Covert Message Authentication in MIMO Communications

Sang Wu Kim, *Senior Member, IEEE*

***Abstract*—We propose a novel covert message authentication technique designed to completely obscure the existence of the digital signature, rendering it secure against integrity attacks. This innovative approach not only thwarts counterfeiting attempts of digital signatures but also effectively evades the scrutiny of potential hackers, thereby protecting the authentication scheme proactively. The core idea involves superimposing the digital signature onto the message and harnessing the capabilities of multiple input multiple output (MIMO) techniques to obfuscate the signature. We demonstrate that the total detection error probability (sum of false alarm and miss detection probability) of the signature approaches unity with an increasing number of transmitter antennas, indicating the undetectability of the signature, regardless of its transmission power. Furthermore, we analyze the impact of this covert verification on the signature decoding error probability and the authenticated message throughput, providing insights into the overall effectiveness of the proposed technique in protecting the authenticity of the message. We also investigate how artificial noise affects the total detection error probability and the authenticated message throughput. Finally, we compare two approaches to signature protection: signature secrecy which prevents eavesdroppers from gaining any meaningful information about the signature and signature covertness which hides the signature transmission.**

***Index Terms*—Covert message authentication, digital signature, MIMO, integrity attack, authenticated message throughput, artificial noise.**

I. Introduction

ENSURING message authenticity is critical in open wireless communication environments, as it encompasses two essential aspects: message integrity and source authenticity. Message integrity ensures that the transmitted message remains unaltered, while source authenticity verifies that the message originates from a trusted sender. Traditionally, digital signatures have been used to guarantee both integrity and authenticity [1]. However, these methods are increasingly vulnerable to evolving technological threats. In particular, the advent of quantum computing poses a severe challenge to current digital signature schemes. Quantum algorithms, such as Shor's algorithm, can derive private keys from public keys, thereby compromising the entire system's security [2], [3].

Received 17 December 2024; revised 16 April 2025 and 29 May 2025; accepted 9 June 2025. Date of publication 26 June 2025; date of current version 25 July 2025. This work was supported by the U.S. National Science Foundation under Grant 2401127 and Grant 2515378. The associate editor coordinating the review of this article and approving it for publication was Dr. Valeria Loscri.

The author is with the Department of Electrical and Computer Engineering, Iowa State University, Ames, IA 50011 USA (e-mail: swkim@iastate.edu).

Digital Object Identifier 10.1109/TIFS.2025.3583434

Moreover, if an attacker collects multiple message-signature (or ciphertext) pairs, they can extract statistical information about the underlying secret key. This gradual reduction in the adversary's uncertainty about the key significantly increases the risk of key recovery attacks [4]. With access to the private key and the signature-generation function, an adversary could then forge valid signatures for arbitrary message-an action commonly referred to as an integrity attack. Given these emerging threats to cryptographic systems, there is an urgent need for more resilient authentication techniques, particularly as quantum computing capabilities advance.

Physical layer authentication (PLA) offers an alternative approach to message authentication by embedding low-power signatures (tags) into communication messages transmitted over noisy channels [5], [6], [7], [8], [9], [10], [11], [12]. This technique aims to enhance security by making it more challenging for adversaries to extract secret keys by reducing signature power. The relationship between the information leaked on the secret key and the signature power, when the content of the message is known (decoded), has been analyzed in [13]. However, reducing the signature power also diminishes the receiver's ability to decode these signatures, which compromises the reliability of message authentication. Furthermore, adversaries aware of the authentication process could exploit it by executing attacks such as replay attacks, man-in-the-middle attacks, or denial-of-service attacks.

This paper introduces a novel paradigm: covert authentication. Unlike conventional authentication methods, covert authentication verifies the authenticity of a message while concealing the existence of the authentication process itself. By masking the presence of authentication, this approach significantly reduces the likelihood of detection by adversaries, thereby proactively protecting the authentication mechanism. The underlying principle is simple yet profound: what cannot be seen cannot be attacked. By remaining undetectable, covert authentication minimizes the risk of key exposure and data tampering, offering enhanced protection in critical applications. These include military communications, intelligence operations, and secure messaging in contested or adversarial environments - contexts where traditional overt authentication may draw unwanted attention and compromise security.

Recent studies have explored the use of multiple-input multiple-output (MIMO) technology to enable covert communications. For example, Wang and Bloch [14] studied the covert capacity of the MIMO AWGN channel under the total variation distance measure, and derived the explicit formula

for the covert transmission rate. Bendary et al. [15] derived the covert capacity of the MIMO AWGN channel under the Kullback-Leibler divergence covertness measure and studied the effect of the number of transmit antennas on the covert capacity, while Bai et al. [16] extend these findings to spatially sparse mmWave massive MIMO channels. While these studies have primarily focused on concealing the existence of data transmission, they largely overlook the challenge of secure and verifiable message authentication in adversarial environments. In contrast, this work leverages the capabilities of MIMO systems not only to support high-rate message transmission without covertness constraints but also to enable the covert embedding of digital signatures. By ensuring that these signatures remain undetectable to unauthorized observers, the proposed scheme strengthens the robustness of the authentication process against targeted attacks. This constitutes a novel integration of physical-layer covert transmission and cryptographic message authentication within a unified framework.

The main contributions of this paper are as follows:

- We propose a novel covert message authentication technique that fully conceals the presence of a digital signature within MIMO communications, ensuring security against integrity attacks. This approach embeds the signature directly into the transmitted message, while using MIMO technology to direct signal energy toward the intended receiver. This minimizes the risk of detection by adversaries while maintaining robust authentication security.
- We show that the total detection error probability (the sum of false alarm and miss detection probabilities) of the signature approaches unity as the number of transmitter antennas increases. This guarantees the undetectability of the signature, regardless of its transmission power. This advancement addresses a limitation of existing PLA methods, which require reduced transmission power for the signature to protect keys, often at the cost of increased decoding errors at the receiver [5], [7], [8], [12].
- We analyze the impact of covert verification on the signature decoding error probability. Our results reveal that the signature decoding error probability approaches zero as the number of transmitter antennas increases, provided that the power allocated to the message remains below a specified threshold. This highlights the method's potential to achieve reliable and undetectable authentication by leveraging a large number of transmitter antennas.
- We introduce and analyze the concept of authenticated message throughput, defined as the average number of correctly received and authenticated message bits per channel use. This metric uniquely integrates communication reliability with message authenticity, which have traditionally been studied independently. Additionally, we investigate the role of artificial noise in enhancing the total detection error probability and the authenticated message throughput in scenarios where the eavesdropper has more antennas than the transmitter.
- We compare two approaches to signature protection: signature secrecy, which employs wiretap codes to keep the signature content confidential, and signature covertness, which conceals the signature's existence. Furthermore, we extend our analysis to eigenbeamforming, where information is transmitted across multiple eigenmodes. We evaluate how multi-stream transmission enhances both the total detection error probability and the authenticated message throughput.

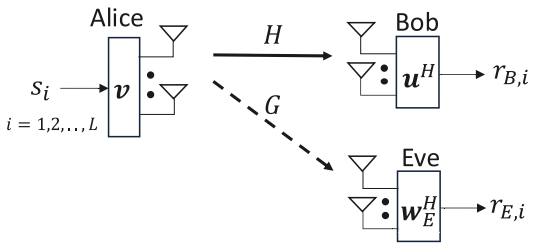

Fig. 1. System model.

These contributions significantly advance the field of secure wireless communications by introducing a covert and robust method for message authentication. The proposed approach provides resilience against a wide range of attacks, addressing the unique challenges of secure and covert authentication in MIMO communication systems.

The paper is structured as follows: Section II outlines the system model. Section III describes the signature detection strategy. Section IV derives the total detection error probability. Section V addresses the signature's decoding outage probability. Section VI focuses on authenticated message throughput. Section VII extends our analysis to include AN. Section VIII compares the signature secrecy and signature covertness. Section IX extends the analysis for the case of eigenbeamforming. Section X concludes with key findings and contributions.

II. System Model

Consider a multiple-input multiple-output (MIMO) communication system involving a transmitter (Alice) and a receiver (Bob), equipped with N_A and N_B antennas, respectively. This communication takes place in the presence of an eavesdropper (Eve), who possesses N_E antennas. The system model is illustrated in Figure 1. The notations and terms used throughout the paper are listed in Table I.

A. Signature Generation and Channel Coding

To ensure the authenticity of a message $\mathbf{m} = e_m(\mathbf{d})$, where $\mathbf{d}$ denotes information data and $e_m(\cdot)$ denotes channel encoding, Alice sends a digital signature $\mathbf{t} = e_t(h(\mathbf{d}, K))$. Here, $h(\cdot)$ is a one-way hash function, K is a private key used by Alice to sign the message, and $e_t(\cdot)$ denotes channel encoding for the signature. The entropy of the signature is typically limited by its size. For example, if the output of a hash function is 128-bit, the maximum entropy of the signature is 128 bits.

We assume that $\mathbf{m}$ and $\mathbf{t}$ are codewords of length L symbols and rate R_m and R_t bits per channel use, respectively. The hash sequence $h(\mathbf{d}, K)$ is typically much shorter than the data sequence $\mathbf{d}$. This inherent difference in length results in a significantly lower code rate for the signature compared to

TABLE I
LIST OF NOTATIONS AND TERMS

α	power allocation factor
$\mathcal{C}$	codebook for $\mathbf{m}$
$\mathbf{d}$	information data
$e_m(\cdot), e_t(\cdot)$	channel encoding
$\mathbf{G}$	channel gain matrix between Alice and Eve
γ	transmit SNR
$\mathbf{g}_w$	$\mathbf{Gv}$
$h(\cdot)$	one-way hash
$\mathbf{H}$	channel gain matrix between Alice and Bob
K	private key
L	code length
λ, λ'	detection threshold
λ_{max}	maximum eigenvalue
Λ, Λ'	log likelihood ratio
$\mathbf{m}$	message
N_A	number of transmitter antennas
N_B	number of receiver antennas
N_E	number of eavesdropper antennas
$\mathbf{n}_{B,i}, \mathbf{n}_{E,i}$	background noise vector
P	signal power
P_F	probability of false alarm
P_M	probability of miss detection
$\mathbf{P}_{o,t}$	signature decoding outage probability
P_{so}	secrecy outage probability
ϕ	power allocation factor in AN scheme
R_m, R_t	transmission rate
σ_n	noise variance
$\mathbf{t}$	digital signature
$\mathbf{u}$	receive weight vector
$\mathbf{v}$	transmit weight vector
W	authenticated message throughput
$\mathbf{x}_i$	transmitted signal
ξ	total detection error probability
$\mathbf{y}_{B,i}, \mathbf{y}_{E,i}$	received signal vector
$\mathbf{z}$	artificial noise vector

the message. This design choice has important implications for the decoding process. Specifically, the lower code rate of the signature provides a crucial advantage: it enables Bob to decode the signature reliably even in the presence of interference and noise.

The code symbols m_i and t_i within $\mathbf{m}$ and $\mathbf{t}$, respectively, are designed to be complex Gaussian distributed with mean zero and variance P, i.e., $m_i, t_i \sim \mathcal{CN}(0, P)$ to maximize the channel capacity [17]. To approximate Gaussian-distributed code symbols in practice, techniques like superposition (e.g., OFDM) [18] or probabilistic amplitude shaping [19] can be employed. For example, higher-amplitude points in a QAM constellation are transmitted less frequently, creating a distribution closer to Gaussian.

We assume that the codebook for the message is known to all users including Eve. But the codebook for the signature is known to Alice and Bob only. This means the message can be decoded, but the signature cannot be decoded by Eve.

B. Transmitted Signal

To conceal the authentication from Eve, we employ a superposition coding technique that embeds the signature within the message. The resulting signal, represented as $\mathbf{s} = (s_1, \ldots, s_L)$, is defined as follows:

$$\mathbf{s} = \begin{cases} \mathbf{m}, & \mathcal{H}_0 \\ \sqrt{\alpha}\mathbf{m} + \sqrt{1-\alpha}\mathbf{t}, & \mathcal{H}_1, \end{cases} \tag{1}$$

where $\mathcal{H}_0$ denotes the null hypothesis that Alice did not send $\mathbf{t}$, $\mathcal{H}_1$ denotes the alternate hypothesis that Alice did send $\mathbf{t}$, and $\alpha \in (0, 1)$ represents the power allocation factor between the message and the signature. In our system, we assume that the signature is included with a probability of π_0 while a fraction of $(1 - \pi_0)$ of transmissions are sent without a signature. Additionally, Alice and Bob share a pre-established secret that allows Bob to identify the specific times or frequencies in which the signature is transmitted. The transmitted signal $\mathbf{x}_i$ is given by

$$\mathbf{x}_i = \mathbf{v}s_i, \tag{2}$$

$i = 1, \ldots, L$, where $\mathbf{v}$ is an $N_A \times 1$ transmit weight vector at Alice.

