be written as

$$\mathbf{x} = \mathbf{h}_E[k] - \mathbf{h}_A[k-1] + \boldsymbol{\epsilon}_A[k-1] - \boldsymbol{\epsilon}_E[k]. \tag{13}$$

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 $\mathcal{H}_1$ is also a zero-mean complex Gaussian random vector with covariance $2\mathbf{R}_A - p\mathbf{R}_A\mathbf{R}_{\mathbf{y}_{BA}}^{-1}\mathbf{R}_A + 2\mathbf{R}_E - p\mathbf{R}_E\mathbf{R}_{\mathbf{y}_{BE}}^{-1}\mathbf{R}_E$. Based on (12) and (13), $\mathbf{C}_i$ can be expressed as (14).

We can see from (14a) and (14b) that $\mathbf{C}_1$ can be decomposed as

$$\mathbf{C}_1 = \mathbf{C}_0 + \mathbf{K}. \tag{15}$$

where $\mathbf{K}$ is

$$\mathbf{K} = 2\mathbf{R}_E - p\mathbf{R}_E\mathbf{R}_{\mathbf{y}_{BE}}^{-1}\mathbf{R}_E + p\mathbf{R}_A\mathbf{R}_{\mathbf{y}_{BA}}^{-1}\mathbf{R}_A - 2(1-\alpha)\mathbf{R}_A. \tag{16}$$

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 [33, Lemma 2.3] as $\Delta\mathbf{Q} = \mathbf{C}_0^{-1}\mathbf{K}\mathbf{C}_1^{-1}$, and then the LRT for the hypothesis test in (11) can be summarized in the following lemma.

Lemma 2. *The LRT for the hypothesis test in* (11) *can be written as*

$$\mathcal{L}(\mathbf{x}) = \mathbf{x}^H\Delta\mathbf{Q}\mathbf{x} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \delta, \tag{17}$$

where $\mathcal{L}(\mathbf{x})$ is sufficient statistic and δ is a decision threshold.

Proof: See Appendix A. ∎

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 2. *To meet extreme data demand growth, it is a promising solution for future wireless systems (e.g., 5G network) 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, and eventually even to lack of fast fading in line of sight dominated cases. The proposed scheme utilizes location-specific channels to authenticate transmitters. Therefore, slower fading or without fading might contribute to improving the authentication performance. This will be verified by numerical results in Section V-C.*

Remark 3. *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 [34] 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 legitimate and illegitimate transmitters, by taking aggregate residual hardware impairments into account. As shown in Section V-C, less fluctuation in channel gain (i.e., tending to hardening) will lead to better authentication performance.*

IV. Modeling of Probabilities for False Alarm and Detection

In this section, we first explore the behaviors of the LRT in (17) for diverse channel covariance models, and then utilize these behavior results to derive analytical expressions of P_F and P_D for the proposed scheme.

According to Section II-B, the channel for the same transmitter-receiver pair can be either spatially independent (uncorrelated) or correlated. Against this backdrop, we need

$$\mathbf{C}_i = \begin{cases} 2(1-\alpha)\mathbf{R}_A + 2(\mathbf{R}_A - p\mathbf{R}_A\mathbf{R}_{\mathbf{y}_{B,A}}^{-1}\mathbf{R}_A), & i=0, \quad (14a) \\ 2\mathbf{R}_A - p\mathbf{R}_A\mathbf{R}_{\mathbf{y}_{BA}}^{-1}\mathbf{R}_A + 2\mathbf{R}_E - p\mathbf{R}_E\mathbf{R}_{\mathbf{y}_{BE}}^{-1}\mathbf{R}_E, & i=1. \quad (14b) \end{cases}$$

to analyze each case in detail to find analytical expressions for P_F and P_D.

A. Spatially independent Channel Components

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

When the temporally correlated channel components are spatially IID (i.e., spatio-temporal), $\mathbf{R}_t$ can be denoted as $\mathbf{R}_t = \sigma_t^2\mathbf{I}$, where σ_t^2 is the variance of $h_{t,m}$. Then, substituting $\mathbf{R}_t$ into (7), $\mathbf{R}_{\mathbf{y}_{Bt}}$ becomes

$$\mathbf{R}_{\mathbf{y}_{Bt}} = \lambda_{\mathbf{y}_{Bt}}\mathbf{I}, \quad (18)$$

where $\lambda_{\mathbf{y}_{Bt}} = \left(p(1+\kappa_t+\kappa_B)\sigma_t^2 + \sigma_I^2 + \sigma_N^2\right)$. Based on (18), we have the following lemma.

Lemma 3. *When the temporally correlated channel components are spatially IID, $\mathbf{C}_i$ given in* (14) *can be further written as*

$$\mathbf{C}_i = \begin{cases} \lambda_{\mathbf{C}_0}\mathbf{I}, & \text{if} \quad i=0, \\ (\lambda_{\mathbf{C}_0} + \lambda_{\mathbf{K}})\mathbf{I}, & \text{if} \quad i=0. \end{cases} \quad (19)$$

where

$$\lambda_{\mathbf{C}_0} = 2(1-\alpha)\sigma_A^2 + 2(\sigma_A^2 - p\sigma_A^4/\lambda_{\mathbf{y}_{BA}}), \quad (20a)$$
$$\lambda_{\mathbf{C}_1} = \lambda_{\mathbf{C}_0} + \lambda_{\mathbf{K}}, \quad (20b)$$
$$\lambda_{\mathbf{K}} = 2\sigma_E^2 - \frac{p\sigma_E^4}{\lambda_{\mathbf{y}_{BE}}} + \frac{p\sigma_A^4}{\lambda_{\mathbf{y}_{BA}}} - 2(1-\alpha)\sigma_A^2. \quad (20c)$$

Proof: 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 (14), (18), and (15), one can see that $\mathbf{C}_i$ are also diagonal matrices. Substituting $\mathbf{R}_t = \sigma_t^2\mathbf{I}$ and $\mathbf{R}_{\mathbf{y}_{Bt}} = \lambda_{\mathbf{y}_{Bt}}\mathbf{I}$ into (14) yields (19). ∎

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

$$\mathbf{R}_t = \text{diag}[\sigma_{t,1}^2, ..., \sigma_{t,M}^2]. \quad (21)$$

Substituting (21) into (7), $\mathbf{R}_{\mathbf{y}_{Bt}}$ becomes

$$\mathbf{R}_{\mathbf{y}_{Bt}} = \text{diag}[\lambda_{\mathbf{y}_{BA,1}}, ..., \lambda_{\mathbf{y}_{BA,M}}], \quad (22)$$

where $\lambda_{\mathbf{y}_{BA,m}} = \left(p(1+\kappa_t+\kappa_B)\sigma_{A,m}^2 + \sigma_I^2 + \sigma_N^2\right)$. Based on (21) and (22), we have the following lemma.

