Received 5 August 2021; accepted 7 September 2021. Date of publication 14 September 2021; date of current version 22 September 2021.

Digital Object Identifier 10.1109/OJCOMS.2021.3112486

Artificially Time-Varying Differential MIMO for Achieving Practical Physical Layer Security

NAOKI ISHIKAWA [1] (Member, IEEE), JEHAD M. HAMAMREH [2] (Member, IEEE), EIJI OKAMOTO [3] (Member, IEEE), CHAO XU [4] (Senior Member, IEEE), AND LIXIA XIAO [5] (Member, IEEE)

[1]Faculty of Engineering, Yokohama National University, Kanagawa 240-8501, Japan

[2]Department of Electrical and Electronics Engineering, Antalya Bilim University, 07468 Antalya, Turkey

[3]Department of Electrical and Mechanical Engineering, Nagoya Institute of Technology, Nagoya 466-8555, Japan

[4]School of Electronics and Computer Science, University of Southampton, Southampton SO17 1BJ, U.K.

[5]Wuhan National Laboratory for Optoelectronics, Huazhong University of Science and Technology, Wuhan 430074, China

CORRESPONDING AUTHOR: N. ISHIKAWA (e-mail: ishikawa-naoki-fr@ynu.ac.jp)

The work of Naoki Ishikawa was supported in part by the Japan Society for the Promotion of Science (JSPS) KAKENHI under Grant 19K14987. The work of Jehad M. Hamamreh was supported in part by the Scientific and Technological Research Council of Turkey (TUBITAK) under Grant 119E392. The work of Lixia Xiao was supported in part by the National Science Foundation of China under Grant 62001179.

ABSTRACT In this paper, we propose a differential multiple-input multiple-output (MIMO) scheme based on the novel concept of chaos-based time-varying unitary matrices to demonstrate—for the first time in the literature—the ability of differential encoding in achieving practical physical layer security even without the need for using channel estimation. In the proposed scheme, an erroneous secret key, which is extracted from the wireless nature, is used to initialize a chaos sequence that is responsible for generating artificially time-varying unitary matrices capable of obfuscating the transmitted data symbols from illegitimate eavesdroppers. Contrary to conventional studies, the key agreement ratio in this study is assumed to be imperfect, which is often true and very realistic in high-mobility scenarios. Following this, we conceive a new calibration algorithm for reconciling the chaotic sequence generated at the legitimate parties, thus making this calibration algorithm a unique, novel solution to the key sharing problem of conventional chaos-based communication techniques, which has been overlooked over the past few decades. It is found out that differential encoding obviates additional complexity and insecurity in dealing with channel estimation, whereas an eavesdropper must tackle the complicated differentially encoded patterns, which have an exponentially increasing complexity order. In addition, the obtained simulation results demonstrate that the proposed scheme can outperform conventional chaos-based MIMO schemes that assume perfect channel knowledge.

INDEX TERMS MIMO, differential modulation, differential space-time block codes, physical layer security, physical layer encryption, chaos theory, phase ambiguity, constrained capacity, secrecy rate, security gap.

I. INTRODUCTION

RADIO waves can propagate over long distances. Even a low-power signal that is transmitted by a household Wi-Fi device can reach as far as 100 meters if a line of sight path is present [1]. In public Wi-Fi networks, it is easy for eavesdroppers to obtain standardized 802.11 frames and retrieve private information streams [2]–[4]. Because the private information is encrypted in the transport layer, we feel safe using wireless network, but this security is not guaranteed forever. For example, the widely used public-key encryption method, RSA [5], has been threatened by the invention of Shor's algorithm [6], which performs integer factoring in polynomial time using a quantum computer [7]. Therefore, it is necessary to invent practical wireless communication method in the physical layer that reinforces security.

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

Operational wireless systems generally rely on encryption-based methods to secure communications, where a secret key is exchanged in advance. Physical layer communication schemes that require a perfect secret key are classifiable into the physical layer encryption (PLE) category [8]. As a pioneering PLE study, Dedieu *et al.* proposed a seminal chaos-based communication system [9], designated as *chaos shift keying* (CSK). The original CSK system of [9] was proposed for wired communications using Chua's analog circuit, which generates a chaotic carrier. Based on this CSK philosophy [9], Okamoto *et al.* proposed the chaos MIMO technique [10]–[13]. This technique generates a Gaussian-distributed constellation that is difficult to perceive and eavesdrop because the Gaussian symbols are naturally hidden by additive Gaussian noise. Furthermore, the chaos MIMO arrangement obtains channel coding gain by exploiting a unique chaos modulation structure. In parallel, Kaddoum *et al.* proposed the MIMO-CSK concept [14] for a 2×2 setup, where chaos is used to spread data symbols. Note that all the above chaos-based schemes [10]–[14] require precise estimates of channel state information (CSI). The estimation of CSI imposes potential risk as the eavesdropper may use pilot symbols to synchronize received signals and obtain accurate CSI between the transmitter and the eavesdropper [15], [16].

Key-based encryption techniques that rely on computational security may be cracked by future supercomputers [8], [17], [18]. In order to overcome this limitation, physical layer security (PLS) methods that frequently update private keys have been conceived [19], which rely on the true randomness of wireless channels. The information-theoretic foundation of secret-key agreement in public channels was first established by Maurer and Wolf [20]–[22]. Most of secret-key agreement or generation methods require the assumptions of time division duplex (TDD) channel reciprocity and near-perfect CSI [23]–[25]. The strong assumptions on CSI have hindered the industrial applications of PLS [26]. When considering realistic CSI errors, the key agreement ratio between legitimate parties is far from perfect in practice [27]–[29]. Hence, it is a challenging task to achieve the key agreement ratio of 100% in high-mobility scenarios. A long-overlooked problem here is that this erroneous key cannot be used for all the conventional encryption methods. To improve the key agreement ratio, the legitimate parties have to exchange thousands of probe symbols [27]. We have to eliminate this channel estimation process because it increases both the communication overhead and risk.