C. Received Signal at Bob

The received signal at Bob is given by

$$\mathbf{y}_{B,i} = \mathbf{H}\mathbf{x}_i + \mathbf{n}_{B,i}, \tag{3}$$

where

$$\mathbf{H} = \begin{bmatrix} \mathbf{h}_1 \\ \vdots \\ \mathbf{h}_{N_B} \end{bmatrix} \tag{4}$$

denotes the channel gain matrix between Alice and Bob. Each element of $\mathbf{h}_j \in \mathcal{C}^{1\times N_A}$, $j = 1, \cdots, N_B$, is a complex Gaussian random variable with mean zero and variance one (representing Rayleigh fading), and $\mathbf{n}_{B,i} \in \mathcal{C}^{N_B\times 1}$ denotes the background noise vector with $\mathbb{E}[\mathbf{n}_{B,i}] = 0$ and $\mathbb{E}[\mathbf{n}_{B,i}^H \mathbf{n}_{B,i}] = \sigma_n^2 \mathbf{I}_{N_B}$.

We assume that Alice employs transmit beamforming (TB) to focus her signal towards Bob and chooses $\mathbf{v}$ as [20]

$$\mathbf{v} = \frac{\mathbf{H}^H\mathbf{u}}{\|\mathbf{H}^H\mathbf{u}\|}, \tag{5}$$

where $\mathbf{u}$ is an $N_B \times 1$ receive weight vector at Bob.[1] Bob employs maximum ratio combining (MRC) to combine the signals received from each of his antennas. This combination is known as MIMO-MRC [20], [21] or single-mode eigenbeamforming. It requires both Alice and Bob possess knowledge of the channel matrix $\mathbf{H}$. In Section IX, we will consider eigenbeamforming in which information is transmitted over multiple eigenmodes.

After applying $\mathbf{u}$ to the received signal $\mathbf{y}_{B,i}$, Bob obtains

$$r_{B,i} = \mathbf{u}^H\mathbf{y}_{B,i} \tag{6}$$

$$= \|\mathbf{H}^H\mathbf{u}\|s_i + \mathbf{u}^H\mathbf{n}_{B,i}. \tag{7}$$

The SNR of (7), given by $(P/\sigma_n^2)\|\mathbf{H}^H\mathbf{u}\|^2/\|\mathbf{u}\|^2$, is maximized when the vector $\mathbf{u}$ is selected as the eigenvector of $\mathbf{H}\mathbf{H}^H$ associated with the maximum eigenvalue $\lambda_{\max}$ of $\mathbf{H}\mathbf{H}^H$ [20]. Utilizing MIMO-MRC, the SNR of $r_{B,i}$ is given by $\lambda_{\max}\gamma$, where

$$\gamma = P/\sigma_n^2 \tag{8}$$

is the transmit SNR.

[1]We assume a passive Eve whose exact location is unknown to Alice. If Eve's location is known or can be estimated, Alice can incorporate this information into her beamforming design - for example, by placing spatial nulls in Eve's direction - to suppress the authentication signal's energy in her direction.

Achievable Rate and Transmission Rate: Unlike previous endeavors primarily aimed at maximizing the rate of covert message (i.e., signature), our focus lies in maximizing the rate of message $\mathbf{m}$ (without the covertness requirement), while ensuring the concealed transmission of the signature $\mathbf{t}$. This strategy is motivated by the typically shorter length of the signature compared to the message, which eliminates the need to maximize the signature's transmission rate.

To achieve a higher rate for the message than the signature, the signature is decoded first, treating the message as interference. Under this scheme, the achievable rate (bits per channel use) of the signature for Bob is given by:

$$I(\mathbf{t};\mathbf{r}_B) = \log_2\left(1+\frac{(1-\alpha)\lambda_{\max}\gamma}{\alpha\lambda_{\max}\gamma+1}\right). \tag{9}$$

We assume that the signature $\mathbf{t}$ is transmitted at a constant rate of R_t (bits per channel use), reflecting the fixed length of signature in practical applications. In contrast, the message, $\mathbf{m}$, is transmitted at the rate:

$$R_m = \log_2\left(1+\alpha\lambda_{\max}\gamma\right). \tag{10}$$

This transmission rate applies under both hypotheses, $\mathcal{H}_0$ and $\mathcal{H}_1$, to ensure the covertness of the signature. Any variation in R_m based on the signature transmission would enable Eve to detect the signature by merely observing changes in the message transmission rate.

D. Received Signal at Eve

The received signal at Eve is given by

$$\mathbf{y}_{E,i} = \mathbf{G}\mathbf{x}_i + \mathbf{n}_{E,i} = \mathbf{g}_w s_i + \mathbf{n}_{E,i}, \tag{11}$$

where

$$\mathbf{G} = \begin{bmatrix} \mathbf{g}_1 \\ \vdots \\ \mathbf{g}_{N_E} \end{bmatrix} \tag{12}$$

is the channel gain matrix between Alice and Eve and $\mathbf{g}_w = \mathbf{G}\mathbf{v}$ is the effective channel gain for Eve. Each element of $\mathbf{g}_j \in \mathcal{C}^{1\times N_A}$, $j = 1,\cdots,N_E$, is a complex Gaussian random variable with mean zero and variance one, and $\mathbf{n}_{E,i} \in \mathcal{C}^{N_E\times 1}$ is the background noise vector with $\mathbb{E}[\mathbf{n}_{E,i}] = 0$ and $\mathbb{E}[\mathbf{n}_{E,i}^H\mathbf{n}_{E,i}] = \sigma_n^2\mathbf{I}_{N_E}$.

1) Estimation of Channel State by Eve: To perform MRC and maximize her received SNR, Eve must estimate her channel state $\mathbf{g}_w$. Since Eve does not have access to the precoding vector $\mathbf{v}$, she resorts to blind channel estimation based on the statistical structure of her received signals.

Specifically, Eve computes the sample covariance matrix: $\mathbf{R}_y = \frac{1}{L}\sum_{i=1}^{L}\mathbf{y}_{E,i}\mathbf{y}_{E,i}^\dagger$, which converges, as $L\to\infty$, to its expectation: $\mathbb{E}[\mathbf{R}_y] = \mathbf{g}_w\mathbf{g}_w^\dagger P + \sigma_n^2\mathbf{I}$. Hence, the dominant eigenvalue of $\mathbf{R}_y$ is approximately $\|\mathbf{g}_w\|^2P + \sigma_n^2$, and the associated principal eigenvector closely approximates the normalized channel direction $\mathbf{g}_w/\|\mathbf{g}_w\|$. This eigenvector serves as a sufficient estimate of the effective channel direction for applying MRC. Accordingly, Eve selects her combining vector as: $\mathbf{w}_E = \mathbf{g}_w/\|\mathbf{g}_w\|$. In addition, the dominant eigenvalue can be used to estimate the channel gain $\|\mathbf{g}_w\|$.

As the number of observations $L\to\infty$, both the covariance matrix and the resulting eigenvalue estimates converge to their true values, ensuring that Eve's estimate of channel state is asymptotically consistent.

2) MRC: After applying MRC, Eve obtains

$$r_{E,i} = \mathbf{w}_E^H\mathbf{y}_{E,i} \tag{13}$$

$$= \|\mathbf{g}_w\|s_i + z_{E,i}, \tag{14}$$

where $z_{E,i}\sim\mathcal{CN}(0,\sigma_n^2)$, $i=1,\ldots,L$. The resulting received SNR is $\|\mathbf{g}_w\|^2\gamma$, where $\|\mathbf{g}_w\|^2\sim\chi^2(2N_E)$ [22].

3) Remark: The multiple antennas at the transmitter provide no advantage to Eve, as her channel gain, $\|\mathbf{g}_w\|$, is independent of the number of antennas, N_A. However, they do benefit Bob, since the largest eigenvalue, λ_{max}, increases with N_A. This occurs because the transmitter can focus energy towards Bob using TB (see Eq. (5)), making it more difficult for Eve to detect the signature.

III. Signature Detection Strategy

Based on her observation vector $\mathbf{r}_E = (r_{E,1},\ldots,r_{E,L})$, Eve seeks to determine whether a signature $\mathbf{t}$ is present in the received signal. A naive approach might involve using a radiometer (i.e., an energy detector) that bases its decision on the average received signal power, $\|\mathbf{r}_E\|^2/L$. However, under both hypotheses $\mathcal{H}_0$ and $\mathcal{H}_1$, the received signal power converges to the same asymptotic value, namely: $\|\mathbf{r}_E\|^2/L \to \|\mathbf{g}_w\|^2P + \sigma_n^2$ as $L\to\infty$. As a result, energy detection is ineffective in this setting and cannot reliably distinguish between $\mathcal{H}_0$ and $\mathcal{H}_1$, even with long observation periods.

Instead, Eve may employ a more powerful detection method - namely, the likelihood ratio test (LRT) - which is known to be optimal in the Neyman-Pearson sense for maximizing detection probability at a fixed false alarm rate [23]. The LRT is given by:

$$\Lambda = \frac{\mathcal{P}(\mathbf{r}_E \mid \|\mathbf{g}_w\|,\mathbf{m},\mathcal{H}_1)}{\mathcal{P}(\mathbf{r}_E \mid \|\mathbf{g}_w\|,\mathbf{m},\mathcal{H}_0)} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \lambda, \tag{15}$$

where λ is the detection threshold.

The probability density functions under the two hypotheses are given by:

$$\mathcal{P}(\mathbf{r}_E \mid \|\mathbf{g}_w\|,\mathbf{m},\mathcal{H}_1) = \left(\frac{1}{\pi(\sigma_n^2+(1-\alpha)\|\mathbf{g}_w\|^2P)}\right)^L \times \exp\left(-\frac{\|\mathbf{r}_E-\|\mathbf{g}_w\|\sqrt{\alpha}\mathbf{m}\|^2}{\sigma_n^2+(1-\alpha)\|\mathbf{g}_w\|^2P}\right), \tag{16}$$

$$\mathcal{P}(\mathbf{r}_E \mid \|\mathbf{g}_w\|,\mathbf{m},\mathcal{H}_0) = \left(\frac{1}{\pi\sigma_n^2}\right)^L \exp\left(-\frac{\|\mathbf{r}_E-\|\mathbf{g}_w\|\mathbf{m}\|^2}{\sigma_n^2}\right). \tag{17}$$

These expressions represent the conditional distributions of Eve's observation vector under each hypothesis.[2]

The LRT essentially compares the likelihood of the received signal under the two hypotheses, accounting for the fact that under $\mathcal{H}_1$, the presence of the unknown $\mathbf{t}$ increases the

[2]Recall that Eve does not know the codebook of $\mathbf{t}$. As a result, $\mathbf{t}$ is treated as noise and its effect is captured through the increased variance in the likelihood under $\mathcal{H}_1$.

effective noise power. Unlike classical energy detection, the LRT exploits statistical knowledge of the signal structure and noise variance under both hypotheses, making it significantly more powerful in distinguishing between them.