Lemma 4. *When the temporally correlated channel components are spatially IUV, $\mathbf{C}_i$ given in* (14) *can be written as*

$$\mathbf{C}_i = \text{diag}[\lambda_{\mathbf{C}_i,1}, ..., \lambda_{\mathbf{C}_i,M}], \quad (23)$$

where

$$\lambda_{\mathbf{C}_0,m} = (4-2\alpha)\sigma_{A,m}^2 - \frac{2p\sigma_{A,m}^4}{\lambda_{\mathbf{y}_{BA,m}}}, \quad (24a)$$
$$\lambda_{\mathbf{C}_1,m} = \lambda_{\mathbf{C}_0,m} + \lambda_{\mathbf{K},m}, \quad (24b)$$
$$\lambda_{\mathbf{K},m} = 2\sigma_{E,m}^2 - 2(1-\alpha)\sigma_{A,m}^2 - \frac{p\sigma_{E,m}^4}{\lambda_{\mathbf{y}_{BE,m}}} + \frac{p\sigma_{A,m}^4}{\lambda_{\mathbf{y}_{BA,m}}}. \quad (24c)$$

Proof: 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 (14), (18), and (15), we know that $\mathbf{C}_i$ is also diagonal matrix. Substituting (21) and (22) into (14), we can obtain (23). ∎

Let $a_m = \frac{\lambda_{\mathbf{K},m}}{\lambda_{\mathbf{C}_0,m}+\lambda_{\mathbf{K},m}}$ 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. Further let δ denote a decision threshold, and $\lambda_{\mathbf{C}_0}$ and $\lambda_{\mathbf{K}}$ are defined in Lemma 3. Based on the above lemmas, we have the following theorem.

Theorem 1. *For the proposed physical layer authentication scheme, P_F under IID and IUV channel components can be given in* (25) *and P_D in* (26)*.*

Proof: See Appendix B. ∎

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 $\mathbf{\Delta Q}$ is a diagonal matrix (due to $\mathbf{C}_i$ being diagonal matrices). 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.

B. Spatially Correlated Channel Components

In practice, 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 [6]; and third, for antenna, there exists spatially dependent patterns when setting short antenna space and large angular spread, causing channels between adjacent antennas spatially correlated [6], [35], [36]. Therefore, for massive MIMO systems, spatial correlation properties of channels between adjacent antennas always exist. We generate channel covariance matrix $\mathbf{R}_t$ ($t = \{A, E\}$) via exponential correlation model in [35]. In fact, it is expressed by a $M \times M$ complex Toeplitz matrix [33]. That is,

$$\mathbf{R}_t = \sigma_t^2 \begin{bmatrix} 1 & \rho_t^* & \cdots & (\rho_t^*)^{M-1} \\ \rho_t & 1 & \cdots & (\rho_t^*)^{M-2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_t^{M-1} & \rho_t^{M-2} & \cdots & 1 \end{bmatrix}, \quad (27)$$

$$P_F = \begin{cases} \mathbf{\Gamma}_{\chi^2_{2M}}\left(\left(\frac{\lambda_{\mathbf{C}_0}}{\lambda_{\mathbf{K}}}+1\right)\delta\right), & \text{if IID}, \quad (25a) \\ \sum_{m=1}^{M}\left[\prod_{\substack{i=1\\ i\neq m}}^{M}\frac{a_m}{a_m-a_i}\right]\exp\left(-\frac{\delta}{a_m}\right), & \text{if IUV}. \quad (25b) \end{cases}$$

$$P_D = \begin{cases} \mathbf{\Gamma}_{\chi^2_{2M}}\left(\frac{\lambda_{\mathbf{C}_0}}{\lambda_{\mathbf{K}}}\delta\right), & \text{if IID}, \quad (26a) \\ \sum_{m=1}^{M}\left[\prod_{\substack{i=1\\ i\neq m}}^{M}\frac{c_m}{c_m-c_i}\right]\exp\left(-\frac{\delta}{c_m}\right), & \text{if IUV}. \quad (26b) \end{cases}$$

where σ_t^2 is arbitrary scaling factor and ρ_t ($0 < |\rho_t| \leq 1$) is correlation coefficient between adjacent antennas. When $|\rho_t| = 0$, channel components are spatially uncorrelated. Note that the eigenvalue spread in $\mathbf{R}_t$ depends on $|\rho_t|$. Hence, we need to consider different $|\rho_t|$ to derive exact expressions for P_F and P_D.

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

$$\mathbf{R}_t = \text{diag}[M\sigma_t^2, 0, ..., 0], \quad (28)$$

and then $\mathbf{R}_A = \text{diag}[M\sigma_A^2, 0, ..., 0]$ and $\mathbf{R}_E = \text{diag}[M\sigma_E^2, 0, ..., 0]$. Substituting $\mathbf{R}_A$ and $\mathbf{R}_E$ into (7) yields

$$\lambda_{\mathbf{y}_{Bt,1}} = p(1+\kappa_t)M\sigma_t^2 + p\kappa_B\sigma_t^2 + \sigma_I^2 + \sigma_N^2. \quad (29)$$

Combining (14) and (16), we can get the following lemma.

Lemma 5. *When the temporally correlated channel components are fully correlated in space (i.e., $|\rho_t| = 1$), $\mathbf{C}_i$ given in* (14) *becomes*

$$\mathbf{C}_i = \text{diag}[\lambda_{\mathbf{C}_i,1}, 0, ..., 0], \quad (30)$$

where

$$\lambda_{\mathbf{C}_0,1} = (4-2\alpha)M\sigma_A^2 - \frac{2pM^2\sigma_A^4}{\lambda_{\mathbf{y}_{BA,1}}}, \quad (31a)$$

$$\lambda_{\mathbf{C}_1,1} = \lambda_{\mathbf{C}_0,1} + \lambda_{\mathbf{K},1}, \quad (31b)$$

$$\lambda_{\mathbf{K},1} = M\left(2\sigma_E^2 - 2(1-\alpha)\sigma_A^2 - \frac{pM\sigma_E^4}{\lambda_{\mathbf{y}_{BE,1}}} + \frac{pM\sigma_A^4}{\lambda_{\mathbf{y}_{BA,1}}}\right). \quad (31c)$$

Proof: When the temporally correlated channel components are fully correlated in space, i.e., $|\rho_t| = 1$, according to (28) and (14), $\mathbf{C}_i$ have one non-zero eigenvalue and $M-1$ zero eigenvalues. Combining (28) and (29), we can obtain (30). ■