The classic differential MIMO can eliminate the channel estimation process [30], [31]. The pioneering scheme [30] established in the early 2000s relies on square unitary matrices. By contrast, in 2017, a new approach of *non-square differential MIMO* [32]–[36] was proposed. This new approach maps the classic unitary matrix to a nonsquare matrix, which beneficially improves the transmission rate linearly. Because of the differential structure, the resultant constellation might reach an infinite cardinality [37]–[39]. This structure is naturally useful for improving the wireless communication security. Nevertheless, no report of the relevant literature has described a study of the differential MIMO in the context of both PLE and PLS.[1]

Against this background, we propose a chaos-based differential MIMO system that is free from the additional complexity and insecurity of dealing with channel estimation. More explicitly, the data-carrying matrices are obfuscated using a specially designed *artificially time-varying unitary matrix* concept, which is generated by a chaos sequence. Unlike the conventional chaos-based PLE family [10]–[14], [40]–[43], our proposed system extracts a noisy key from the wireless channel, which is used as the initial condition of a chaos sequence. Following this invention, a low-complexity *chaos calibration* is conceived to continue estimation of the original chaos sequence. We also conceive a simple real-valued key generation method and analyze the minimum security level achieved by the proposed system.

The major contributions of this paper are summarized in twofold.

1) Our proposed scheme is the first chaos-based scheme that is free from key establishment in advance. All of the schemes of the conventional chaos-based family must rely on the perfect pre-shared key because the chaos sequence is sensitive to error. For example, the small error of 2^{-1022} added to the initial condition causes mismatches between the legitimate parties.[2] We resolve this issue by conceiving a chaos calibration algorithm that updates a chaos sequence at the receiver. This calibration process can be interpreted as information reconciliation of a chaotic sequence. The advantage of this process is that there is no additional overhead because the reconciliation is performed using data matrices instead of transmitting probe signals.
2) We prove and reveal for the first time that the classic differential encoding is suitable for achieving practical PLS. The security level depends on the length of differentially encoded matrices. This special encoding expands the search space exponentially. Most existing PLS methods assume perfect knowledge of CSI both at transmitter and receiver, which is impractical in high-mobility scenarios. We resolve this CSI issue by the nonsquare differential structure, which mitigates both the communication overhead and risk.

We must report that the proposed arrangement also has the following shortcomings, as discussed in Section VI.

1) For a case in which the eavesdropper is in the same position as the legitimate receiver, the proposed scheme cannot provide security because the eavesdropper would have a near-perfect estimate of the channel coefficients

1. The differential counterpart of MIMO-CSK [40] has also been proposed for the synchronizing a chaos sequence both at a transmitter and receiver. However, it is noteworthy that the term *differential* differs from the classic modulation definition, as described in Section III-B

2. 2^{-1022} is the absolute minimum of a 64-bit floating point number, which is specified by IEEE 754.

between the legitimate transmitter and receiver. In this case, the legitimate user can physically eliminate the eavesdropping device or can halt secret communications.

2) Despite the fact that the search space is increased exponentially because of the employment of differential encoding, the proposed scheme might become susceptible to a brute-force attack when the length of differentially encoded matrices is short. In this unrealistic case, an eavesdropper having perfect channel state information (PCSI) can obtain a part of private information.

The remainder of this paper is organized as follows. Section II defines the common system model used for this study. Section III reviews the classic chaos theory and the conventional chaos-based MIMO schemes. Section IV proposes our chaos-based differential MIMO that relies on the novel time-varying basis and the calibration algorithm. Section V presents an attack algorithm for the proposed scheme. Section VI demonstrates the performance superiority over conventional schemes in terms of secrecy rate and reliability. Finally, Section VII concludes this paper.

We note that italicized symbols represent scalar values. Bold symbols represent vectors and matrices. Table 1 presents a list of mathematical symbols used for this study.

II. SYSTEM MODEL

This section presents a description of a general system model common to Sections III and IV. Without loss of generality, a narrow-band system model is considered in this paper, but extension to the wide-band scenario in the context of orthogonal frequency division multiplexing (OFDM) is straightforward.[3]

We assume that the legitimate transmitter, *Alice*, is equipped with M antennas, whereas the legitimate receiver, *Bob*, is equipped with N antennas. Additionally, we assume that eavesdropper *Eve* has unlimited capabilities of computers, such as cloud-based computing resources and supercomputers. The received signal block at Bob is given as [44]

$$\mathbf{Y}(i) = \mathbf{H}(i)\mathbf{S}(i) + \mathbf{V}(i) \in \mathbb{C}^{N\times M}, \tag{1}$$

where i represents a transmission index, $\mathbf{H}(i) \in \mathbb{C}^{N\times M}$ denotes a channel matrix that obeys the i.i.d. Rayleigh fading $\mathcal{CN}(0, 1)$, and $\mathbf{S}(i) \in \mathbb{C}^{M\times T}$ stands for a space-time codeword. The codeword $\mathbf{S}(i)$ is transmitted by M antennas over T time slots. Specific construction methods for $\mathbf{S}(i)$ and the corresponding detectors are described in Sections III and IV. Furthermore, the additive noise $\mathbf{V}(i)$ is assumed to follow the i.i.d. complex Gaussian distribution, $\mathcal{CN}(0, \sigma_v^2)$. The signal-to-noise ratio (SNR) is calculated as $1/\sigma_v^2$, i.e., $10 \cdot \log_{10}(1/\sigma_v^2)$ [dB] because we have the power constraint of $\mathrm{E}_i[\|\mathbf{S}(i)\|_\mathrm{F}^2]/T = 1$. The transmission index starts from $i = 1$ and finishes at a frame length $i = W$. After extracting a new private key from the channel, it restarts from $i = 1$.