A. Case of Message Decoding Failure

When Eve does not know (fails to decode) the message $\mathbf{m}$, the optimal strategy is to marginalize the likelihood function to eliminate the unknown parameter $\mathbf{m}$ [23]. More precisely, one computes the marginal distribution of $\mathbf{r}_E$:

$$\mathcal{P}(\mathbf{r}_E \mid \|\mathbf{g}_w\|, \mathcal{H}_l) \tag{18}$$

$$= \sum_{i=1}^{|\mathcal{C}|} \mathcal{P}(\mathbf{r}_E \mid \|\mathbf{g}_w\|, \bar{\mathbf{m}}_i, \mathcal{H}_l) Pr(\bar{\mathbf{m}}_i|\mathcal{H}_l), \tag{19}$$

$l = 0, 1$, where $\mathcal{C} = \{\bar{\mathbf{m}}_1, \bar{\mathbf{m}}_2, \cdots, \bar{\mathbf{m}}_{|\mathcal{C}|}\}$ is the codebook for $\mathbf{m}$ and $|\mathcal{C}|$ is the codebook size. Then, Eve performs the marginalized likelihood ratio test (MLRT):

$$\Lambda = \frac{\mathcal{P}(\mathbf{r}_E \mid \|\mathbf{g}_w\|, \mathcal{H}_1)}{\mathcal{P}(\mathbf{r}_E \mid \|\mathbf{g}_w\|, \mathcal{H}_0)} \underset{\mathcal{H}_1}{\overset{\mathcal{H}_0}{\lessgtr}} \lambda. \tag{20}$$

If the codewords $\bar{\mathbf{m}}_1, \bar{\mathbf{m}}_2, \cdots, \bar{\mathbf{m}}_{|\mathcal{C}|}$ are independent and identically distributed and each codeword is sent with equal probability, i.e. $Pr(\bar{\mathbf{m}}_i|\mathcal{H}_l) = 1/|\mathcal{C}|$, then for large $|\mathcal{C}|$ it can be shown that [24]

$$\mathcal{P}(\mathbf{r}_E \mid \|\mathbf{g}_w\|, \mathcal{H}_l) = \left(\frac{1}{\pi(\|\mathbf{g}_w\|^2 P + \sigma_n^2)}\right)^L e^{-\frac{\|\mathbf{r}_E\|^2}{\|\mathbf{g}_w\|^2 P + \sigma_n^2}} \tag{21}$$

for both $\mathcal{H}_0$ and $\mathcal{H}_1$. That is, the marginalized probability distribution of Eve's observation are the same whether the signature is transmitted or not. That is, $\Lambda = 1$ under both hypotheses. Therefore, Eve cannot detect the transmission of the signature if she fails to decode the message $\mathbf{m}$.

B. Case of Message Decoding Success

If Eve knows (succeeds in decoding) the message $\mathbf{m}$, we show that she can detect the transmission of the signature. The LRT in (15) can be reformulated as

$$\Lambda' = \frac{\frac{1}{L}\|\mathbf{r}_E - \|\mathbf{g}_w\|\mathbf{m}\|^2}{\sigma_n^2} - \frac{\frac{1}{L}\|\mathbf{r}_E - \|\mathbf{g}_w\|\sqrt{\alpha}\mathbf{m}\|^2}{\sigma_n^2 + (1-\alpha)\|\mathbf{g}_w\|^2 P} \underset{\mathcal{H}_1}{\overset{\mathcal{H}_0}{\lessgtr}} \lambda', \tag{22}$$

for a detection threshold λ'. It can be shown from the law of large numbers [24] that Λ' converges to Λ_0 under hypothesis $\mathcal{H}_0$ and Λ_1 under hypothesis $\mathcal{H}_1$ as $L \to \infty$, where

$$\Lambda_0 = \frac{2\sqrt{\alpha}(1-\sqrt{\alpha})\|\mathbf{g}_w\|^2 P}{(1-\alpha)\|\mathbf{g}_w\|^2 P + \sigma_n^2} \tag{23}$$

$$\Lambda_1 = \frac{2(1-\sqrt{\alpha})\|\mathbf{g}_w\|^2 P}{\sigma_n^2}. \tag{24}$$

By selecting λ' within the range (Λ_0, Λ_1), Eve can detect the signature transmission if she successfully decodes the message $\mathbf{m}$ and observes the received signal for an extended period. In Section VII, we will explore methods to transmit artificial noise to Eve, making it challenging for her to decode $\mathbf{m}$ and consequently detect the signature.

Estimation of α by Eve: To compute the LRT in (22), Eve must estimate the power allocation factor α used by Alice. When Eve does not know whether $\mathcal{H}_0$ or $\mathcal{H}_1$ is true, estimating α becomes a problem of inference under model uncertainty. This can be solved using a Bayesian inference strategy based on her received observation $\mathbf{r}_E$.

The posterior distribution of α is modeled using a Bayesian mixture that accounts for both hypotheses, $\mathcal{H}_0$ and $\mathcal{H}_1$:

$$\mathcal{P}(\alpha \mid \mathbf{r}_E) = \mathcal{P}(\mathcal{H}_1 \mid \mathbf{r}_E) \cdot \mathcal{P}(\alpha \mid \mathbf{r}_E, \mathcal{H}_1) + \mathcal{P}(\mathcal{H}_0 \mid \mathbf{r}_E) \cdot \delta(\alpha - 1), \tag{25}$$

where $\delta(\cdot)$ is the Dirac delta function, reflecting that $\alpha = 1$ deterministically under $\mathcal{H}_0$.

The posterior probability of $\mathcal{H}_1$ is computed using Bayes' rule:

$$\mathcal{P}(\mathcal{H}_1 \mid \mathbf{r}_E) = \frac{\mathcal{P}(\mathbf{r}_E \mid \mathcal{H}_1) \cdot \mathcal{P}(\mathcal{H}_1)}{\mathcal{P}(\mathbf{r}_E \mid \mathcal{H}_0) \cdot \mathcal{P}(\mathcal{H}_0) + \mathcal{P}(\mathbf{r}_E \mid \mathcal{H}_1) \cdot \mathcal{P}(\mathcal{H}_1)}, \tag{26}$$

with $\mathcal{P}(\mathcal{H}_0) = \mathcal{P}(\mathcal{H}_1) = 0.5$ in the absence of prior information. The complementary posterior probability is then $\mathcal{P}(\mathcal{H}_0 \mid \mathbf{r}_E) = 1 - \mathcal{P}(\mathcal{H}_1 \mid \mathbf{r}_E)$. To compute $\mathcal{P}(\alpha \mid \mathbf{r}_E, \mathcal{H}_1)$, Bayes' theorem can be applied by treating α as a random variable, conditioned on the hypothesis $\mathcal{H}_1$.

Eve can then estimate α using either the posterior mean or the maximum a posteriori (MAP) estimate. As the number of observations L increases, the posterior distribution becomes increasingly concentrated around the true value of α, leading to asymptotically consistent estimation.

IV. Total Detection Error Probability

In this section, we derive the total detection error probability, which is the sum of the probability of false alarm and miss detection. The total detection error probability serves as a metric to assess the effectiveness of the adversary's detection capability [25], [26].

According to Shannon's channel coding theorem, the receiver becomes unable to decode the message $\mathbf{m}$ if its transmission rate is higher than the capacity. Otherwise, the receiver can decode the message. Therefore, the total detection error probability, averaged over the events that $\mathbf{m}$ is decoded and not decoded by Eve, is given by

$$\begin{aligned} P_F + P_M = {} & Pr(\Lambda > \lambda, I(\mathbf{m}; \mathbf{r}_E) < R_m \mid \mathcal{H}_0) \\ & + Pr(\Lambda' > \lambda', I(\mathbf{m}; \mathbf{r}_E) \geq R_m \mid \mathcal{H}_0) \\ & + Pr(\Lambda \leq \lambda, I(\mathbf{m}; \mathbf{r}_E) < R_m \mid \mathcal{H}_1) \\ & + Pr(\Lambda' \leq \lambda', I(\mathbf{m}; \mathbf{r}_E) \geq R_m \mid \mathcal{H}_1), \end{aligned} \tag{27}$$

where

$$I(\mathbf{m}; \mathbf{r}_E) = \begin{cases} \log_2(1 + \|\mathbf{g}_w\|^2 \gamma), & \mathcal{H}_0 \\ \log_2\left(1 + \dfrac{\|\mathbf{g}_w\|^2 \alpha\gamma}{\|\mathbf{g}_w\|^2(1-\alpha)\gamma + 1}\right), & \mathcal{H}_1 \end{cases} \tag{28}$$

is the achievable rate of $\mathbf{m}$ for Eve.

Eve's goal is to minimize $P_F + P_M$ by choosing the detection thresholds, λ and λ', properly. By selecting $\lambda' \in (\Lambda_0, \Lambda_1)$, the second and fourth terms in (27) can be nullified. Since the mutual information $I(\mathbf{m}; \mathbf{r}_E)$ under the null hypothesis $\mathcal{H}_0$ is larger than under the alternative hypothesis $\mathcal{H}_1$, we have

$Pr(I(\mathbf{m};\mathbf{r}_E) < R_m \mid \mathcal{H}_0) < Pr(I(\mathbf{m};\mathbf{r}_E) < R_m \mid \mathcal{H}_1)$. Hence, the sum of the first and third terms in (27) can be minimized by setting $\lambda < \Lambda$ ($\Lambda = 1$), which renders the third term zero. Therefore, the minimum total detection error probability, achieved by the optimum choice of λ and λ', is given by

$$\xi = \min_{\lambda,\lambda'} P_F + P_M \tag{29}$$

$$= Pr(I(\mathbf{m};\mathbf{r}_E) < R_m \mid \mathcal{H}_0) \tag{30}$$

$$= Pr\left(\|\mathbf{g}_w\|^2 < \alpha\lambda_{\max}\right) \tag{31}$$

$$= 1 - \int_0^\infty \int_{\alpha x}^\infty f_{\|\mathbf{g}_w\|^2}(y)dy \cdot f_{\lambda_{\max}}(x)dx \tag{32}$$

$$= 1 - \sum_{k=0}^{N_E-1} \sum_{i=1}^{N_B} \sum_{m=N_A-N_B}^{(N_A+N_B)i-2i^2} d_{i,m} \frac{(k+m)!}{k!m!} \cdot \left(\frac{i}{i+\alpha}\right)^{m+1} \left(\frac{\alpha}{i+\alpha}\right)^{k} \tag{33}$$

where

$$f_{\|\mathbf{g}_w\|^2}(y) = \frac{y^{N_E-1}e^{-y}}{(N_E-1)!},\ y > 0 \tag{34}$$

is the probability density function (PDF) of $\|\mathbf{g}_w\|^2$ [27] and

$$f_{\lambda_{\max}}(x) = \sum_{i=1}^{N_B} \sum_{m=N_A-N_B}^{(N_A+N_B)i-2i^2} \frac{i^{m+1}d_{i,m}}{m!} x^m e^{-ix},\ x > 0 \tag{35}$$

is the PDF of $\lambda_{\max}$ for $N_A \geq N_B$ [20]. The coefficients $d_{i,m}$ are listed in [20, Tables I–IV], with the property that the summation of $d_{i,m}$ over all i and m is unity. The proof of (33) is provided in Appendix A.

1) Covertness: For covert signature transmission, we require $\xi \geq 1 - \epsilon$, where ϵ is a small value. When $\xi = 1$, the signature transmission becomes completely undetectable to the adversary, as their ability to detect it is no better than random guessing. In this case, the detection outcome provides no information about the signature transmission. A detailed proof of this is provided in Appendix B.

2) Proposition: The minimum total detection error probability, ξ, approaches one as the number of transmitter antennas, N_A, increases. This behavior is evident from (33) when $N_A \gg N_B$, where the term $\left(\frac{i}{i+\alpha}\right)^{m+1}$ converges to zero as N_A (and consequently m) increases, regardless of α. This indicates that the signature becomes undetectable, regardless of the power allocated to it, when the transmitter has a sufficiently large number of antennas.

When both N_A and N_B increase (e.g., $N_A = N_B$), the channel rank also increases. This higher channel rank results in a larger maximum eigenvalue, leading to an increase in the detection error probability, as described by (31). Consequently, ξ converges more rapidly to 1 when both N_A and N_B grow. This phenomenon arises because the maximum eigenvalue, λ_{max}, increases with larger N_A and N_B (its mean is $(\sqrt{N_A} + \sqrt{N_B})^2$ [28]), while Eve's channel gain, $\|\mathbf{g}_w\|$, remains unchanged. Fig. 2 illustrates and supports this proposition.

3) Remark: This feature represents a significant improvement over previous physical layer authentication methods, which require minimal transmission power for the signature to protect keys [5], [7]. Our proposed approach eliminates the need to reduce the signature's transmission power for concealment purposes. This distinctive characteristic allows for increasing the signature's transmission power to reduce Bob's decoding outage (error) probability of the signature without risking detection by Eve. It is also evident from (31) that ξ remains independent of the total transmission power, γ.