When $0 < |\rho_t| < 1$, the eigenvalues of $\mathbf{R}_t$ are distinct and can be found numerically. Let the eigendecomposition of $\mathbf{R}_t$ be

$$\mathbf{R}_t = \mathbf{u}_t\mathbf{\Lambda}_t\mathbf{u}_t^H, \quad (32)$$

where $\mathbf{u}_t$ is an $M \times M$ matrix [37]; and $\mathbf{\Lambda}_t = \text{diag}[\lambda_{t,1}, ..., \lambda_{A,M}]$ with $\lambda_{t,m}$ denoting the m^{th} eigenvalue of $\mathbf{R}_t$. From (7), we can see that the eigendecomposition of $\mathbf{R}_{\mathbf{y}_{Bt}}$ is

$$\mathbf{R}_{\mathbf{y}_{Bt}} = \mathbf{u}_t\mathbf{\Lambda}_{\mathbf{y}_{Bt}}\mathbf{u}_t^H, \quad (33)$$

where $\mathbf{\Lambda}_{\mathbf{y}_{Bt}} = \text{diag}[\lambda_{\mathbf{y}_{Bt,1}}, ..., \lambda_{\mathbf{y}_{Bt,M}}]$ with

$$\lambda_{\mathbf{y}_{Bt,m}} = p(1+\kappa_t)\lambda_{t,m} + p\kappa_B\sigma_t^2 + \sigma_I^2 + \sigma_N^2. \quad (34)$$

To analyze the behavior of the LRT defined in (17) 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 carry out eigendecomposition for $\mathbf{C}_0$, that is,

$$\mathbf{C}_0 = \mathbf{u}_A\mathbf{\Lambda}_0\mathbf{u}_A^H, \quad (35)$$

where $\mathbf{\Lambda}_0 = \text{diag}[\lambda_{\mathbf{C}_0,1}, ..., \lambda_{\mathbf{C}_0,M}]$ with $\lambda_{\mathbf{C}_0,m}$ representing the m^{th} eigenvalue of $\mathbf{C}_0$. It can be easily seen from (27) 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 $\mathcal{H}_0$ 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 $\mathcal{H}_0$ is $\mathbf{I}$. On $\mathcal{H}_1$, its covariance matrix is denoted by

$$\mathbf{R}_{1\mathbf{w}} = \mathbb{E}\{\mathbf{x}_\mathbf{w}\mathbf{x}_\mathbf{w}^H|\mathcal{H}_1\} = \mathbf{w}^H\mathbf{C}_1\mathbf{w} = \mathbf{w}^H\mathbf{D}\mathbf{w} + \mathbf{I}. \quad (36)$$

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 $\mathbf{w}^H$ can not decorrelate $\mathbf{D}$. Therefore, we now need to carry out an eigendecomposition of $\mathbf{R}_{\mathbf{Dw}}$:

$$\mathbf{R}_{\mathbf{Dw}} = \mathbf{u}_{\mathbf{Dw}}\mathbf{\Lambda}_{\mathbf{Dw}}\mathbf{u}_{\mathbf{Dw}}^H, \quad (37)$$

where $\mathbf{u}_{\mathbf{Dw}}$ is an $M \times M$ modal matrix; and $\mathbf{\Lambda}_{\mathbf{Dw}} = \text{diag}[\lambda_{\mathbf{Dw},1}, ..., \lambda_{\mathbf{Dw},M}]$ 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 a full rank matrix. Hence, we augment the eigenvectors if its rank is not M. The eigendecomposition of $\mathbf{R}_{1\mathbf{w}}$ is

$$\mathbf{R}_{1\mathbf{w}} = \mathbf{u}_{\mathbf{Dw}}[\Lambda_{\mathbf{Dw}} + \mathbf{I}]\mathbf{u}_{\mathbf{Dw}}^H = \mathbf{u}_{\mathbf{Dw}}\mathbf{\Lambda}_{1\mathbf{w}}\mathbf{u}_{\mathbf{Dw}}^H, \quad (38)$$

where

$$\mathbf{\Lambda}_{1\mathbf{w}} = \text{diag}[\lambda_{\mathbf{Dw},1} + 1, ..., \lambda_{\mathbf{Dw},M} + 1]. \quad (39)$$

Let $\lambda_{\mathbf{wu},m} = \frac{\lambda_{\mathbf{Dw}_m}}{\lambda_{\mathbf{Dw},m}+1}$, and then based on the above lemmas, P_F and P_D under spatially correlated channel are summarized in the following theorem.

Theorem 2. *For the proposed physical layer authentication scheme, P_F under spatially correlated channel components can be given in* (40) *and P_D in* (41)*.*

Proof: See Appendix C. ∎

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, utilizing eigendecomposition and diagonalizing operations we can transform an arbitrary channel covariance matrix model to the case in which $\mathbf{\Delta Q}$ 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) can affect authentication performance.

C. 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.

$$\begin{aligned}\mathcal{L}(\mathbf{x}) &= \frac{1}{\sigma_N^2 + \sigma_I^2} \sum_{m=1}^{M} |x_m|^2 \\ &= \frac{1}{\sigma_N^2 + \sigma_I^2} \sum_{m=1}^{M} |\hat{h}_{t,m}[k] - \hat{h}_{A,m}[k-1]|^2 \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \delta. \quad (42)\end{aligned}$$

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

V. Numerical Results

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

A. System Parameters and Simulation Setting

System parameters that determine authentication performance (P_F, P_D) are listed in Table II. 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 Alice and Eve is defined as $\kappa = \frac{\kappa_E}{\kappa_A}$. According to the EVM ranges introduced in Section II-C, we consider four typical levels of impairments: κ_A, κ_B, $\kappa_E \in \{0^2, 0.05^2, 0.1^2, 0.15^2\}$. 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 is spatial correlation coefficient between adjacent antennas for $\mathbf{h}_A$ and ρ_E is that for $\mathbf{h}_E$. 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 of [38] and the exponential correlation model of [35] are exploited to generate time-varying MIMO channels and covariance matrices of such channels, respectively. The quantity of temporal correlation of underlying channels depends on normalized Doppler frequency, which is determined by the speed of entities 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 $\alpha = 0.8$) [39]. For Monte-Carlo experiments, 10^5 independent trials are conducted to obtain average results.

B. Models of P_F and P_D 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. 2. Fig. 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. 2, 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. 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 the smallest one can approach 0.3 under the same channel covariance matrix. This clearly reveals that hardware impairments greatly deteriorate authentication performance.

Another important observation from Fig. 2 is that the choice of covariance model has a large impact on performance. The reason is that: for the spatially uncorrelated covariance model (Fig. 2(a) and Fig. 2(b)), we have $2M$ real observations of channel component estimation; decreasing ρ results in lower spatial correlation and thus improves P_D; and for the spatially correlated covariance model (Fig. 2(c)), we have no more than $2M$ real observations, especially when $\rho = 1$ we only have two real observations. It is proved in [40] that the quantity of spatial correlation determines the number of observations for channel component estimation and this is consistent with our results.