3. Note that the nonsquare differential coding described later is particularly suitable for high-mobility OFDM scenarios [33], [34].

TABLE 1. List of important mathematical symbols.

$\mathbb{B}$		Binary numbers
$\mathbb{R}$		Real numbers
$\mathbb{C}$		Complex numbers
$\mathbb{Z}$		Integers
M	$\in \mathbb{Z}$	Number of transmit antennas
N	$\in \mathbb{Z}$	Number of receive antennas
T	$\in \mathbb{Z}$	Number of time slots in a codeword
W	$\in \mathbb{Z}$	Frame length
D	$\in \mathbb{Z}$	Number of data-carrying codewords ($= W - M$)
B	$\in \mathbb{Z}$	Input bitwidth
R	$\in \mathbb{R}$	Transmission rate
R^{eff}	$\in \mathbb{R}$	Effective transmission rate
Y	$\in \mathbb{Z}$	Non-zero integer value that determines security
$\mathbf{Y}(i)$	$\in \mathbb{C}^{N\times T}$	Bob's received signal block
$\mathbf{H}(i)$	$\in \mathbb{C}^{N\times M}$	Bob's channel matrix
$\mathbf{V}(i)$	$\in \mathbb{C}^{N\times T}$	Bob's additive noise
$\mathbf{Y}_\mathrm{E}(i)$	$\in \mathbb{C}^{N\times T}$	Eve's received signal block
$\mathbf{H}_\mathrm{E}(i)$	$\in \mathbb{C}^{N\times M}$	Eve's channel matrix
$\mathbf{V}_\mathrm{E}(i)$	$\in \mathbb{C}^{N\times T}$	Eve's additive noise
$\mathbf{S}(i)$	$\in \mathbb{C}^{M\times T}$	Space-time codeword
$\mathbf{X}(i)$	$\in \mathbb{C}^{M\times M}$	Space-time codeword
$\tilde{\mathbf{S}}(i)$	$\in \mathbb{C}^{M\times M}$	Unitary space-time codeword
$\mathbf{E}_1(i)$	$\in \mathbb{C}^{M\times 1}$	Square-to-nonsquare projection
$\mathbf{W}_1$	$\in \mathbb{C}^{M\times 1}$	First column of DFT matrix
$\hat{\mathbf{Y}}(i)$	$\in \mathbb{C}^{N\times M}$	Estimation of $\mathbf{H}(i)\tilde{\mathbf{S}}(i)$
$\mathbf{b}$	$\in \mathbb{B}^B$	B-length input bits
σ_v^2	$\in \mathbb{R}$	Noise variance
ϵ_e	$\in \mathbb{R}$	Accuracy of shared real-valued keys
ρ	$\in \mathbb{R}$	Channel correlation coefficient
n	$\in \mathbb{Z}$	Bitwidth of a floating-point number
i	$\in \mathbb{Z}$	Transmission index ($\leq W$)
h, h'	$\in \mathbb{C}$	Single channel coefficient
$\alpha(i)$	$\in \mathbb{R}$	Forgetting factor
x_i	$\in \mathbb{R}$	Pure chaos solution
$\hat{x}_i$	$\in \mathbb{R}$	Calibrated counterpart of x_i
x_i^t, x_i^r	$\in \mathbb{R}$	x_i at transmitter and receiver
x_i'	$\in \mathbb{R}$	Second-order Chebyshev polynomial function
y_i	$\in \mathbb{R}$	Chaos sequence having a uniform distribution
$\hat{y}_i$	$\in \mathbb{R}$	Calibrated counterpart of y_i

Similar to Bob, the received signal block at Eve is given as

$$\mathbf{Y}_\mathrm{E}(i) = \mathbf{H}_\mathrm{E}(i)\mathbf{S}(i) + \mathbf{V}_\mathrm{E}(i) \in \mathbb{C}^{N\times M}, \tag{2}$$

where Eve's channel matrix is defined as

$$\mathbf{H}_\mathrm{E}(i) = \rho\mathbf{H}(i) + \sqrt{1-\rho^2}\mathbf{H}'(i), \tag{3}$$

and $\mathbf{H}'(i)$ is an independent channel matrix following $\mathcal{CN}(0, 1)$. Here, the channel correlation ρ represents the similarity between the Alice–Bob channel $\mathbf{H}(i)$ and the Alice–Eve channel $\mathbf{H}_\mathrm{E}(i)$. When Eve is near Bob, ρ is close to 1. Eve's channel matrix $\mathbf{H}_\mathrm{E}(i)$ becomes similar to Bob's channel matrix $\mathbf{H}(i)$.

In this paper, we model calculation complexity based on Donald Knuth's big Omega $\Omega(\cdot)$ notation [45], which represents an asymptotic lower bound. Conventional studies [44] often only evaluate the total number of real-valued multiplications. They ignore other operations such as division and elementary functions, which are as costly as multiplication. This asymptotic analysis is useful for estimating the implementation complexity and power consumption of circuits [44], [46]. According to [47], the addition and subtraction cost $\Omega(n)$, where n is the bitwidth of a floating-point number. The multiplication costs $\Omega(n \log n)$, while the division costs $\Omega(n \log n \log n)$. The square root operation $\sqrt{\cdot}$ costs $\Omega(n \log n)$. The elementary functions such as $\exp(\cdot)$, $\arcsin(\cdot)$, and $\arctan(\cdot)$ cost $\Omega(n \log n \log n)$. For example, the complexity of $(a + bj)(a' + b'j)$ for $a, b, a', b' \in \mathbb{R}$ is calculated as $4n \log n + 4n \geq \Omega(n \log n)$. The calculation of $\mathbf{H}(i)\mathbf{S}(i)$ costs $NT(4Mn \log n + 2(M-1)n) \geq \Omega(NTMn \log n)$.