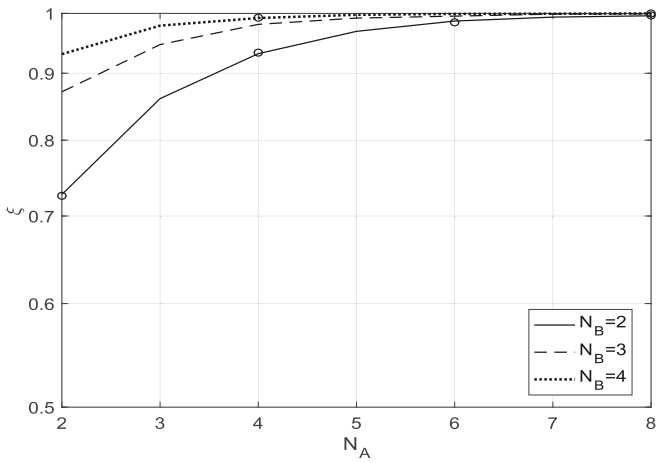

Fig. 2. Total detection error probability, ξ, versus number of transmitter's antennas, N_A; $N_E = 2$, $\alpha = 0.9$.

A. Numerical Results

Fig. 2 illustrates the relationship between the total detection error probability, ξ, and the number of transmitter antennas, N_A, for various values of N_B. The plot demonstrates that as N_A increases, the total detection error probability also increases, eventually converging to 1, which signifies perfect covertness of the signature transmission. Moreover, the convergence of ξ to 1 occurs more rapidly for larger N_B, indicating that increasing the intended receiver's antenna accelerates the convergence. This behavior is attributed to the increase in the maximum eigenvalue, λ_{max}, for larger N_B. This trend highlights the significant impact of the intended receiver's antenna size on Eve's detection performance. Additionally, it is important to note that there is no requirement to reduce transmission power to achieve covertness of signature transmission. This means the transmission power does not affect the concealment of the signature's existence. The analytical result in (33), represented by circles, aligns closely with the simulation results, depicted by lines.

Fig. 3 depicts the relationship between the total detection error probability, ξ, and the fractional power allocated to the message, α, for different values of N_A. The plot reveals that as N_A increases, ξ becomes increasingly robust to changes in α. For sufficiently large N_A, the power allocation to the signature has no impact on the total detection error probability, thereby validating the proposition.

Fig. 4 illustrates the relationship between the total detection error probability, ξ, and the number of eavesdropper antennas, N_E, for different numbers of receiver antennas, N_B. The graph shows that ξ decreases as N_E increases, indicating enhanced eavesdropping capability. However, this effect becomes less pronounced with larger N_B values, suggesting

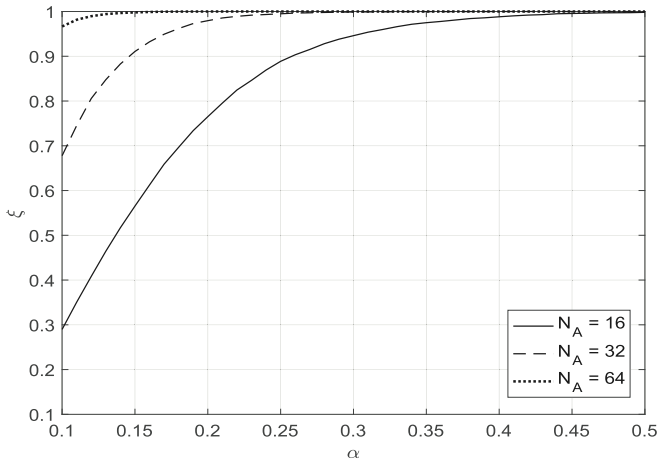

Fig. 3. Total detection error probability, ξ, versus α for different values of N_A; $N_B = N_E = 4$.

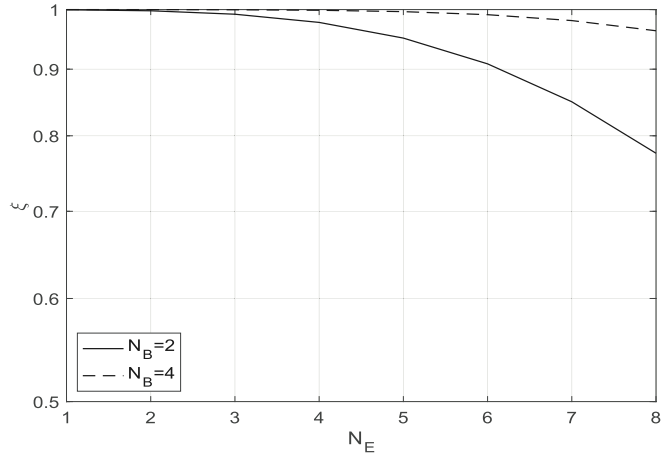

Fig. 4. Total detection error probability, ξ, versus number of eavesdropper's antennas, N_E; $N_A = 8$, $\alpha = 0.99$.

that increasing the legitimate receiver's antennas can mitigate the eavesdropper's advantage. Additionally, the introduction of artificial noise can make it more challenging for an adversary to detect the authentication process. A more detailed analysis of this effect will be presented in Section VII.

V. Decoding Outage Probability of Signature

The decoding outage (error) probability of the signature at Bob is given by:

$$P_{o,t} = Pr(I(\mathbf{t};\mathbf{r}_B) < R_t) \tag{36}$$

$$= Pr\left(\lambda_{\max} < \frac{2^{R_t}-1}{(1-\alpha 2^{R_t})\gamma}\right). \tag{37}$$

Applying (35) to (37), we obtain

$$P_{o,t} = \sum_{i=1}^{N_B} \sum_{m=N_A-N_B}^{(N_A+N_B)i-2i^2} \frac{d_{i,m}}{m!}\Gamma\left(m+1, \frac{i(2^{R_t}-1)}{(1-\alpha 2^{R_t})\gamma}\right) \tag{38}$$

$$= \sum_{i=1}^{N_B} \sum_{m=N_A-N_B}^{(N_A+N_B)i-2i^2} d_{i,m}$$

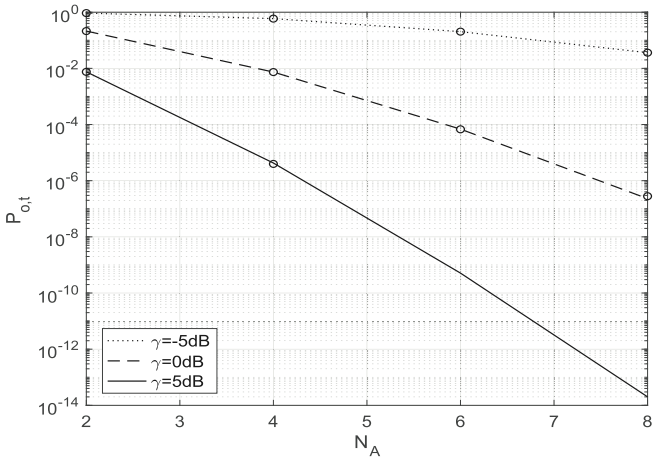

Fig. 5. Decoding outage probability of signature at Bob, $P_{o,t}$, versus number of transmitter's antennas, N_A; $N_B = 2$, $R_t = 0.1$, and $\alpha = 0.9$.

$$\cdot \left[1 - e^{-\frac{i(2^{R_t}-1)}{(1-\alpha 2^{R_t})\gamma}} \sum_{k=0}^{m} \frac{1}{k!}\left(\frac{i(2^{R_t}-1)}{(1-\alpha 2^{R_t})\gamma}\right)^k\right] \tag{39}$$

for $R_t < \log_2(1/\alpha)$. For $\alpha \geq 2^{-R_t}$, or equivalently, $R_t \geq \log_2(1/\alpha)$, $P_{o,t} = 1$.

1) Asymptotic Analysis: Let

$$\mu = (\sqrt{N_A} + \sqrt{N_B})^2 \tag{40}$$

$$\sigma = (\sqrt{N_A} + \sqrt{N_B})\left(\frac{1}{\sqrt{N_A}} + \frac{1}{\sqrt{N_B}}\right)^{\frac{1}{3}}. \tag{41}$$

For $N_A, N_B \to \infty$, the probability density function of λ_{max} converges to

$$\frac{\lambda_{max} - \mu}{\sigma} \to \mathcal{X}_2 \tag{42}$$

where $\mathcal{X}_2$ is Tracy-Widom distribution of order 2 [28]. Therefore, it follows from (37) that the decoding outage probability of signature is given by

$$P_{o,t} \to F_{\mathcal{X}_2}\left(\frac{\zeta-\mu}{\sigma}\right), \tag{43}$$

for large N_A, N_B, where $F_{\mathcal{X}_2}(\cdot)$ is the cumulative distribution function (CDF) of $\mathcal{X}_2$ and $\zeta = \frac{2^{R_t}-1}{(1-\alpha 2^{R_t})\gamma}$.

2) Proposition: The decoding outage probability of the signature ($P_{o,t}$) approaches zero as the number of transmitter antennas (N_A) increases if $R_t < \log_2(1/\alpha)$, or equivalently, $\alpha < 2^{-R_t}$. This trend is evident from (39) when $N_A \gg N_B$, where the summation inside the bracket converges to $e^{\frac{i(2^{R_t}-1)}{(1-\alpha 2^{R_t})\gamma}}$ as N_A (and consequently m) increases. This convergence causes the term inside the bracket to approach zero. When both N_A and N_B increase (e.g., $N_A = N_B$), the channel rank also increases, leading to a larger maximum eigenvalue. This results in a lower decoding outage probability, as indicated by (37). The numerical result shown in Fig. 5 supports the proposition.

This proposition has significant implications for our authentication system. It suggests that with a sufficiently large number of transmitter antennas, we can achieve two crucial objectives simultaneously. First, the intended receiver, Bob, can decode the signature without any errors. Second, the

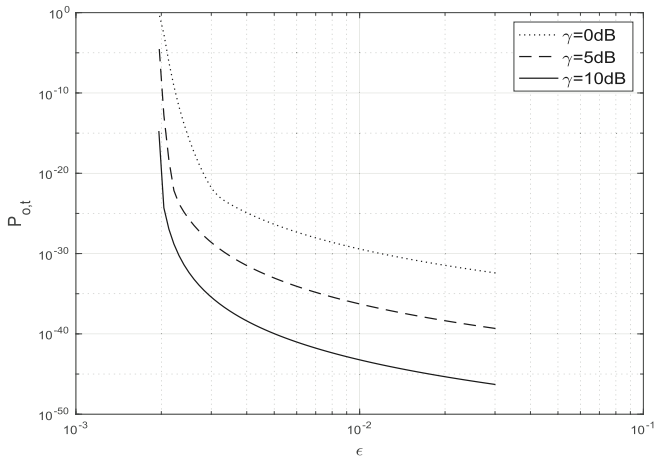

Fig. 6. Decoding outage probability of signature at Bob, $P_{o,t}$, versus total detection error probability constraint, ϵ; $N_A = 8$, $N_B = N_E = 4$, $R_t = 0.1$.

eavesdropper, Eve, cannot detect the presence of the signature at all, as discussed in Section IV. In essence, this demonstrates the potential of our scheme to provide reliable authentication without being detected by leveraging a large number of transmitter antennas. This finding highlights the effectiveness of our approach in enhancing the authentication security and reliability.

3) Remark: Equation (31) shows that Eve's total detection error probability, ξ, increases with the power allocation factor α. However, as Equation (37) indicates, a higher α also increases Bob's signature decoding outage probability, $P_{o,t}$. To ensure that $P_{o,t}$ vanishes as the number of transmit antennas grows, it is necessary to keep $\alpha < 2^{-R_t}$. Therefore, ξ can be maximized while maintaining reliable signature decoding by selecting α close to 2^{-R_t}.

Since the length of the hash-based signature, $h(\mathbf{d}, K)$, is typically much shorter than that of the information data $\mathbf{d}$, the signature rate R_t can be very small. For example, with a 256-bit hash (before encoding) and a 19200-bit message (after encoding), as specified in 5G NR, we obtain $R_t = 0.0133$, which permits choosing α as high as 0.99.

A. Numerical Results

Fig. 5 shows the decoding outage probability of the signature at Bob, $P_{o,t}$, plotted against the number of transmitter antennas, N_A, for different values of SNR, γ. The plot clearly demonstrates that as N_A increases, $P_{o,t}$ decreases, almost exponentially, converging to zero. This behavior confirms the validity of Proposition in Section V-2. Furthermore, it is noteworthy that the rate of decrease becomes steeper for larger values of γ, indicating that higher SNR facilitates more reliable signature decoding by Bob. However, the increase of transmission power does not affect the total detection error probability of Eve, ξ, as described by (31). The simulation result, represented by circles, aligns closely with the analytical result in (39), depicted by lines.