C. Authentication Performance Analysis

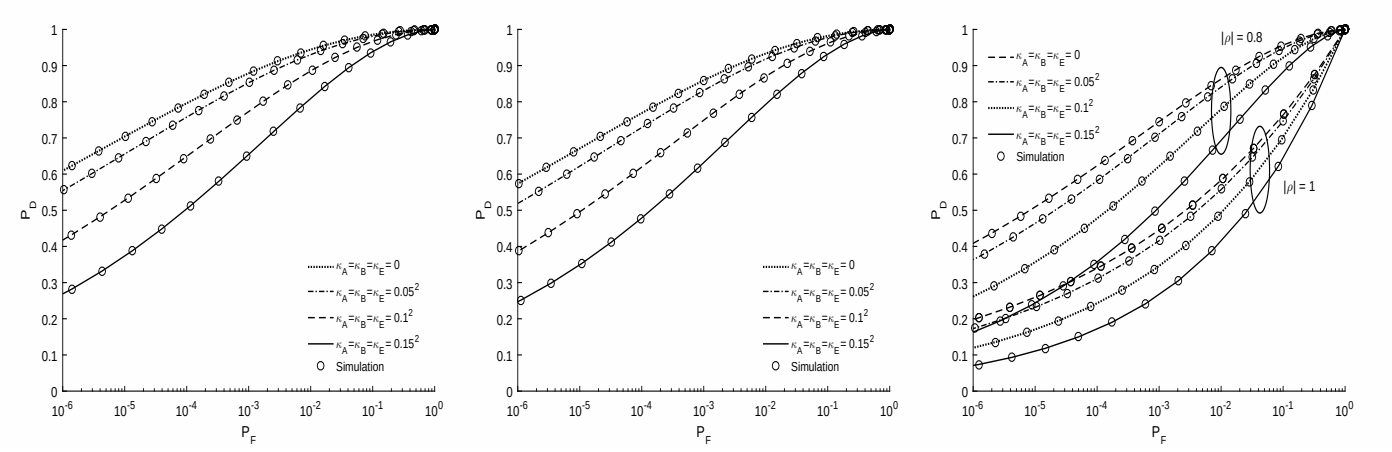
Based on theoretical models for P_F and P_D, we explore how the system parameters (e.g., κ, SINR, γ, α, and M) affect the authentication performance under diverse channel covariance matrix models. Meanwhile, we also examine the

$$P_F = \begin{cases} \exp\left(-\delta\left(1+\dfrac{\lambda_{\mathbf{C}_0,1}}{\lambda_{\mathbf{K},1}}\right)\right), & \text{if} \quad |\rho_A| = |\rho_E| = 1, \quad (40a) \\ \displaystyle\sum_{m=1}^{M}\left[\prod_{\substack{i=1 \\ i\neq m}}^{M}\frac{\lambda_{\mathbf{wu},m}}{\lambda_{\mathbf{wu},m}-\lambda_{\mathbf{wu},i}}\right]\exp\left(-\frac{\delta}{\lambda_{\mathbf{wu},m}}\right), & \text{if} \quad 0<|\rho_{\mathrm{A}}|,|\rho_{\mathrm{E}}|<1. \quad (40b) \end{cases}$$

$$P_D = \begin{cases} \exp\left(-\dfrac{\delta\lambda_{\mathbf{C}_0,1}}{\lambda_{\mathbf{K},1}}\right), & \text{if} \quad |\rho_A| = |\rho_E| = 1, \quad (41a) \\ \displaystyle\sum_{m=1}^{M}\left[\prod_{\substack{i=1 \\ i\neq m}}^{M}\frac{\lambda_{\mathbf{Dw},m}}{\lambda_{\mathbf{Dw},m}-\lambda_{\mathbf{Dw},i}}\right]\exp\left(-\frac{\delta}{\lambda_{\mathbf{Dw},m}}\right), & \text{if} \quad 0<|\rho_{\mathrm{A}}|,|\rho_{\mathrm{E}}|<1. \quad (41b) \end{cases}$$

TABLE II
SYSTEM PARAMETERS

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

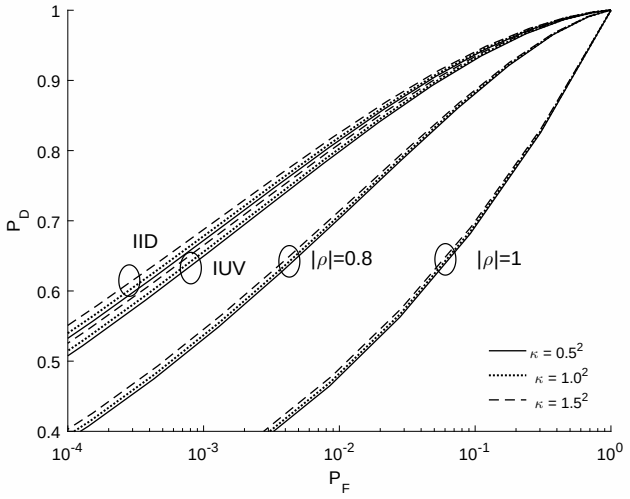


(a) IID (b) IUV (c) Spatial correlation

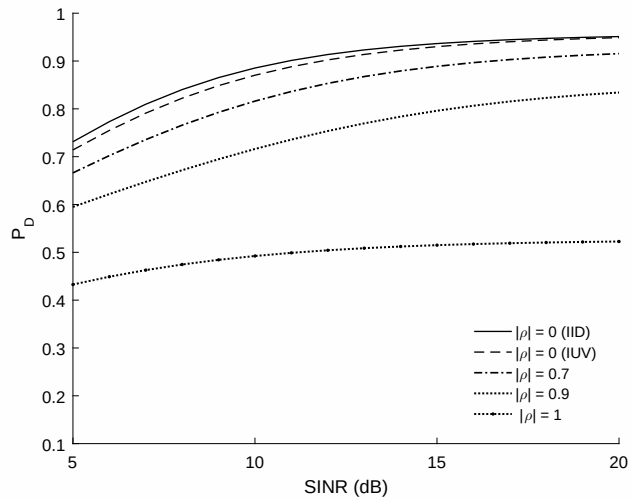
Fig. 2. ROC curves of the proposed scheme with the settings of ($\gamma = 0$ dB, $\kappa = 1.0^2$, $M = 5$, SINR $= 10$ dB, and $\alpha = 0.9$).

performance under unknown parameters via numerical simulations.

We first explore how κ can impact the performance for both scenarios (i.e., spatially uncorrelated and correlated channel components). We summarize in Fig. 3(a) the ROC curves with some representative values of κ for spatially uncorrelated and correlated channel components. As shown in Fig. 3(a), for all channel covariance matrix models, the performance monotonically improves as κ increases. In particular, when $\kappa = 1.5^2$, we have the best performance of the proposed scheme; 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 suggests that we should choose hardware with smaller level of impairments for secure wireless communications.