III. CONVENTIONAL CHAOS THEORY AND ITS APPLICATIONS TO MIMO

Since 1993, chaos theory has been applied to wireless communications for enhancing security [9]–[15], [41]–[43]. Although a chaotic mathematical model is deterministic, it is sensitive to initial conditions. Moreover, it is almost impossible to predict future trajectory. This sensitivity can be quantified by the Lyapunov exponent [48]. Additionally, the chaotic sequence is bounded within a region and is non-periodic. Therefore, the initial condition works as a secret key for secure communications.

In von Neumann's seminal study [49], the logistic map

$$x_{i+1} = 4x_i(1 - x_i) \tag{4}$$

is used to generate a random digit. The initial condition x_0 must be $0 < x_0 < 1$. The chaotic sequence of (4) is called a pure chaos solution. Its Lyapunov exponent is calculated as $\log_e(2) \approx 0.6931$, which is higher than the value of 0.0714 [48] of the simplified Rossler equation [50]. The probabilistic distribution of x_i is a non-uniform function of [49]

$$p(x) = \frac{1}{\pi\sqrt{x(1-x)}}, \tag{5}$$

which differs from a uniform distribution. Because the exact solution of (4) is given as [51][4]

$$x_i = \sin^2 2^i \arcsin\left(\sqrt{x_0}\right) := f(i, x_0), \tag{6}$$

the chaotic sequence of x_i can be transformed into a uniform distribution as [51]

$$y_i = \frac{2}{\pi} \arcsin \sqrt{x_i}, \tag{7}$$

where we have the uniform distribution of $p(y) = 1$ for $0 < y < 1$.

As a simple example, Fig. 1 portrays the transition of x_i, defined in (4), where the index is increased from $i = 0$ to 25. Three initial values were considered as shown in Fig. 1: $x_0 = 0.24$, $0.24 + 10^{-3}$ and $0.24 + 10^{-6}$. As shown in Fig. 1, both the initial values $x_0 = 0.24$ and $0.24 + 10^{-6}$ caused almost identical x_i for $0 \leq i \leq 13$, whereas both sequences exhibited significantly different transitions for $i > 13$. In the $x_0 + 10^{-3}$ case, it exhibited different transitions for $i > 4$. Therefore, the accuracy of the initial value determines the agreement interval of the chaos sequence. Based on this fact, we propose a new practical calibration algorithm in Section IV-F.

4. This equation might overflow because of the calculation of 2^i. For our simulations, we calculate $f(i, x)$ in a recursive manner, i.e., $f(0, x) = x$ and $f(i, x) = 4f(i-1, x)(1 - f(i-1, x))$.

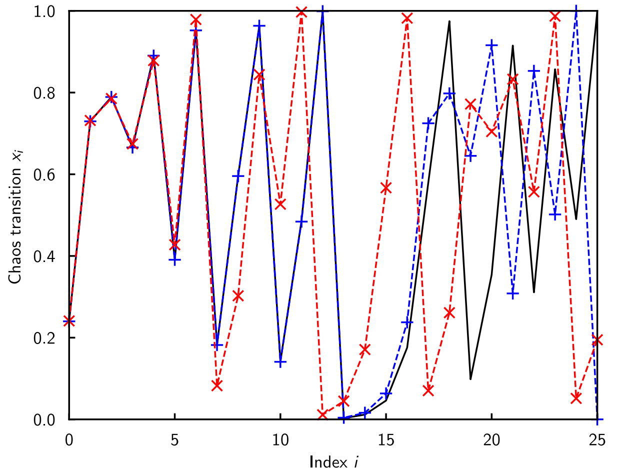


FIGURE 1. Chaos transition x_i over time, for which the initial values of $x_0 = 0.24$, $x_0 + 10^{-6}$, and $x_0 + 10^{-3}$ were considered.

A. MIMO-CSK [14]

The MIMO-CSK scheme of [14] uses the second-order Chebyshev polynomial function of

$$x'_i = 1 - 2x'^2_{i-1}, \tag{8}$$

where the initial condition x'_0 must be within $[-1, 1]$. It is noteworthy that the mean of x'_i is zero. The variance is 0.5. The original contribution of [14] considers the direct sequence spread spectrum instead of OFDM. The chaotic sequence of (8) is used as a spread sequence over the time domain. The spreading factor is varied from $\beta = 2$ to 50 in [14]. As described in this paper, we limit the spreading factor to $\beta = 1$ for simplicity.[5] In this case, the 2×2 space-time matrix is generated by $\mathbf{S}(i) = \mathbf{P}(i)\mathbf{X}(i)$ [14], where we have

$$\mathbf{P}(i) = \sqrt{2} \cdot \mathrm{diag}\left(x'_{2i}, x'_{2i+1}\right) \tag{9}$$

and the orthogonal space-time block code (OSTBC) of [52]

$$\mathbf{X}(i) = \frac{1}{\sqrt{2}} \begin{bmatrix} s_1(i) & -s_2^*(i) \\ s_2(i) & s_1^*(i) \end{bmatrix}. \tag{10}$$

Here, $s_1(i)$ and $s_2(i)$ denote complex-valued symbols with L-ary phase-shift keying (PSK) or quadrature amplitude modulation (QAM). The transmission rate of (10) is $R = \log_2(L)$. Note that the mean transmission power is calculated as $\mathrm{E}_i[\|\mathbf{S}(i)\|_\mathrm{F}^2]/T = 2/2 = 1$, which is the same as in other MIMO techniques. The maximum-likelihood detector is given as

$$\hat{\mathbf{X}}(i) = \arg\min_{\mathbf{X}} \|\mathbf{Y}(i) - \mathbf{H}(i)\mathbf{P}(i)\mathbf{X}\|_\mathrm{F}^2, \tag{11}$$

where the chaotic sequence $\mathrm{diag}(x'_{2i}, x'_{2i+1})$ is known perfectly at the receiver with no noise.