Fig. 6 illustrates the relationship between the decoding outage probability of the signature at Bob, $P_{o,t}$, and the covertness constraint, ϵ, for scenarios where $\xi \geq 1 - \epsilon$. A lower value of ϵ implies a stricter requirement for signature covertness. The plot is generated by varying the power allocation factor, α, across the range $[0.5, 2^{-R_t}]$. The plot reveals that $P_{o,t}$ increases as the covertness constraint becomes more stringent (i.e., as ϵ decreases). This increase is attributed to the reduced power allocated to the signature (larger α) when enforcing a stricter covertness requirement. Interestingly, it is observed that increasing the transmission power (γ) can actually reduce $P_{o,t}$ without compromising the signature's covertness (maintaining a fixed ϵ). This result stands in contrast to earlier findings, such as those in [13], which reported that higher transmission power leads to reduced key equivocation - i.e., increased information leakage about the key - thereby weakening security. In contrast, our result demonstrates that transmission power can be strategically increased to improve the reliability of signature decoding by the intended receiver while preserving the signature's covertness, thereby supporting both reliability and security in message authentication.

VI. Authenticated Message Throughput

The authenticated message throughput signifies the average number of correctly received and validated message bits per channel use. This metric combines both communication reliability and trustworthiness, which traditionally have been studied separately, and provides a more comprehensive evaluation of system performance.

The message, transmitted at the rate R_m as expressed in (10), can be successfully decoded (after canceling the signature) and authenticated if the signature is decoded. The latter event occurs when the achievable rate of the signature $I(\mathbf{t}; \mathbf{r}_B)$ exceeds the transmission rate R_t, or equivalently, $\lambda_{\max}\gamma \geq \frac{2^{R_t}-1}{1-\alpha 2^{R_t}}$ for $R_t < \log_2(1/\alpha)$. If $\lambda_{\max}\gamma < \frac{2^{R_t}-1}{1-\alpha 2^{R_t}}$, the signature cannot be decoded and consequently the message transmitted at the rate in (10) cannot be decoded and authenticated. Therefore, by defining $X := \lambda_{\max}\gamma$, the authenticated message throughput (bits per channel use) can be expressed as:

$$W = \int_{\frac{2^{R_t}-1}{(1-\alpha 2^{R_t})\gamma}}^{\infty} \log_2(1 + \alpha\gamma x)\, f_{\lambda_{max}}(x)dx \tag{44}$$

for $R_t < \log_2(1/\alpha)$. For $R_t \geq \log_2(1/\alpha)$, the authenticated message throughput is zero.

Remark:

1) Since (44) increases as a function of α for $\alpha < 2^{-R_t}$ and abruptly drops to zero when $\alpha \geq 2^{-R_t}$, the optimal value of α that maximizes the authenticated throughput is $\alpha_{\text{opt}} = 2^{-R_t}$.
2) The authenticated message throughput with covertness requirement can be computed from W subject to $\xi \geq 1 - \epsilon$.
3) The message throughput without authentication requirement ($\alpha = 1$), denoted as $W_{\text{no auth}}$, can be obtained from (10) as

$$W_{\text{no auth}} = \int_0^{\infty} \log_2(1 + \gamma x)\, f_{\lambda_{max}}(x)dx. \tag{45}$$

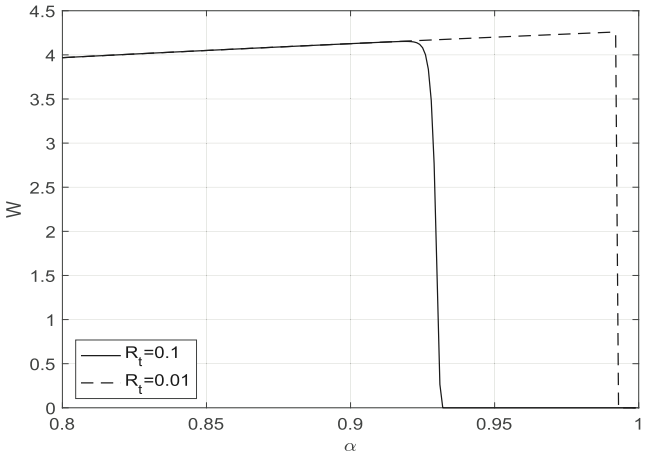

Fig. 7. Authenticated message throughput, W, versus power allocation to the message, α; $N_A = 4$, $N_B = N_E = 2$, $\gamma = 5$ dB.

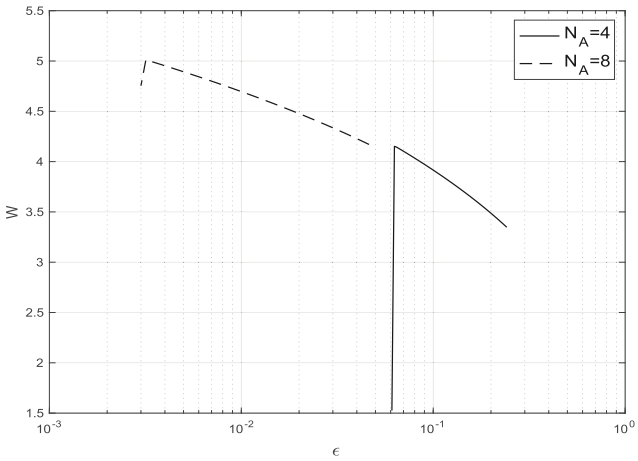

Fig. 8. Authenticated message throughput, W, versus total detection error probability constraint, ϵ; $N_B = N_E = 2$, $R_t = 0.1$, $\gamma = 5$ dB.

Comparing (44) and (45), we can quantify the throughput loss due to message authentication. For $R_t < \log_2(1/\alpha)$, this loss is expressed as

$$W_{\text{loss}} = W_{\text{no auth}} - W_{\text{auth}} \tag{46}$$

$$= \int_0^{\frac{2^{R_t}-1}{(1-\alpha 2^{R_t})\gamma}} \log_2(1+\gamma x)\, f_{\lambda_{max}}(x)dx + \int_{\frac{2^{R_t}-1}{(1-\alpha 2^{R_t})\gamma}}^{\infty} \log_2\left(\frac{1+\gamma x}{1+\alpha\gamma x}\right) f_{\lambda_{max}}(x)dx. \tag{47}$$

A. Numerical Results

Fig. 7 illustrates the relationship between authenticated message throughput (W) and the power allocation factor for the message (α) across different signature rates (R_t). The graph reveals a crucial insight: for each R_t value, there exists an optimal α that maximizes W, typically close to 2^{-R_t}. As we increase the power allocated to the message (α), we observe a corresponding increase in message throughput. However, this comes at a cost: the probability of signature decoding failure rises due to the diminished power available for the signature. To ensure successful message authentication, the signature must be correctly decoded, which requires maintaining α below 2^{-R_t}. To achieve the highest authenticated message throughput, it is ideal to set α near, but not exceeding, 2^{-R_t}. This finding has exciting practical implications, particularly in scenarios where R_t is small - a common occurrence due to the typically short signature length relative to the information data length. In such cases, the throughput loss incurred to provide message authentication can be minimal because α can be close to 1.

Fig. 8 shows the authenticated message throughput, W, plotted against the covertness constraint, ϵ, for scenarios where the total detection error probability ξ is greater than or equal to $1-\epsilon$. The plot is generated by varying the power allocation factor to the message, α, across the range $[0.5, 2^{-R_t}]$. The analysis considers different numbers of transmitter antennas, specifically $N_A = 4$ and 8, while keeping the number of receiver and eavesdropper antennas fixed at $N_B = N_E = 2$.

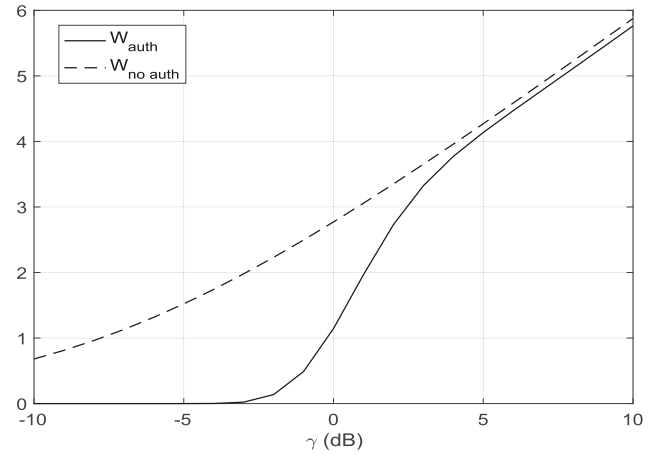

Fig. 9. Authenticated message throughput, W_{auth}, and message throughput without authentication, $W_{\text{no auth}}$, versus SNR, γ; $N_A = 4$, $N_B = N_E = 2$, $R_t = 0.1$, $\alpha = 2^{-R_t} - 0.01$.

The plot reveals an interesting interplay between authenticated throughput and covertness: W initially increases as the covertness constraint becomes more stringent (i.e., as ϵ decreases). This occurs because allocating more power to the message (increasing α) improves authenticated throughput while simultaneously reducing the probability of signature detection. However, this trend only holds until a certain ϵ threshold is reached. Beyond this threshold, a decline in W is observed. This drop can be attributed to the signature decoding failure when α approaches 2^{-R_t}, where the signature's decoding outage probability reaches 1. Consequently, when the signature cannot be decoded, message authentication is not possible, resulting in zero authenticated message throughput. Importantly, increasing the number of transmitter antennas can mitigate this decline by increasing the probability of decoding the signature. This ultimately leads to a higher authenticated throughput while still maintaining the desired level of covertness.

Fig. 9 illustrates the relationship between authenticated message throughput and SNR under two conditions: with

authentication (W_{auth}) and without authentication ($W_{\text{no auth}}$). A notable observation is the substantial difference between $W_{\text{no auth}}$ and W_{auth} at lower SNR values, indicating a significant loss of throughput due to the authentication of messages in this range. This phenomenon can be attributed to the challenges in decoding the signature when signal quality is poor, resulting in failed message authentication attempts. However, as the SNR increases, the gap between authenticated and non-authenticated throughput narrows considerably. This convergence suggests that the impact of authentication on throughput becomes minimal under favorable signal conditions.

VII. Artificial Noise

In this section, we analyze the impact of artificial noise (AN) on the total detection error probability of the signature and the authenticated message throughput. AN is intentionally generated interference added to the transmitted signal to degrade the channel of potential eavesdroppers while minimizing impact on the legitimate receiver. It helps mask the presence of the actual transmission, making it harder for an adversary to detect that communication is occurring, especially when the eavesdropper has an advantage in terms of the number of antennas.

A. Transmitted Signal

The transmitted signal is given by

$$\mathbf{x}_i = \mathbf{w}s_i + \mathbf{H}^{\perp}\mathbf{z} \tag{48}$$

where $\mathbf{H}^{\perp}$ is an $N_A \times (N_A - 1)$ matrix composed of $N_A - 1$ orthonormal column vectors of length N_A which are in the null space of $\mathbf{H}$, i.e. $\mathbf{H}^T\mathbf{H}^{\perp} = \mathbf{0}$, and $\mathbf{z} \sim \mathcal{CN}(0, \sigma_z^2\mathbf{I}_{N_a-1})$ is an $(N_A - 1) \times 1$ AN vector. AN is statistically identical to the background noise so that Eve cannot distinguish it from the received noise.

The average transmission power is given by

$$E[|x_i|^2] = \sigma_s^2 + (N_A - 1)\sigma_z^2 = P, \tag{49}$$

where $\sigma_s^2 = E[|s_i|^2]$ denotes the average signal power. One important design parameter is the ratio of power allocated to the information bearing signal and the artificial noise. We denote the fraction of total power allocated to the information signal as ϕ. Hence, we have the following relationships:

$$\sigma_s^2 = \phi P \tag{50}$$

$$\sigma_z^2 = (1 - \phi)P/(N_A - 1). \tag{51}$$

In the rest of this paper, we investigate the impact of ϕ on the tradeoff between the total detection error probability and the authenticated message throughput.