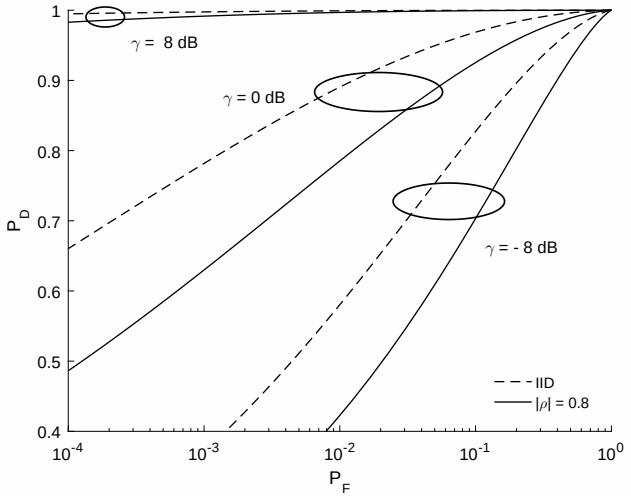


(a) Impact of κ on performance under SINR $= 10$ dB

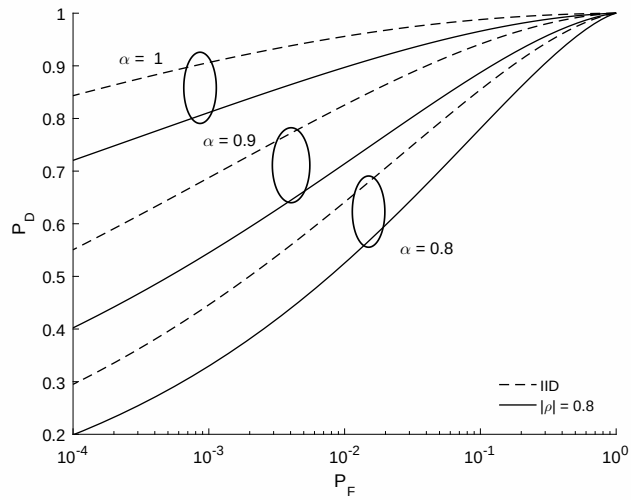


(b) P_D vs. SINR under $\kappa = 1.0^2$ and $P_F = 10^{-2}$

Fig. 3. Authentication performance with the settings of ($\gamma = 0$ dB, $M = 5$, SINR $= 10$ dB, $\alpha = 0.9$).



(a) Impact of γ on ROC curve under $\alpha = 0.9$



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

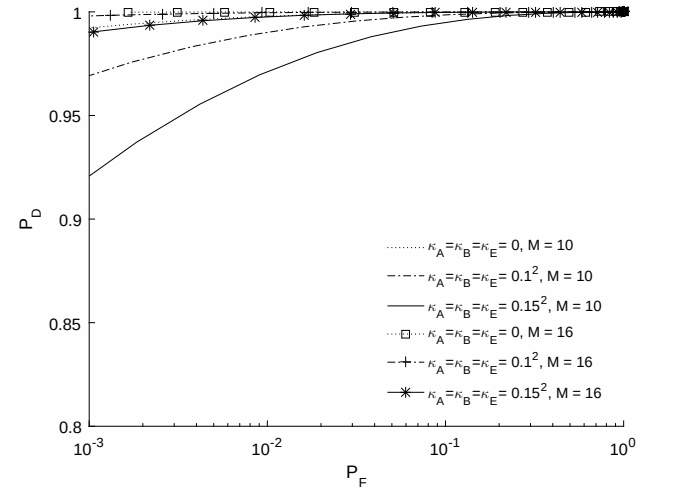
Fig. 4. Impacts of (γ, α) on ROC curve with the settings of (SINR $= 10$ dB, $\kappa = 1.5^2$, $M = 5$).

Next, we investigate the impact of SINR on P_D for a fixed P_F. Fig. 3(b) illustrates how P_D varies with SINR with the settings of ($\gamma = 0$ 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 varying tendencies of P_D for different channel model. In particular, P_D improves with SINR; the curves for spatially uncorrelated channel model (i.e., $|\rho| = 0$ for IID and IUV) have a better slope than that for spatially correlated channel models. For spatially correlated channel model, the curves for $\rho = 0.7$ and $\rho = 0.9$ have the same slope while the curve 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 a 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 network applications, transmit power is limited to a certain level due to energy constraint.

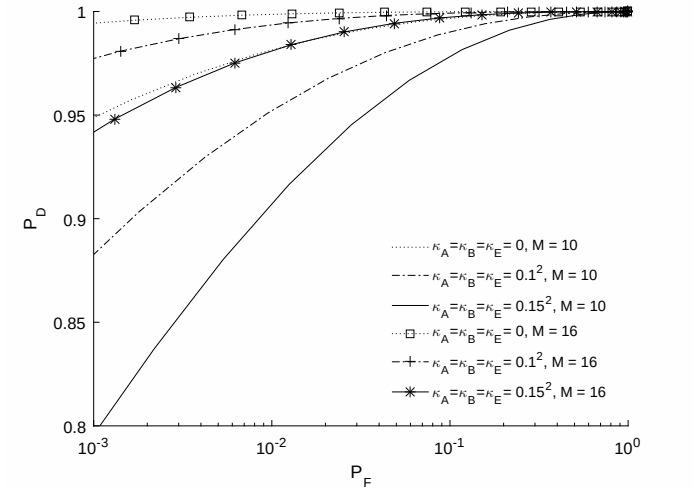
Fig. 4(a) shows how performance varies with $\gamma \in \{-8 \text{ dB}, 0 \text{ dB}, 8 \text{ dB}\}$, given that SINR $= 10$ dB, $\kappa = 1.5^2$, $M = 5$, and $\alpha = 0.9$. It is interesting to see from Fig. 4(a) that for both channel covariance matrix models, the performance monotonically rises as γ increases. More specifically, when γ = 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. 4(b) demonstrates the impact of channel fading status on the 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. 4(b), the authentication performance under case I outperforms that under other cases (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. 5 the impact of $M \in \{10, 16\}$ on the authentication performance under IID channel components