5. As one might expect, this limitation worsens performance. Although MIMO-CSK is the product of an important pioneering study, it is not a performance baseline.

Although the original contributions of [14] considered only the $M = 2$ case, it can be extended to the $M > 2$ cases as

$$\mathbf{S}(i) = \mathbf{P}(i)\mathbf{X}(i), \tag{12}$$

where we have

$$\mathbf{P}(i) = \sqrt{2} \cdot \mathrm{diag}\left(x'_{i \cdot M}, x'_{i \cdot M+1}, \ldots, x'_{i \cdot M+M-1}\right) \tag{13}$$

and an $M \times M$ data-carrying matrix $\mathbf{X}(i)$. We must multiply $\sqrt{2}$ in (13) for any M because the variance of x'_i is 0.5. For example, the Bell Laboratories layered space-time scheme [53] is defined as $\mathbf{X}(i) = [s_1(i)\ s_2(i)\ \cdots\ s_M(i)]^{\mathrm{T}}/\sqrt{M} \in \mathbb{C}^{M \times 1}$. In the OSTBC case having $M = 4$, the data-carrying matrix is given as [54]

$$\mathbf{X}(i) = \frac{1}{\sqrt{2}} \begin{bmatrix} s_1(i) & -s_2^*(i) & 0 & 0 \\ s_2(i) & s_1^*(i) & 0 & 0 \\ 0 & 0 & s_1(i) & -s_2^*(i) \\ 0 & 0 & s_2(i) & s_1^*(i) \end{bmatrix}. \tag{14}$$

The transmission rate of (14) is $R = \log_2(L)/2$. Following the complexity model described in Section II, the detection complexity of (11) is lower bounded by $\Omega(2^R NMn \log n)$.

B. MIMO-DCSK [41]

The MIMO differential CSK (DCSK) scheme [41] has been proposed for spread spectrum communications, which require no CSI either at the transmitter or receiver. The conventional MIMO-CSK [14] requires the receiver to reproduce the original chaos sequence generated by the transmitter. To address this synchronization issue, Kaddoum *et al.* proposed DCSK for a MIMO setup. It generates a space-time codeword of [41]

$$\mathbf{S}(i) = \begin{bmatrix} x'_{4i} & s_1(i)x'_{4i+1} & x'_{4i+2} & -s_2^*(i)x'_{4i+3} \\ x'_{4i} & s_2(i)x'_{4i+1} & x'_{4i+2} & s_1^*(i)x'_{4i+3} \end{bmatrix} \tag{15}$$

when the spreading factor is 1. By transmitting a chaos sequence directly, this scheme helps the receiver to reproduce the chaos sequence. Such DCSK-related studies [41], [43], [55], [56] differ from the differential STBC concept established in the early 2000s [30], [31]. Therefore, we do not consider the DCSK family in our performance comparisons.

C. C-MIMO [10]–[13]

The C-MIMO scheme has been proposed for improving the security of MIMO communications, where PCSI is necessary at the receiver. An initial condition is generated using a pre-shared key. The key is processed many times by the Dirac transformation [10], [13]. The resultant constellation follows the complex-valued Gaussian distribution.

The C-MIMO scheme requires a pre-shared key $c_0 \in \mathbb{C}$ that obeys $0 < \mathrm{Re}[c_0] < 1$ and $0 < \mathrm{Im}[c_0] < 1$. The $B = MT$-length input $\mathbf{b} = [b_1, b_2, \ldots, b_B] \in \mathbb{B}^B$ is mapped to a set of complex-valued symbols $\mathbf{s} = [s_1, s_2, \ldots, s_B] \in \mathbb{C}^B$. This set is then mapped to an $M \times T$ space-time codeword. Each symbol s_k for $k = 1, 2, \ldots$ Also, B is defined by two independent chaos sequences $\mathrm{Re}[z_l]$ and $\mathrm{Im}[z_l]$. Both sequences are initialized by $\mathrm{Re}[z_0] = \Gamma(\mathrm{Re}[c_{k-1}], b_{k-1})$ and $\mathrm{Im}[z_0] = \Gamma(\mathrm{Im}[c_{k-1}], b_{k \bmod B})$, where we have [13]

$$\Gamma(a, b) = \begin{cases} a, & (b = 0) \\ 1 - a, & (b = 1 \text{ and } a > 1/2) \\ a + 1/2, & (b = 1 \text{ and } a \leq 1/2). \end{cases} \tag{16}$$

Then, both sequences are generated as[6]

$$\begin{aligned} \mathrm{Re}[z_l] &= 2 \cdot \mathrm{Re}\left[z_{l-1}\right] \bmod \left(1 - 10^{-16}\right) \quad \text{and} \\ \mathrm{Im}[z_l] &= 2 \cdot \mathrm{Im}\left[z_{l-1}\right] \bmod \left(1 - 10^{-16}\right) \end{aligned} \tag{17}$$

for $l = 1, 2, \ldots, N_s, N_s + 1$, and $N_s = 100$ [13]. The C-MIMO symbol s_k for $k = 1, 2, \ldots, B$ is given by [13]

$$s_k = \sqrt{-\log\left(c_k^{(x)}\right)}\left(\cos\left(2\pi c_k^{(y)}\right) + j \sin 2\left(\pi c_k^{(y)}\right)\right), \tag{18}$$

where we have [13]

$$\begin{cases} c_k = \mathrm{Re}\left[z_{N_s + b_{(k+B/2) \bmod B}}\right] + j\mathrm{Im}\left[z_{N_s + b_{(k+B/2+1) \bmod B}}\right] \\ c_k^{(x)} = \arccos(\cos(37\pi(\mathrm{Re}[c_k] + \mathrm{Im}[c_k])))/\pi \\ c_k^{(y)} = \arcsin(\sin(43\pi(\mathrm{Re}[c_k] - \mathrm{Im}[c_k])))/\pi + \frac{1}{2}. \end{cases} \tag{19}$$