B. Total Detection Error Probability

The received signal at Bob and Eve are given by

$$\mathbf{y}_{B,i} = \mathbf{H}\mathbf{w}s_i + \mathbf{n}_{B,i} \tag{52}$$

$$\mathbf{y}_{E,i} = \mathbf{G}\mathbf{w}s_i + \mathbf{G}\mathbf{H}^{\perp}\mathbf{z} + \mathbf{n}_{E,i}, \tag{53}$$

respectively. After applying MRC to $\mathbf{y}_{B,i}$ and $\mathbf{y}_{E,i}$, we obtain $r_{B,i}$ in (7) and

$$r_{E,i} = \|\mathbf{g}_w\|s_i + \mathbf{g}_w^H\mathbf{G}\mathbf{H}^{\perp}\mathbf{z}/\|\mathbf{g}_w\| + \mathbf{g}_w^H\mathbf{n}_{E,i}/\|\mathbf{g}_w\|, \tag{54}$$

where $\mathbf{g}_w^H\mathbf{G}\mathbf{H}^{\perp}\mathbf{z}/\|\mathbf{g}_w\| \sim \mathcal{CN}(0, N_A\sigma_g^2\sigma_z^2)$ and $\mathbf{g}_w^H\mathbf{n}_{E,i}/\|\mathbf{g}_w\| \sim \mathcal{CN}(0, \sigma_n^2)$.

It follows from (54) that the achievable rate of the message $\mathbf{m}$ for Eve under $\mathcal{H}_0$ (i.e., $s_i = m_i$) is given by

$$I(\mathbf{m}; \mathbf{r}_E) = \log_2\left(1 + \frac{\|\mathbf{g}_w\|^2\sigma_s^2}{N_A\sigma_g^2\sigma_z^2 + \sigma_n^2}\right) \tag{55}$$

$$= \log_2\left(1 + \frac{\|\mathbf{g}_w\|^2\phi\gamma}{N_A\sigma_g^2(1 - \phi)\gamma/(N_A - 1) + 1}\right). \tag{56}$$

Assuming the message is transmitted at the maximum rate:

$$R_m = \log_2(1 + \alpha\lambda_{\max}\phi\gamma), \tag{57}$$

it follows from (56) and (57) that the minimum total detection error probability is given by

$$\xi = Pr(I(\mathbf{m}; \mathbf{r}_E) < R_m \mid \mathcal{H}_0) \tag{58}$$

$$= Pr(\|\mathbf{g}_w\|^2 < \alpha'\lambda_{\max}), \tag{59}$$

where

$$\alpha' = \alpha(N_A\sigma_g^2(1 - \phi)\gamma/(N_A - 1) + 1). \tag{60}$$

Hence, the minimum total detection error probability is given by (33) with α replaced by α'.

1) Remark: It should be noted that α' increases with increasing γ, hence ξ also increases. This implies that a higher transmission power makes it more difficult for potential adversaries to detect the presence of the signature. This is due to the increased interference caused by AN towards potential adversaries when the transmit power increases, while the intended receiver remains unaffected by this interference. Consequently, the higher transmission power benefits covert transmission of the signature by hindering detection by adversaries. This is in stark contrast to conventional covert transmission techniques, where the total detection error probability typically decreases with increasing transmission power.

C. Decoding Outage Probability of Signature

Since $E[|s_i|^2] = \phi P$, it follows from (9) that the achievable rate of the signature for Bob is given by

$$I(\mathbf{t}; \mathbf{r}_B) = \log_2\left(1 + \frac{(1 - \alpha)\lambda_{\max}\phi\gamma}{\alpha\lambda_{\max}\phi\gamma + 1}\right). \tag{61}$$

For a fixed transmission rate of R_t for the signature, the decoding outage probability of the signature at Bob is given by

$$P_{o,t} = Pr(I(\mathbf{t}; \mathbf{r}_B) < R_t) \tag{62}$$

$$= Pr\left(\lambda_{\max} \le \frac{2^{R_t} - 1}{(1 - \alpha 2^{R_t})\phi\gamma}\right). \tag{63}$$

Hence, the decoding outage probability of the signature is given by (39) with γ replaced by $\phi\gamma$.

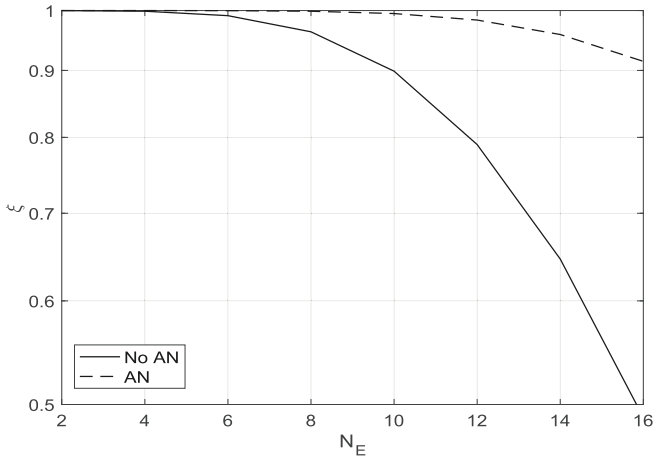

Fig. 10. Total detection error probability, ξ, versus number of eavesdropper's antennas, N_E; $N_A = 8$, $N_B = 4$, $\alpha = 0.99$, $\phi = 0.95$, $\gamma = 10$ dB.

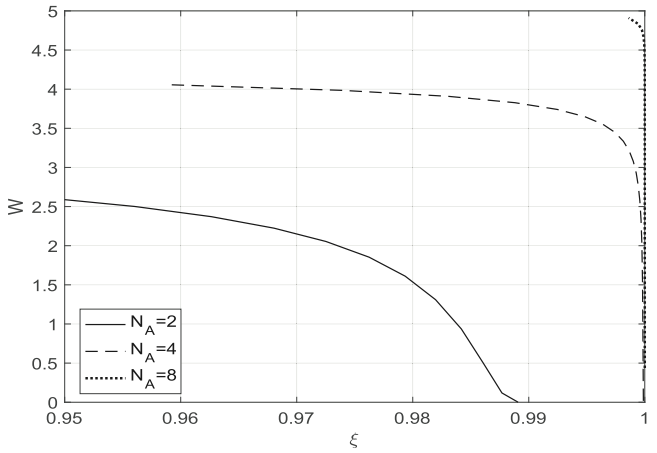

Fig. 11. Authenticated message throughput, W, versus total detection error probability, ξ, for different values of N_A; $N_B = N_E = 2$, $\alpha = 0.9$, $R_t = 0.1$, $\gamma = 5$ dB.

D. Authenticated Message Throughput

The signature can be decoded successfully, allowing Bob to decode and authenticate the message, if $I(\mathbf{t};\mathbf{r}_B) \geq R_t$, or equivalently $X := \phi\lambda_{\max}\gamma \geq \frac{2^{R_t}-1}{(1-\alpha 2^{R_t})}$. Hence, for the message transmitted at the rate of (57), the authenticated message throughput (bits per channel use) is given by:

$$W = \int_{\frac{2^{R_t}-1}{1-\alpha 2^{R_t}}}^{\infty} \log_2(1+\alpha x)\, f_X(x)dx \tag{64}$$

for $R_t < \log_2(1/\alpha)$, where

$$f_X(x) = \frac{1}{\phi\gamma} f_{\lambda_{\max}}\left(\frac{x}{\phi\gamma}\right). \tag{65}$$

E. Numerical Results

Fig. 10 illustrates the relationship between the total detection error probability, ξ, and the number of eavesdropper antennas, N_E, comparing scenarios with and without artificial noise (AN). The graph clearly demonstrates that the transmission of AN results in a higher ξ, thereby making it more challenging for an adversary to detect the signature. This effect becomes especially significant as the number of antennas at the eavesdropper's disposal increases. Overall, the figure highlights the effectiveness of AN in obscuring the authentication process from well-equipped adversaries.

Fig. 11 illustrates the relationship between authenticated message throughput (W) and total detection error probability (ξ) with AN for various numbers of transmitter antennas (N_A), as the power allocation factor for the signal (ϕ) varies. The plot reveals an inverse relationship: as more power is allocated to AN, ξ increases while W decreases. This demonstrates how adjusting ϕ allows for a balance between the covertness of the signature transmission and the authenticated message throughput, highlighting the trade-off between these two factors.

VIII. Signature Secrecy

Unlike covert transmission, which aims to conceal the presence of the signature, physical layer security (PLS) using wiretap codes is focused on ensuring the confidentiality of the signature itself. PLS prevents eavesdroppers from gaining any meaningful information about the signature, rather than preventing them from detecting its transmission. In this section, we will compare the concepts of secrecy (keeping the content confidential) and covertness (hiding the existence) of the signature.

A. Secrecy Outage Probability

We consider the wiretap code [29] to secure the transmission of signatures. There are two rate parameters for the signature, namely, the codeword transmission rate, R_t, and the confidential information rate, R_s, for the signature. The rate difference $R_t - R_s$ reflects the cost of securing the signature transmission against eavesdropping. The secrecy capacity of MIMO broadcast channel is computed in [30].

When instantaneous channel state information (CSI) of the eavesdropper is unknown at the transmitter (Alice), perfect secrecy is not achievable. The secrecy outage probability is adopted to measure the secrecy performance of the signature transmission. The secrecy outage probability of the signature is given by [31] and [32]

$$P_{so} = Pr(R_t - R_s < I(\mathbf{t};\mathbf{r}_E)), \tag{66}$$

where

$$I(\mathbf{t};\mathbf{r}_E) = \log_2\left(1 + \frac{(1-\alpha)\|\mathbf{g}_w\|^2\gamma}{\alpha\|\mathbf{g}_w\|^2\gamma + 1}\right) \tag{67}$$

is the achievable rate of the signature for Eve. P_{so} measures the probability that a transmitted signature fails to achieve perfect secrecy. It can be shown from (66) and (67) that

$$P_{so} = Pr\left(\|\mathbf{g}_w\|^2 > \frac{2^{R_t-R_s}-1}{(1-\alpha 2^{R_t-R_s})\gamma}\right) \tag{68}$$

$$= 1 - \frac{\Gamma\left(N_E, \frac{2^{R_t-R_s}-1}{(1-\alpha 2^{R_t-R_s})\gamma}\right)}{(N_E-1)!}, \tag{69}$$

for $\alpha < 2^{-(R_t-R_s)}$ and $P_{so} = 1$ for $\alpha \geq 2^{-(R_t-R_s)}$.

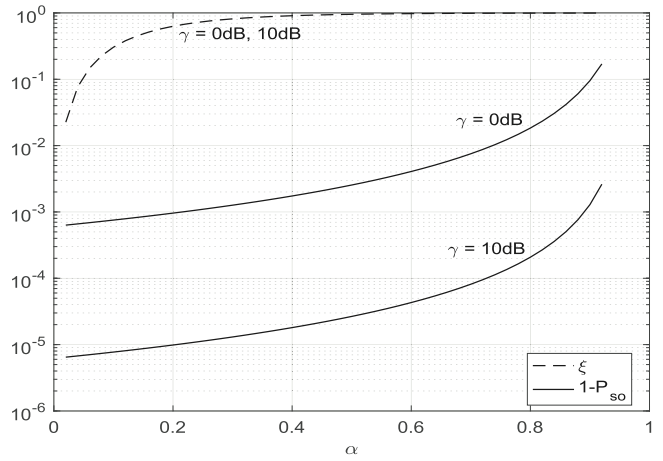

Fig. 12. Secrecy probability, $1 - P_{so}$, and covertness probability, ξ, versus α; $R_t = 0.1$, $R_s = 0.05$, $N_A = 8$, $N_B = 2$, $N_E = 2$.

B. Comparing Secrecy and Covertness in Signature Transmission

We compare two probabilities that determine how effective secrecy and covertness are when transmitting signatures. The first probability, secrecy probability, tells us how likely it is that Eve cannot extract any meaningful information about signatures. This probability is $1 - P_{so}$. Perfect secrecy is achieved if $P_{so} = 0$.

The second probability, which we call the covertness probability, tells us how likely it is that Eve cannot detect the transmission of signatures. This probability is ξ. The perfect covertness is achieved if $\xi = 1$.

By comparing these two probabilities - the secrecy probability ($1 - P_{so}$) and the covertness probability (ξ) - we can get a quantitative sense of how the secrecy approach (which aims to keep the signature's content secret) compares to the covertness approach (which aims to hide the signature's very existence). This comparison helps us understand the trade-offs and effectiveness of these two different security strategies.

C. Numerical Result

Figure 12 illustrates the relationship between secrecy probability ($1 - P_{so}$) and covertness probability (ξ) as functions of $\alpha \in (0, 2^{-R_t})$. For values of $\alpha > 2^{-R_t}$, the signature becomes undecodable. The graph shows that ξ consistently exceeds $1 - P_{so}$, indicating that the probability that the signature goes undetected is higher than the probability that Eve will extract information about it. This suggests that Eve is more likely to intercept information about the signature than to detect its transmission. This comparison assumes $R_s < R_t$, which means that the rate of signature bits transmitted covertly is greater than that transmitted securely. Furthermore, while the secrecy probability decreases as the SNR, γ, increases, the covertness probability remains unaffected by γ.