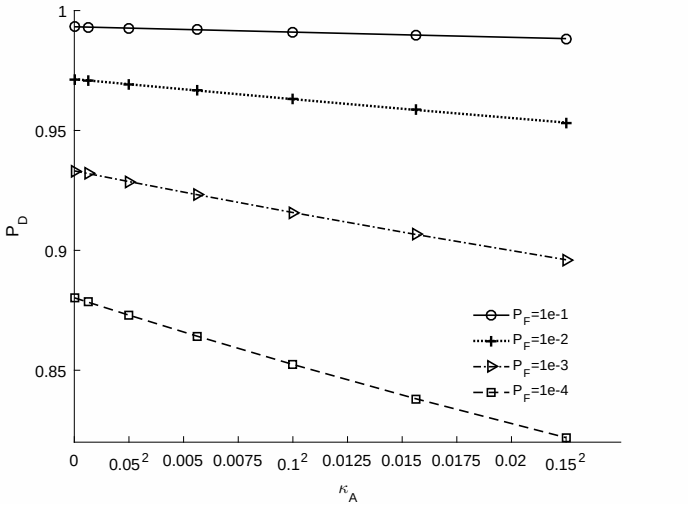


(a) IID

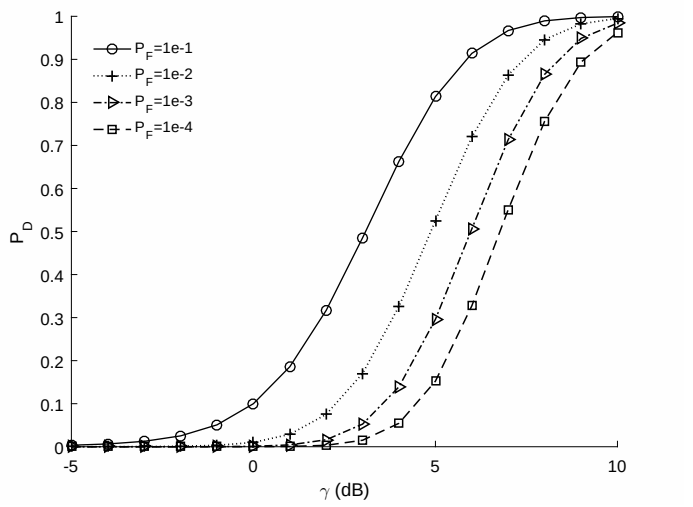


(b) $|\rho| = 0.8$

Fig. 5. Impact of $M \in \{10, 16\}$ on performance, given that $\gamma = 0$ dB, SINR $= 10$ dB, $\alpha = 0.9$, and $\kappa_A = \kappa_B = \kappa_E \in \{0, 0.1^2, 0.15^2\}$.



(a) P_D vs. κ_A under $\kappa = 1.0^2$, $\alpha = 0.9$, $\gamma = 0$ dB



(b) P_D vs. γ under unknown parameter case

Fig. 6. Authentication performance with the settings of (SINR $= 10$ dB, $M = 5$).

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

Furthermore, we explore how P_D varies with κ_A for IID with the settings of ($\kappa = 1.0^2$, SINR $= 5$ dB, $\alpha = 0.9$, $\gamma = 0$ dB, and $M = 5$). Fig. 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. 6(b) via numerical simulations. Fig. 6(b) demonstrates that P_D varies with γ under the unknown parameter case with the settings of (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 the 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 (42) 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 unaware of 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.

VI. Conclusion

We proposed an improved 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, which is demonstrated in the literature. In this paper, 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 individually used to improve the security of massive MIMO systems.

Appendix A
Proof of Lemma 2

Using (14a) and (14b), the probability density functions (PDFs) of $\mathbf{x}$ on the two hypotheses can be written as

$$f(\mathbf{x}|\mathcal{H}_i) = \frac{1}{\pi^M \det(\mathbf{C}_i)} \exp(-\mathbf{x}^H \mathbf{Q}_i^{-1} \mathbf{x}), \quad i = 0, 1. \tag{43}$$

We use $\mathcal{L}_0(\mathbf{x})$ to denote a LRT and δ_0 for threshold. Neyman–Pearson Criterion in [40] leads to a LRT, which can be written as

$$\mathcal{L}_0(\mathbf{x}) \triangleq \frac{f(\mathbf{x}|\mathcal{H}_1)}{f(\mathbf{x}|\mathcal{H}_0)} = \frac{\det(\mathbf{C}_0)}{\det(\mathbf{C}_1)} \frac{\exp(-\mathbf{x}^H \mathbf{Q}_1 \mathbf{x})}{\exp(-\mathbf{x}^H \mathbf{Q}_0 \mathbf{x})} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \delta_0. \tag{44}$$

Taking logarithms and retaining only data-dependent terms, we can obtain logarithmic LRT as (17).

Appendix B
Proof of Theorem 1

Proof of Theorem 1 for IID: Using (14), we can obtain the LRT in (17) under IID channel components as

$$\mathcal{L}(\mathbf{x}) \triangleq \frac{\lambda_{\mathbf{D}}}{\lambda_{\mathbf{C}_0}(\lambda_{\mathbf{C}_0} + \lambda_{\mathbf{D}})} \sum_{m=1}^{M} |x_m|^2 \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \delta. \tag{45}$$

Based on the above results, we now derive expressions for P_F and P_D under IID channel components. Since $x_m/\sqrt{\lambda_{\mathbf{C}_0}}$ on $\mathcal{H}_0$ 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, $\sum_{m=1}^{M} |x_m/\sqrt{\lambda_{\mathbf{C}_0}}|^2 \sim \chi^2_{2M}$. Under IID channel components, P_F can be given by

$$\begin{aligned} P_F &= \mathbf{Pr}(\mathcal{L}(\mathbf{x}) > \delta|\mathcal{H}_0) \\ &= \mathbf{Pr}\left(\sum_{m=1}^{M} \left| \frac{x_m}{\sqrt{\lambda_{\mathbf{C}_0}}} \right|^2 > \left(\frac{\lambda_{\mathbf{C}_0}}{\lambda_{\mathbf{D}}} + 1 \right) \delta | \mathcal{H}_0 \right). \end{aligned} \tag{46}$$

Substituting the right-tail probability function of chi-square random variable into (46) yields (25a).

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

$$P_D = \mathbf{Pr}(\mathcal{L}(\mathbf{x}) > \delta|\mathcal{H}_1). \tag{47}$$

Substituting the right-tail probability function of chi-square random variable into (47) yields (26a) under IID.