By virtue of the Box–Muller transform in (18), s_k follows the complex Gaussian distribution $\mathcal{CN}(0, 1)$. Finally, the codeword associated with $B = MT$-length bits $\mathbf{b}$ is given as

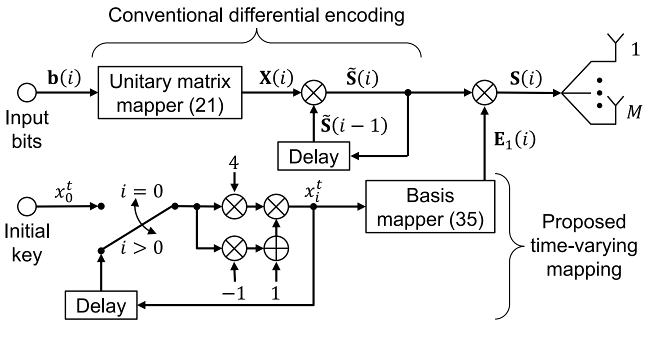
$$\mathbf{S}(i) = \frac{1}{\sqrt{M}} \begin{bmatrix} s_1 & s_{M+1} & \cdots & s_{MT-M+1} \\ s_2 & s_{M+2} & \cdots & s_{MT-M+2} \\ \vdots & \vdots & \ddots & \vdots \\ s_M & s_{2M} & \cdots & s_{MT} \end{bmatrix} \in \mathbb{C}^{M \times T}. \tag{20}$$

The normalized transmission rate is calculated as $R = M$ [bit/symbol]. The detection complexity is lower bounded by $\Omega(2^{RT} NMn \log n)$, where the complexity of generating (20) is ignored for simplicity.

IV. PROPOSED CHAOS-BASED DIFFERENTIAL MIMO

The proposed scheme has a common structure with the conventional nonsquare differential scheme of [32]–[34], [36]. It invokes a chaos-based time-varying basis and a chaos calibration algorithm. Fig. 2 shows (a) the transmitter and (b) the receiver of our proposed system. As shown in Fig. 2, our system extracts a secret initial key from the wireless channel, which is denoted by x_0^t at the transmitter and x_0^r at the receiver. At the transmitter, an input bit sequence $\mathbf{b}(i)$ is mapped to a differentially-encoded square matrix $\tilde{\mathbf{S}}(i)$. In parallel, a chaos sequence x_i^t is used to generate a time-varying basis $\mathbf{E}_1(i) \in \mathbb{C}^{M \times 1}$. Then, the square matrix $\tilde{\mathbf{S}}(i) \in \mathbb{C}^{M \times M}$ is mapped into a nonsquare matrix $\tilde{\mathbf{S}}(i)\mathbf{E}_1(i) \in \mathbb{C}^{M \times 1}$. Conventional studies [32]–[34] adopted a static basis $\mathbf{E}_1 \in \mathbb{C}^{M \times 1}$ instead of this time-varying counterpart. At the receiver, the chaos sequence x_i^r is initialized

6. The modulo operation is extended to real numbers as $x \bmod y := x - y \cdot \lfloor x/y \rfloor$.



(a) Transmitter

(b) Receiver

FIGURE 2. **Schematic of the proposed system.**

by x_0^r. It is used to estimate the private bit sequence $\hat{\mathbf{b}}(i)$. Because the chaos sequence x_i^r might include errors, it is calibrated by the proposed algorithm and is used for the next time slot.

A. ENCODING AT THE TRANSMITTER

We first introduce the encoding process which supports the time-varying basis. The B-length input bit sequence $\mathbf{b} \in \mathbb{B}^B$ is associated with an $M \times M$ square data-carrying matrix of [30]

$$\mathbf{X}(i) = \mathrm{diag}\left[\exp\left(j\frac{2\pi b}{2^B}u_1\right), \ldots, \exp\left(j\frac{2\pi b}{2^B}u_M\right)\right] \tag{21}$$

$$:= \mathbf{X}^{(b)}, \tag{22}$$

which is known as diagonal unitary code (DUC). Here, the code index is given as $b = (\mathbf{b})_{10}$, where $(\cdot)_{10}$ denotes the binary to decimal conversion. Additionally, M diversity-maximizing factors $0 < u_1 \leq \cdots \leq u_M \leq 2^B/2 \in \mathbb{Z}$ are designed to maximize the diversity product of [30]

$$\min_{b \in \{1,\ldots,2^B-1\}} \left| \prod_{m=1}^{M} \sin\left(\frac{\pi b u_m}{2^B}\right) \right|^{\frac{1}{M}}. \tag{23}$$

Although this optimization is a time-consuming task, the designed factors are available from the open-source library used in [57].[7] For example, in the $(M, B) = (4, 4)$ case, the designed factors are $[u_1, u_2, u_3, u_4] = [1, 3, 5, 7]$. Note that (21) can be replaced with all of the sophisticated differential family which relies on sparse unitary matrices [37]–[39], [58], [59]. However, to simplify our analysis, we limit $\mathbf{X}(i)$ of (21) to the classic DUC.