IX. Eigenbeamforming

In this section, we consider eigenbeamforming in which Alice transmits multiple data streams $\mathbf{d}_1, \mathbf{d}_2, \cdots \mathbf{d}_r$ in parallel, achieving the maximum rate of $\sum_{i=1}^{r} \log_2(1 + \lambda_i \gamma_i)$, where $r = \min\{N_A, N_B\}$ is the channel rank, λ_i is the i-th largest eigenvalue of the matrix $\mathbf{H}\mathbf{H}^H$, and $\gamma_i = P_i/\sigma_n^2$ represents the SNR for $\mathbf{d}_i$.

In this setting, the hash value can be generated based on $\mathbf{d}_1, \cdots, \mathbf{d}_r$ as $h(\mathbf{d}_1, \cdots, \mathbf{d}_r, K)$, and embedded onto the signal $\mathbf{m}_1 = e_m(\mathbf{d}_1)$, which is then transmitted over the channel corresponding to the largest eigenvalue $\lambda_1 (= \lambda_{max})$.

The eigenbeamforming requires multiple RF chains, while MIMO-MRC technique only requires a single RF chain, reducing hardware complexity and power consumption. The latter also minimizes interference in multi-user scenarios by directing the transmission toward the intended receiver.

A. Transmitted Signal

The transmitted signal is represented by

$$\mathbf{x} = \mathbf{V}[\mathbf{s}_1 \cdots \mathbf{s}_r]^T, \tag{70}$$

where $\mathbf{V} = [\mathbf{v}_1 \cdots \mathbf{v}_r]$ is a $N_A \times r$ unitary matix with $\|\mathbf{v}_i\|^2 = 1$, $i = 1, \cdots, r$, generated from the singular value decomposition of $\mathbf{H}$, i.e., $\mathbf{H} = \mathbf{U}\mathbf{D}\mathbf{V}^H$, and $\mathbf{s}_i$ is the signal transmitted over the channel associated with the i-th eigenvalue. We assume

$$\mathbf{s}_1 = \begin{cases} \mathbf{m}_1, & \mathcal{H}_0 \\ \sqrt{\alpha}\mathbf{m}_1 + \sqrt{1-\alpha}\mathbf{t}, & \mathcal{H}_1 \end{cases} \tag{71}$$

where $\mathbf{t} = e_t(h(\mathbf{d}_1, \cdots, \mathbf{d}_r, K))$ is the signature for $\mathbf{d}_1, \cdots, \mathbf{d}_r$ and

$$\mathbf{s}_i = \mathbf{m}_i, \ i = 2, \cdots, r. \tag{72}$$

B. Received Signal

The received signal at Bob is:

$$\mathbf{y}_{B,i} = \sqrt{\lambda_i}\mathbf{s}_i + \mathbf{z}_{B,i}, \ i = 1, \cdots, r, \tag{73}$$

where $\mathbf{z}_{B,i}$ is noise. Assuming equal power allocation among the data streams, i.e., $\|\mathbf{s}_i\|^2/L = P/r$, the transmission rate for $\mathbf{m}_1$ is:

$$R_{m,1} = \log_2(1 + \alpha\lambda_1\gamma/r). \tag{74}$$

The received signal at Eve is given by:

$$\mathbf{y}_E = \mathbf{G}\mathbf{x} + \mathbf{n}_E \tag{75}$$

$$= \mathbf{G}\mathbf{v}_1\mathbf{s}_1 + \sum_{i=2}^{r} \mathbf{G}\mathbf{v}_i\mathbf{s}_i + \mathbf{n}_E. \tag{76}$$

C. MMSE Filtering

The MMSE filter for Eve can be derived by minimizing the mean square error between the transmitted signal vector $[\mathbf{s}_1, \cdots, \mathbf{s}_r]^T$ and the estimate at Eve. The MMSE filter is given by:

$$W_{\text{MMSE}} = \left((\mathbf{G}\mathbf{V})^H\mathbf{G}\mathbf{V} + \frac{r}{\gamma}I\right)^{-1}(\mathbf{G}\mathbf{V})^H.$$

To detect $\mathbf{s}_1$, Eve applies the MMSE filter corresponding to $\mathbf{s}_1$, which is represented by the first row of W_{MMSE}, denoted as $\mathbf{w}_1$. The output of the MMSE filter is:

$$\mathbf{r}_{E,1} = \mathbf{w}_1\mathbf{y}_E \tag{77}$$

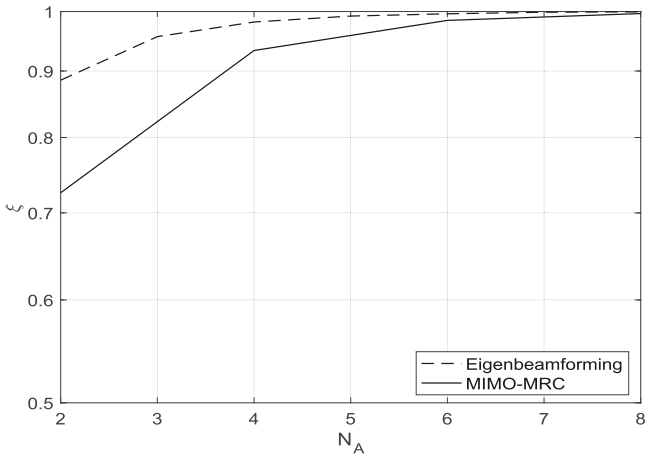

Fig. 13. Total detection error probability, ξ, versus number of transmitter's antennas, N_A, for $\alpha = 0.9$; $N_B = N_E = 2$, $\gamma = 10$ dB.

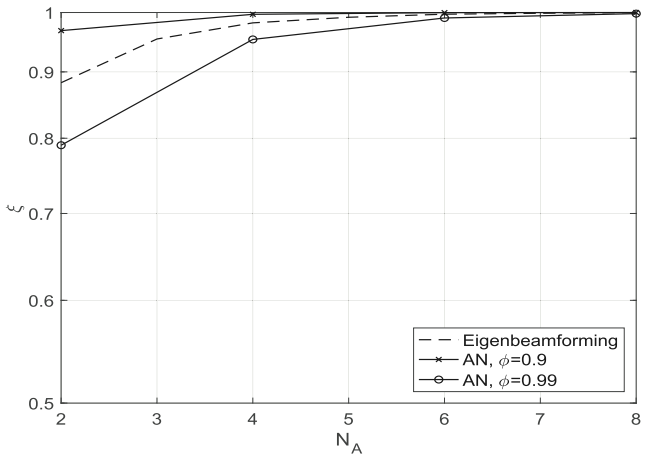

Fig. 14. Total detection error probability, ξ, versus number of transmitter's antennas, N_A, for $\alpha = 0.9$; $N_B = N_E = 2$, $\gamma = 10$ dB.

$$= \mathbf{w}_1\mathbf{G}\mathbf{v}_1\mathbf{s}_1 + \sum_{i=2}^{r} \mathbf{w}_1\mathbf{G}\mathbf{v}_i\mathbf{s}_i + \mathbf{w}_1\mathbf{n}_E. \tag{78}$$

The SINR for $\mathbf{s}_1$ at Eve is given by:

$$\Gamma_1 = \frac{|\mathbf{w}_1\mathbf{G}\mathbf{v}_1|^2 \frac{\gamma}{r}}{\sum_{i=2}^{r} |\mathbf{w}_1\mathbf{G}\mathbf{v}_i|^2 \frac{\gamma}{r} + \|\mathbf{w}_1\|^2}. \tag{79}$$

Thus, the achievable rate of $\mathbf{m}_1$ for Eve, under $\mathcal{H}_0$, is:

$$I(\mathbf{m}_1; \mathbf{r}_E) = \log_2(1 + \Gamma_1).$$

D. Total Detection Error Probability

The total detection error probability is:

$$\xi = \Pr\left(I(\mathbf{m}_1; \mathbf{r}_E) < R_{m,1} \mid \mathcal{H}_0\right) \tag{80}$$

$$= \Pr\left(\frac{|\mathbf{w}_1\mathbf{G}\mathbf{v}_1|^2}{\sum_{i=2}^{r} |\mathbf{w}_1\mathbf{G}\mathbf{v}_i|^2 \frac{\gamma}{r} + \|\mathbf{w}_1\|^2} < \alpha\lambda_1\right). \tag{81}$$

E. Authenticated Message Throughput

The authenticated message throughput with equal power allocation is:

$$W = E_{\lambda_1 \geq \frac{2^{R_t}-1}{(1-\alpha 2^{R_t})\gamma}}[\log_2(1 + \alpha\lambda_1\gamma/r)] + \sum_{i=2}^{r} E_{\lambda_i}[\log_2(1 + \lambda_i\gamma/r)], \tag{82}$$

where the constraint $\lambda_1 \geq \frac{2^{R_t}-1}{(1-\alpha 2^{R_t})\gamma}$ is for successful decoding of the signature.

F. Numerical Results

Fig. 13 presents a comparison of the total detection error probability between MIMO-MRC (also known as single-mode eigenbeamforming) and eigenbeamforming techniques. It is evident that eigenbeamforming results in a higher total detection error probability. This is primarily due to the interfering terms from $\mathbf{s}_2, \cdots, \mathbf{s}_r$ in (78), which make it more difficult for Eve to detect the signature embedded in $\mathbf{s}_1$. However, as the number of transmitter antennas increases, the total detection error probability for both techniques converges to similar values.

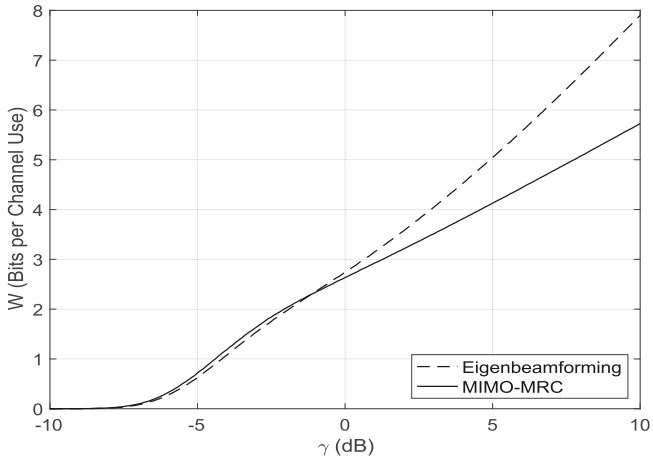

Fig. 15. Authenticated message throughput, W, versus SNR, γ (dB); $N_A = 4$, $N_B = 2$, $R_t = 0.1$, $\alpha = 0.9$.

Fig. 14 compares the total detection error probability ξ between eigenbeamforming and AN-aided transmission. The results show that the performance of the AN scheme is highly sensitive to the power allocated to AN. When $\phi = 0.99$, meaning only 1% of the power is allocated to AN, eigenbeamforming achieves a better covertness (higher total detection error probability) than the AN scheme. However, when $\phi = 0.9$, with 10% of the power allocated to AN, the AN-aided scheme yields better covertness. It is important to note that the AN approach requires a non-zero portion of the power to be reserved for AN to match the covertness performance of eigenbeamforming. In contrast, eigenbeamforming directs all power toward the legitimate receiver ($\phi = 1$), enabling it to deliver higher authenticated throughput than AN for the same level of covertness. This comparison highlights that eigenbeamforming is more efficient in scenarios prioritizing high throughput, whereas AN may be preferred in environments where enhanced covertness is critical and some throughput loss is acceptable.

Fig. 15 compares the authenticated message throughput of MIMO-MRC and eigenbeamforming techniques across various SNR ranges. At low SNR levels, MIMO-MRC performs slightly better than eigenbeamforming, primarily due to the suboptimality of equal power allocation in eigenbeamforming under these conditions. However, at high SNR levels, eigenbeamforming significantly outperforms MIMO-MRC in terms of authenticated message throughput, as equal power allocation becomes optimal in high SNR scenarios. These findings highlight the SNR-dependent performance of these techniques, with eigenbeamforming exhibiting superior efficiency in high SNR environments.

X. Conclusion

In this work, we proposed a novel covert message authentication method tailored to secure MIMO communications against integrity attacks. Our approach embeds the signature directly within the transmitted message, enabling covert verification without requiring additional power or bandwidth.