Proof of Theorem 1 for IUV: Using (23), we can obtain the LRT in (17) under IUV channel components as

$$\mathcal{L}(\mathbf{x}) \triangleq \frac{\lambda_{\mathbf{D}_m}}{\lambda_{\mathbf{C}_0,m}(\lambda_{\mathbf{C}_0,m} + \lambda_{\mathbf{D}_m})} \sum_{m=1}^{M} |x_m|^2 \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \delta. \tag{48}$$

Under $\mathcal{H}_0$, the characteristic function of $|x_m|^2$ is

$$\begin{aligned} M_{|x_m|^2|\mathcal{H}_0}(j\omega) &= \mathbb{E}\{\exp(j\omega|x_m|^2)|\mathcal{H}_0\} \\ &= \int_{-\infty}^{\infty} \frac{\exp\left((j\omega - \frac{1}{\lambda_{\mathbf{C}_0,m}})|x_m|^2 \right)}{\pi \lambda_{\mathbf{C}_0,m}} dx_m \\ &= (1 - j\omega\lambda_{\mathbf{C}_0,m})^{-1}. \end{aligned} \tag{49}$$

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 $\mathcal{H}_0$ as

$$M_{\mathcal{L}(\mathbf{x})|\mathcal{H}_0}(j\omega) = \prod_{m=1}^{M} (1 - j\omega a_m)^{-1}. \tag{50}$$

We use a partial fraction expansion [41] of (50) to obtain

$$M_{\mathcal{L}(\mathbf{x})|\mathcal{H}_0}(j\omega) = \sum_{m=1}^{M} b_m (1 - j\omega a_m)^{-1}, \tag{51}$$

where

$$b_m = \prod_{\substack{i=1 \\ i \neq m}}^{M} \frac{a_m}{a_m - a_i}. \tag{52}$$

As observed from (51), the characteristic function of $\mathcal{L}(\mathbf{x})$ is a weighted superposition of exponentially distributed characteristic functions. After taking inverse Fourier transform for (51), we can see that the PDF of $\mathcal{L}(\mathbf{x})$ is also a weighted superposition of exponentially distributed PDFs [40], that is,

$$f(\mathcal{L}(\mathbf{x})|\mathcal{H}_0) = \sum_{m=1}^{M} \frac{b_m}{a_m} \exp\left(-\frac{\mathcal{L}}{a_m} \right). \tag{53}$$

Under IUV channel components, P_F can be obtained by integrating the following formula as follows

$$P_F = \mathbf{Pr}(\mathcal{L}(\mathbf{x}) > \delta|\mathcal{H}_0) = \int_{\delta}^{\infty} f(\mathcal{L}(\mathbf{x})|\mathcal{H}_0) d\mathcal{L}(\mathbf{x}). \tag{54}$$

Then, substituting (51) into (54) yields (25b).

Similarly, the characteristic function of $\mathcal{L}(\mathbf{x})$ on $\mathcal{H}_1$ can be expressed as follows

$$M_{\mathcal{L}(\mathbf{x})|\mathcal{H}_1}(j\omega) = \prod_{m=1}^{M} (1 - j\omega c_m)^{-1}. \quad (55)$$

Following the same steps, we can obtain P_D as (26b) integrating the following formula (56).

$$P_D = \mathbf{Pr}(\mathcal{L}(\mathbf{x}) > \delta|\mathcal{H}_1) = \int_{\delta}^{\infty} f(\mathcal{L}(\mathbf{x})|\mathcal{H}_1) d\mathcal{L}(\mathbf{x}). \quad (56)$$

APPENDIX C
PROOF OF THEOREM 2

When $|\rho_A| = |\rho_E| = 1$, the characteristic functions of $\mathcal{L}(\mathbf{x})$ on the two hypotheses can be obtained by using Lemma 31 as

$$M_{\mathcal{L}(\mathbf{x})|\mathcal{H}_0}(j\omega) = \left(1 - j\omega \frac{\lambda_{2_1}}{\lambda_{0_1} + \lambda_{2_1}}\right)^{-1}, \quad (57)$$

$$M_{\mathcal{L}(\mathbf{x})|\mathcal{H}_1}(j\omega) = \left(1 - j\omega \frac{\lambda_{2_1}}{\lambda_{0_1}}\right)^{-1}. \quad (58)$$

We can see that $\mathcal{L}(\mathbf{x})$ is an exponentially distributed random variable. Similar to the derivations of (25b) and (26b), we can obtain (40a) and (41a).

When $0 < |\rho_{\mathrm{t}}| < 1$, using $\mathbf{u}_{\mathbf{Dw}}^H$, we transform from $\mathbf{x_w}$ to $\mathbf{x_{wu}} = [x_{\mathbf{wu},1} \cdots x_{\mathbf{wu},M}]^T$, which is denoted by

$$\mathbf{x_{wu}} = \mathbf{u}_{\mathbf{Dw}}^H \mathbf{x_w} = \mathbf{u}_{\mathbf{Dw}}^H \mathbf{w}^H \mathbf{x}. \quad (59)$$

Now, we explore the covariance matrices of $\mathbf{x_{wu}}$ on the two hypotheses. On $\mathcal{H}_0$, its covariance matrix is given by

$$\begin{aligned} \mathrm{Cov}(\mathbf{x_{wu}}|\mathcal{H}_0) &= \mathbb{E}\{\mathbf{x_{wu}}\mathbf{x_{wu}}^H|\mathcal{H}_0\} = \mathbb{E}\{\mathbf{u}_{\mathbf{Dw}}^H \mathbf{x_w} \mathbf{x_w}^H \mathbf{u_{Dw}}|\mathcal{H}_0\} \\ &= \mathbf{u}_{\mathbf{Dw}}^H \mathbf{I} \mathbf{u_{Dw}} = \mathbf{I}. \end{aligned} \quad (60)$$

Similarly, on $\mathcal{H}_1$ the covariance matrix of $\mathbf{x_{wu}}$ is given by

$$\begin{aligned} \mathrm{Cov}(\mathbf{x_{wu}}|\mathcal{H}_1) &= \mathbb{E}\{\mathbf{x_{wu}}\mathbf{x_{wu}}^H|\mathcal{H}_1\} = \mathbf{u}_{\mathbf{Dw}}^H \mathbf{R}_{1\mathbf{w}} \mathbf{u_{Dw}} \\ &= \mathbf{u}_{\mathbf{Dw}}^H \mathbf{u_{Dw}} [\Lambda_{\mathbf{Dw}} + \mathbf{I}] \mathbf{u}_{\mathbf{Dw}}^H \mathbf{u_{Dw}} \\ &= \Lambda_{\mathbf{Dw}} + \mathbf{I}. \end{aligned} \quad (61)$$

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

$$\begin{aligned} \mathbf{Q_{wu}} &\triangleq \mathrm{diag}\left[\frac{\lambda_{\mathbf{Dw},1}}{\lambda_{\mathbf{Dw},1}+1}, ..., \frac{\lambda_{\mathbf{Dw},M}}{\lambda_{\mathbf{Dw},M}+1}\right] \\ &= \mathrm{diag}\left[\lambda_{\mathbf{wu},1}, ..., \lambda_{\mathbf{wu},M}\right]. \end{aligned} \quad (62)$$