7. https://github.com/ishikawalab/wiphy/blob/master/wiphy/code/duc.py

The time-varying basis $\mathbf{E}_1(i) \in \mathbb{C}^{M\times 1}$ varies as the transmission index increases from $i = 1$ to $i = W$, where W is the frame length. Details of the construction method of $\mathbf{E}_1(i)$ are presented in Section IV-C. This basis $\mathbf{E}_1(i)$ is the first column of $\mathbf{E}(i) \in \mathbb{C}^{M\times M}$. Later, other columns are represented by $\mathbf{E}(i) = [\mathbf{E}_1(i)\ \mathbf{E}_2(i)\ \cdots\ \mathbf{E}_M(i)] \in \mathbb{C}^{M\times M}$. For $1 \leq i \leq M$ blocks, the baseband symbol of (1) is defined as

$$\mathbf{S}(i) = \mathbf{E}_i(M) \in \mathbb{C}^{M\times 1}. \tag{24}$$

This equation implies that the unitary matrix $\mathbf{E}(M) \in \mathbb{C}^{M\times M}$ is transmitted in the first M time slots. Although this matrix $\mathbf{E}(M) \in \mathbb{C}^{M\times M}$ is equivalent to the conventional reference symbol, the overall performance will not change when increasing W [34], which is similar to the classic differential MIMO family. For $M + 1 \leq i \leq W$ blocks, the baseband symbol is defined as

$$\mathbf{S}(i) = \tilde{\mathbf{S}}(i)\mathbf{E}_1(i) \in \mathbb{C}^{M\times 1}, \tag{25}$$

where we have the $M \times M$ matrix of

$$\tilde{\mathbf{S}}(i) = \begin{cases} \mathbf{I}_M & (i \leq M) \\ \tilde{\mathbf{S}}(i-1)\mathbf{X}(i) & (i > M). \end{cases} \tag{26}$$

The effective transmission rate is calculated as $R^{\mathrm{eff}} = (W - M)/W \cdot R = (1 - M/W) \cdot R$, whereas the ideal transmission rate is $R = B$ [bit/symbol]. In this paper, we use $W = 20 \cdot M$ to keep the rate loss at 5%. The frame lengths of $W = 100 \cdot M$ and $1000 \cdot M$ are also possible. However, these simulations might become time-consuming.

B. DECODING AT THE RECEIVER

For $1 \leq i \leq M$ blocks, the estimate of $\mathbf{H}(i)\tilde{\mathbf{S}}(i) \in \mathbb{C}^{N\times M}$ is updated as

$$\hat{\mathbf{Y}}(i) = \hat{\mathbf{Y}}(i-1) + \mathbf{Y}(i)\mathbf{E}_i^{\mathrm{H}}(M) \in \mathbb{C}^{N\times M}, \tag{27}$$

where the initial value is a zero matrix, i.e., $\hat{\mathbf{Y}}(0) = \mathbf{0}_{N\times M}$. Then, for $i > M$ blocks, the data-carrying matrix $\mathbf{X}(i)$ of (21), which is associated with the input bits $\mathbf{b}$, is estimated by the maximum-likelihood detector of[8]

$$\hat{\mathbf{X}}(i) = \arg\min_{\mathbf{X}} \left\| \mathbf{Y}(i) - \hat{\mathbf{Y}}(i-1)\mathbf{X}\mathbf{E}_1(i) \right\|_{\mathrm{F}}^2, \tag{28}$$

where we have

$$\hat{\mathbf{Y}}(i) = \hat{\mathbf{Y}}(i-1)\hat{\mathbf{X}}(i) + 1 - \alpha(i)\mathbf{D}(i)\mathbf{E}_1^{\mathrm{H}}(i) \in \mathbb{C}^{N\times M}, \tag{29}$$

$$\alpha(i) = \min(N \cdot \sigma_v^2/\|\mathbf{D}(i)\|_{\mathrm{F}}^2, 0.99) \in \mathbb{R}, \tag{30}$$

and

$$\mathbf{D}(i) = \mathbf{Y}(i) - \hat{\mathbf{Y}}(i-1)\hat{\mathbf{X}}(i)\mathbf{E}_1(i) \in \mathbb{C}^{N\times 1}. \tag{31}$$

The adaptive forgetting factor $\alpha(i)$ must be within the range of $(0, 1)$, which minimizes the error of $\left\|\hat{\mathbf{Y}}(i) - \mathbf{H}(i)\tilde{\mathbf{S}}(i)\right\|_{\mathrm{F}}^2$. As given, the estimate of CSI, $\mathbf{H}(i)$, is not included

8. If the coherent time is extremely short, the noncoherent detection will cause an error floor, which is similar to the coherent detection. Refer to [33] for a study in high-mobility scenarios and [34] for a study in millimeter-wave scenarios.

in (28). Instead of $\mathbf{H}(i)$, this detector is used to calculate $\hat{\mathbf{Y}}(i) \approx \mathbf{H}(i)\tilde{\mathbf{S}}(i)$, which yields a low-complexity detection. Specifically, the detection complexity of (28) is lower bounded by $\Omega(2^R NM^2 n \log n)$. This complexity is higher than those of the conventional chaos-based schemes having $T = 1$. However, these schemes must carry out complex channel estimation that is not considered for this study.