Our analysis revealed that the total detection error probability for the signature approaches unity, indicating that the signature is undetectable, as the number of transmitter antennas increases. This result is independent of the power allocated to the signature or the total transmission power, provided the transmitter has a sufficient number of antennas. This finding contrasts with earlier works, where the signature's transmit power had to be restricted to ensure its security against adversaries, which also constrained the intended receiver's (Bob's) ability to decode the signature. With our proposed scheme, covert authentication is achieved without reducing the transmission power of the signature, allowing Bob's decoding error probability of the signature to be reduced by increasing transmission power, without increasing the risk of detection by an adversary (Eve).

Additionally, we identified the optimal power allocation factor for the message that maximizes authenticated message throughput, which is closely approximated by 2^{-R_t} where R_t is the signature transmission rate. We further examined the interactions among signature covertness, signature decoding outage probability, and authenticated message throughput.

Furthermore, we analyzed the role of artificial noise in enhancing both the total detection error probability and the authenticated message throughput, particularly when Eve has more antennas than Alice. Finally, we compared two approaches for signature protection: signature secrecy, which relies on wiretap codes to keep the signature content confidential, and signature covertness, which conceals the existence of the signature itself. Our findings indicate that Eve is more likely to intercept information about the signature than to detect its transmission, highlighting the effectiveness of the proposed covert authentication scheme.

Appendix A

In this Appendix we provide the proof of (33). The cumulative distribution function of the Chi-square random variable is given by [33]

$$F_{\|\mathbf{g}_w\|^2}(u) = 1 - e^{-u} \sum_{k=0}^{N_E-1} \frac{u^k}{k!}. \tag{83}$$

Hence, we obtain

$$\int_{\alpha x}^{\infty} f_{\|\mathbf{g}_w\|^2}(y)dy = e^{-\alpha x} \sum_{k=0}^{N_E-1} \frac{(\alpha x)^k}{k!}. \tag{84}$$

Applying (84) and (35) to (32) yields

$$\begin{aligned} &\int_0^{\infty} \int_{\alpha x}^{\infty} f_{\|\mathbf{g}_w\|^2}(y)dy \cdot f_{\lambda_{\max}}(x)dx \\ &= \sum_{k=0}^{N_E-1} \frac{\alpha^k}{k!} \sum_{i=1}^{N_B} \sum_{m=N_A-N_B}^{(N_A+N_B)i-2i^2} \frac{d_{i,m} i^{m+1}}{m!} \\ &\cdot \int_0^{\infty} x^{k+m} e^{-(\alpha+i)x} dx \end{aligned} \tag{85}$$

$$\begin{aligned} &= \sum_{k=0}^{N_E-1} \sum_{i=1}^{N_B} \sum_{m=N_A-N_B}^{(N_A+N_B)i-2i^2} d_{i,m} i^{m+1} \\ &\cdot \frac{(k+m)!}{k!m!} \frac{\alpha^k}{(i+\alpha)^{k+m+1}} \end{aligned} \tag{86}$$

$$\begin{aligned} &= \sum_{k=0}^{N_E-1} \sum_{i=1}^{N_B} \sum_{m=N_A-N_B}^{(N_A+N_B)i-2i^2} d_{i,m} \frac{(k+m)!}{k!m!} \\ &\cdot \left(\frac{i}{i+\alpha}\right)^{m+1} \left(\frac{\alpha}{i+\alpha}\right)^k, \end{aligned} \tag{87}$$

where (86) follows from $\int_0^{\infty} x^n e^{-ax} dx = n!/a^{n+1}$.

Appendix B

In this Appendix, we prove that the signature transmission is perfectly covert, regardless of how often signature transmissions occur, when $P_F + P_M = 1$. Let $U = 0$ denote the event that the signature is absent ($\mathcal{H}_0$) and $U = 1$ denote the signature is present ($\mathcal{H}_1$). Similarly, let $V = 0$ denote the event that the detector decides that the signature is absent and $V = 1$ denote the event that the detector decides that the signature is present. Using these definitions, we derive the following conditional entropies:

$$H(U|V=0) = H_2\left(\frac{(1-P_F)\pi_0}{(1-P_F-P_M)\pi_0 + P_M}\right) \tag{88}$$

$$H(U|V=1) = H_2\left(\frac{P_F\pi_0}{(P_F+P_M-1)\pi_0 + 1 - P_M}\right), \tag{89}$$

where $H_2(x) = -x\log_2(x) - (1-x)\log_2(1-x)$ is the binary entropy function and π_0 is the probability that the signature is not embedded, i.e., $Pr(U = 0)$. Using these conditional entropies, the overall conditional entropy $H(U|V)$ is expressed as:

$$\begin{aligned} H(U|V) = &H(U|V=0)Pr(V=0) \\ &+ H(U|V=1)Pr(V=1). \end{aligned} \tag{90}$$

If $P_F + P_M = 1$, substituting into equations (88) and (89) into (90) yields:

$$H(U|V) = H_2(\pi_0) = H(U). \tag{91}$$

This result implies that the detector output (V) provides no information about the presence of the signature (U), regardless of π_0 (i.e., independent of how often signature transmissions occur), when $P_F + P_M = 1$.

REFERENCES

[1] C. Paar and J. Pelzl, *Understanding Cryptography: A Textbook for Students and Practitioners*. Cham, Switzerland: Springer, 2009.

[2] M. Mosca, "Cybersecurity in an era with quantum computers: Will we be ready?," *IEEE Secur. Privacy*, vol. 16, no. 5, pp. 38–41, Sep. 2018.

[3] L.-J. Wang et al., "Experimental authentication of quantum key distribution with post-quantum cryptography," *NPJ Quantum Inf.*, vol. 7, no. 1, p. 67, May 2021.

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

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

[6] 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.

[7] P. L. Yu, G. Verma, and B. M. Sadler, "Wireless physical layer authentication via fingerprint embedding," *IEEE Commun. Mag.*, vol. 53, no. 6, pp. 48–53, Jun. 2015.

[8] P. Zhang, J. Liu, Y. Shen, H. Li, and X. Jiang, "Lightweight tag-based PHY-layer authentication for IoT devices in smart cities," *IEEE Internet Things J.*, vol. 7, no. 5, pp. 3977–3990, May 2020.

[9] N. Xie, M. Sha, T. Hu, and H. Tan, "Multi-user physical-layer authentication and classification," *IEEE Trans. Wireless Commun.*, vol. 22, no. 9, pp. 6171–6184, Sep. 2023.

[10] N. Xie, C. Chen, and Z. Ming, "Security model of authentication at the physical layer and performance analysis over fading channels," *IEEE Trans. Depend. Secure Comput.*, vol. 18, no. 1, pp. 253–268, Jan. 2021.

[11] N. Xie, Z. Li, and H. Tan, "A survey of physical-layer authentication in wireless communications," *IEEE Commun. Surveys Tuts.*, vol. 23, no. 1, pp. 282–310, 1st Quart., 2020.

[12] 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.

[13] P. L. Yu, B. M. Sadler, G. Verma, and J. S. Baras, "Fingerprinting by design: Embedding and authentication," *Digit. Fingerprinting*, vol. 1, pp. 69–88, Jan. 2016.

[14] S.-Y. Wang and M. R. Bloch, "Covert MIMO communications under variational distance constraint," *IEEE Trans. Inf. Forensics Security*, vol. 16, pp. 4605–4620, 2021.

[15] A. Bendary, A. Abdelaziz, and C. E. Koksal, "Achieving positive covert capacity over MIMO AWGN channels," *IEEE J. Sel. Areas Inf. Theory*, vol. 2, no. 1, pp. 149–162, Mar. 2021.

[16] L. Bai, J. Xu, and L. Zhou, "Covert communication for spatially sparse mmWave massive MIMO channels," *IEEE Trans. Commun.*, vol. 71, no. 3, pp. 1615–1630, Mar. 2023.

[17] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Hoboken, NJ, U.S.: Wiley-Interscience, 2006.

[18] H. Ochiai and H. Imai, "On the distribution of the peak-to-average power ratio in OFDM signals," *IEEE Trans. Commun.*, vol. 49, no. 2, pp. 282–289, Feb. 2001.

[19] G. Böcherer, F. Steiner, and P. Schulte, "Bandwidth efficient and rate-matched low-density parity-check coded modulation," *IEEE Trans. Commun.*, vol. 63, no. 12, pp. 4651–4665, Dec. 2015.

[20] P. A. Dighe, R. K. Mallik, and S. S. Jamuar, "Analysis of transmit-receive diversity in Rayleigh fading," *IEEE Trans. Commun.*, vol. 51, no. 4, pp. 694–703, Apr. 2003.

[21] M. Kang and M.-S. Alouini, "A comparative study on the performance of MIMO MRC systems with and without cochannel interference," *IEEE Trans. Commun.*, vol. 52, no. 8, pp. 1417–1425, Aug. 2004.

[22] J. Tang et al., "Associating MIMO beamforming with security codes to achieve unconditional communication security," *IET Commun.*, vol. 10, no. 12, pp. 1522–1531, Aug. 2016.

[23] E. Axell, G. Leus, E. G. Larsson, and H. V. Poor, "Spectrum sensing for cognitive radio: State-of-the-art and recent advances," *IEEE Signal Process. Mag.*, vol. 29, no. 3, pp. 101–116, May 2012.

[24] S. W. Kim and H. Q. Ta, "Covert communications over multiple overt channels," *IEEE Trans. Commun.*, vol. 70, no. 2, pp. 1112–1124, Feb. 2022.

[25] B. A. Bash, D. Goeckel, and D. Towsley, "Limits of reliable communication with low probability of detection on AWGN channels," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 9, pp. 1921–1930, Sep. 2013.

[26] M. R. Bloch, "Covert communication over noisy channels: A resolvability perspective," *IEEE Trans. Inf. Theory*, vol. 62, no. 5, pp. 2334–2354, May 2016.

[27] T. M. Semkow et al., "Chi-square distribution: New derivations and environmental application," *J. Appl. Math. Phys.*, vol. 7, no. 8, pp. 1786–1799, 2019.

[28] I. M. Johnstone, "On the distribution of the largest eigenvalue in principal components analysis," *Ann. Statist.*, vol. 29, no. 2, pp. 295–327, Apr. 2001.

[29] A. D. Wyner, "The wire-tap channel," *Bell Syst. Tech. J.*, vol. 54, no. 8, pp. 1355–1387, Oct. 1975.

[30] F. Oggier and B. Hassibi, "The secrecy capacity of the MIMO wiretap channel," *IEEE Trans. Inf. Theory*, vol. 57, no. 8, pp. 4961–4972, Aug. 2011.

[31] X. Zhou et al., "Rethinking the secrecy outage formulation: A secure transmission design perspective," *IEEE Commun. Lett.*, vol. 15, no. 3, pp. 302–304, Mar. 2011.

[32] B. He, A. Liu, N. Yang, and V. K. N. Lau, "On the design of secure non-orthogonal multiple access systems," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 10, pp. 2196–2206, Oct. 2017.

[33] J. G. Proakis, *Digital Communications*. New York, NY, USA: McGraw-Hill, 2008.

Sang Wu Kim (Senior Member, IEEE) received the B.S. degree in electronic engineering from Yonsei University, Seoul, South Korea, in 1981, the M.S. degree in electrical engineering from KAIST, in 1983, and the Ph.D. degree in electrical engineering from the University of Michigan, Ann Arbor, in 1987. He is currently a Professor at the Department of Electrical and Computer Engineering, Iowa State University. He previously held academic and industry positions, including a Professor at Korea Advanced Institute of Science and Technology (KAIST), Daejeon, South Korea; a Visiting Associate Professor at California Institute of Technology (Caltech), Pasadena; and a Technical Consultant with the Wireless Systems Research Department, AT&T Labs, Middletown, NJ. His research interests include wireless communications and security. He received multiple honors for his contributions, including the Best Paper Award at the IEEE International Symposium on Spread Spectrum Techniques and Applications (ISSSTA), Parsippany, NJ, in 2000, the Best Paper Award at the IEEE International Conference on Mobile Ad Hoc and Sensor Systems (MASS), in 2005, and the Best Poster Award at the IEEE Conference on Sensor, Mesh, and Ad Hoc Communications and Networks (SECON), in 2007. He was the Technical Program Committee Co-Chair of the IEEE GLOBECOM Communication Theory Symposium in 2013 and the Publications Chair for IEEE WCNC in 2017. He served as an Associate Editor for IEEE COMMUNICATIONS LETTERS (2000–2006) and IEEE COMMUNICATIONS SURVEYS AND TUTORIALS (2013–2023).