Applying some derivations similar to that in [40, Chapter 3], under spatially correlated channel components, the LRT $\mathcal{L}(\mathbf{x})$ in (17) becomes

$$\begin{aligned} \mathcal{L}(\mathbf{x}) &\triangleq \mathcal{L}(\mathbf{x_{wu}}) = \mathbf{x_{wu}}^H \mathbf{Q_{wu}} \mathbf{x_{wu}} \\ &= \sum_{m=1}^{M} \frac{\lambda_{\mathbf{Dw},m}}{\lambda_{\mathbf{Dw},m}+1} |x_{\mathbf{wu},m}|^2 \\ &= \sum_{m=1}^{M} \lambda_{\mathbf{wu},m} |x_{\mathbf{wu},m}|^2 \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \delta. \end{aligned} \quad (63)$$

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 follows

$$\begin{aligned} P_F &= \mathbf{Pr}(\mathcal{L}(\mathbf{x}) > \delta|\mathcal{H}_0) = \int_{\delta}^{\infty} f(\mathcal{L}(\mathbf{x})|\mathcal{H}_0) d\mathcal{L}(\mathbf{x}) \\ &\triangleq \int_{\delta}^{\infty} f(\mathcal{L}(\mathbf{x_{wu}})|\mathcal{H}_0) d\mathcal{L}(\mathbf{x_{wu}}), \end{aligned} \quad (64)$$

$$\begin{aligned} P_D &= \mathbf{Pr}(\mathcal{L}(\mathbf{x}) > \delta|\mathcal{H}_1) = \int_{\delta}^{\infty} f(\mathcal{L}(\mathbf{x})|\mathcal{H}_1) d\mathcal{L}(\mathbf{x}) \\ &\triangleq \int_{\delta}^{\infty} f(\mathcal{L}(\mathbf{x_{wu}})|\mathcal{H}_1) d\mathcal{L}(\mathbf{x_{wu}}). \end{aligned} \quad (65)$$

Following a similar method as that of in Section IV-A. we can obtain P_F and P_D as (40b) and (41b) for $0 < |\rho_{\mathrm{A}}|, |\rho_{\mathrm{E}}| < 1$.

REFERENCES

[1] 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.

[2] Y. Shen, T. Zhang, Y. Wang, H. Wang, and X. Jiang, "Microthings: A generic IoT architecture for flexible data aggregation and scalable service cooperation," *IEEE Commun. Mag.*, vol. 55, no. 9, pp. 86–93, Sep. 2017.

[3] T. Taleb, K. Samdanis, B. Mada, H. Flinck, S. Dutta, and D. Sabella, "On multi-access edge computing: A survey of the emerging 5G network edge cloud architecture and orchestration," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 3, pp. 1657–1681, thirdquarter 2017.

[4] J. Zhu, D. W. K. Ng, N. Wang, R. Schober, and V. K. Bhargava, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590–3600, Mar. 2010.

[5] ——, "Analysis and design of secure massive MIMO systems in the presence of hardware impairments," *IEEE Trans. Wireless Commun.*, vol. 16, no. 3, pp. 2001–2016, Mar. 2017.

[6] 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.

[7] 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.

[8] G. Yang, Q. Huang, D. S. Wong, and X. Deng, "Universal authentication protocols for anonymous wireless communications," *IEEE Trans. Wireless Commun.*, vol. 9, no. 1, pp. 168–174, Jan. 2010.

[9] 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.

[10] 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.

[11] 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.

[12] 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.

[13] 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.

[14] 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.

[15] 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.

[16] 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.

[17] 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.

[18] 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.

[19] Y. Zheng, S. S. Dhabu, and C. H. Chang, "Securing IoT monitoring device using PUF and physical layer authentication," in *Proc. IEEE ISCAS*, May 2018, pp. 1–5.

[20] 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.

[21] 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.

[22] 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.

[23] 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.

[24] A. Hakkarainen, J. Werner, K. R. Dandekar, and M. Valkama, "Analysis and augmented spatial processing for uplink OFDMA MU-MIMO receiver with transceiver I/Q imbalance and external interference," *IEEE Trans. Wireless Commun.*, vol. 15, no. 5, pp. 3422–3439, May 2016.

[25] O. Raeesi, A. Gokceoglu, Y. Zou, E. Björnson, and M. Valkama, "Performance analysis of multi-user massive MIMO downlink under channel non-reciprocity and imperfect CSI," *IEEE Trans. Commun.*, vol. 66, no. 6, pp. 2456–2471, Jun. 2018.

[26] S. Cui, A. J. Goldsmith, and A. Bahai, "Energy-constrained modulation optimization," *IEEE Tran. Wireless Commun.*, vol. 4, no. 5, pp. 2349–2360, Sep. 2005.

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

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

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

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

[31] 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.

[32] 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.

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

[34] 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/

[35] 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.

[36] E. Björnson, 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.

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

[38] 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.

[39] 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.

[40] 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.

[41] 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.

Pinchang Zhang received his B.S. degree in electronic information engineering from Wuyi University, China, in 2009, and M.S. degree in electronic and communication engineering from Kunming University of Science and Technology, China, in 2012. From 2012-2017, he was worked in Chuzhou University, China. From 2017, he is currently pursuing his Ph.D. degree in Future University Hakodate, Hakodate, Japan. His research interests include wireless security and physical layer authentication.

Tarik Taleb received his B.E. degree (with distinction) in information engineering in 2001, and M.Sc. and Ph.D. degrees in information sciences from Tohoku University, Sendai, Japan, in 2003, and 2005, respectively. He is currently a Professor with the School of Electrical Engineering, Aalto University, Espoo, Finland. He is the founder and the Director of the MOSA!C Lab. He is the Guest Editor-in-Chief for the IEEE JSAC series on network Softwarization and enablers. His research interests lie in the field of architectural enhancements to mobile core networks (particularly 3GPP's), network softwarization & slicing, mobile cloud networking, network function virtualization, software defined networking, mobile multimedia streaming, inter-vehicular communications, and social media networking.

Xiaohong Jiang received his B.S., M.S. and Ph.D degrees in 1989, 1992, and 1999 respectively, all from Xidian University, China. He is currently a full professor of Future University Hakodate, Japan. Before joining Future University, Dr. Jiang was an Associate professor, Tohoku University, from Feb. 2005 to Mar. 2010. Dr. Jiang's research interests include computer communications networks, mainly wireless networks and optical networks, network security, routers/switches design, etc. He has published over 300 technical papers at premium international journals and conferences, which include over 70 papers published in top IEEE journals and top IEEE conferences, like IEEE/ACM Transactions on Networking, IEEE Journal of Selected Areas on Communications, IEEE Transactions on Parallel and Distributed Systems, IEEE INFOCOM.

Bin Wu received his Ph.D. degree in Electrical and Electronic Engineering from the University of Hong Kong (Pokfulam, Hong Kong) in 2007. He worked as a postdoctoral research fellow from 2007-2012 in the ECE Dept. at University of Waterloo (Waterloo, Canada). He is currently a Professor with the College of Intelligence and Computing at Tianjin University (Tianjin, China). His research interests include computer systems and networking as well as communication system design.