C. PROPOSED CHAOS BASIS

The proposed chaos basis inherits a basic property from the conventional basis proposed in [33]. Similarly to the conventional method, the following $M \times M$ static discrete Fourier transform (DFT) matrix is generated [33]:

$$\mathbf{W} = \frac{1}{\sqrt{M}} \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & \omega & \cdots & \omega^{M-1} \\ 1 & \omega^2 & \cdots & \omega^{2(M-1)} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \omega^{M-1} & \cdots & \omega^{(M-1)(M-1)} \end{bmatrix}, \quad (32)$$

where $\omega = \exp(-2\pi j/M)$. Later, the first column of $\mathbf{W}$ is denoted as

$$\mathbf{W}_1 = \underbrace{[1\ 1\ \cdots\ 1]^{\mathrm{T}}}_{M \text{ rows}} / \sqrt{M} \quad (33)$$

for simple notation. The conventional static DFT basis is generated by [33]

$$\mathbf{E}(i) = \mathbf{W}, \quad (34)$$

whereas the proposed time-varying unitary matrix is generated as

$$\begin{aligned} \mathbf{E}(i) &= \exp(j2\pi y_i Y)\mathbf{W} && (35)\\ &= \exp\left(j4 \arcsin \sqrt{x_i} Y\right)\mathbf{W} && (36)\\ &:= \exp(j\theta(x_i))\mathbf{W} && (37) \end{aligned}$$

where y_i is a chaos sequence defined in (7) and Y is a non-zero arbitrary integer. If $|Y| \geq 2$, (35) becomes a non-invertible function of y_i, which improves security. The tradeoff between security and reliability is discussed in Section VI. Finally, the time-varying DFT basis is generated as $\mathbf{E}_1(i) = \exp(2\pi y_i Y j)\mathbf{W}_1 \in \mathbb{C}^{M \times 1}$.

Fig. 3 shows the transitions of the chaos DFT basis (35) having $(M, T) = (4, 1)$ and $Y = 8$, where the time index was increased from $i = M/T + 1 = 5$ to $W = 80$. As shown in Fig. 3, the first row of $\mathbf{E}_1(i)$ was distributed uniformly on a circle.

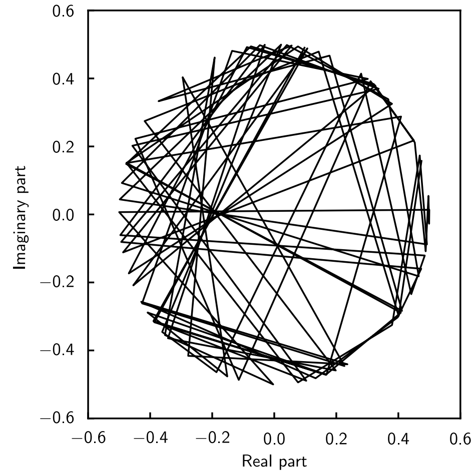


FIGURE 3. Transition of the chaos DFT basis $\mathbf{E}_1(i) \in \mathbb{C}^{4\times 1}$ (35), where the first row of $\mathbf{E}_1(i)$ was presented in the I/Q domain.

D. GENERATION OF A REAL-VALUED SECRET KEY FROM THE TRUE RANDOMNESS OF WIRELESS NATURE

In the literature, several key generation methods have been proposed. They rely on the true randomness of wireless nature [19], with characteristics such as the random received signal strength (RSS) [60], the random channel coefficients [27], the diversity of MIMO [28], the MIMO channel fluctuations [29], and distributed antennas [61]. Most of these methods are highly reliable. However, achieving the key agreement ratio of 100% in high-mobility scenarios, where the benefits of differential coding can be exploited, is a challenging task. As described in Section III, any chaos-based system requires at least one initial real-valued key, which is represented as a 64-bit floating-point number, for example. Here, optimal mapping between a shared 64-bit key and a real-valued key remains unknown. One bit error might result in a large difference in the real-valued counterpart. Therefore, we must consider a real-valued key generation method that differs from conventional binary key generation methods.

Because the proposed scheme is free from the channel estimation process, we opt to use the randomness of RSS to generate a real-valued key, which is inspired by the key generation method presented earlier in the literature [60]. The following RSS-based key generation method is used to model the error of shared keys, the effects of SNR and the channel correlation between Alice-Bob and Alice-Eve channel matrices.

RSS-based example: We specifically examine a single receive antenna. Its RSS value is mapped to a $[0, 1]$ real-valued key. To be more specific, the initial condition of (4) at the transmitter is modeled as

$$x_0^t = \exp\left(-|h + \epsilon_e \sigma_v v|^2\right), \quad (38)$$

where $h \sim \mathcal{CN}(0, 1)$ represents a single channel coefficient, $v \sim \mathcal{CN}(0, 1)$ stands for an additive noise, and $\epsilon_e \in \mathbb{R}$ denotes the accuracy of shared keys. Because the Rayleigh fading channel is assumed for this study, $|h|$ follows the Rayleigh distribution; also, x_0^t follows a uniform distribution $[0, 1]$ at high SNRs. Obviously, the phase information of h can be useful in the same manner as [62], which is not considered in this paper for simplicity. By contrast, at the receiver, the corresponding real-valued key is generated as

$$x_0^r = \exp\left(-\left|\rho h + \sqrt{1 - \rho^2} h' + \epsilon_e \sigma_v v'\right|^2\right), \quad (39)$$

where $h' \sim \mathcal{CN}(0,1)$ and $v' \sim \mathcal{CN}(0,1)$ respectively denote single channel and noise components. The channel correlation $\rho \in [0,1]$ also determines the accuracy. Bob invariably has $\rho = 1$, whereas Eve might have $\rho \in [0,1)$. The error between x_0^t and x_0^r improves as SNR $= 1/\sigma_v^2$ increases. In the MIMO context, the methods described in several papers rely on the assumption of PCSI at both the transmitter and the receiver, i.e., $\epsilon_e = 0$. This assumption is optimistic because the effects of the additive noise [63] and the mismatch of the TDD channel reciprocity cannot be ignored. As shown in Fig. 1, even a small error induces a mismatch in the chaos sequences at the transmitter and the receiver.

As might be inferred from (38) and (39), when Eve is near Bob, i.e., ρ is close to 1, Eve can estimate the secret key generated at Bob. Assuming this simple but vulnerable RSS-based method, one can discuss the minimum security level that can be guaranteed under the proposed system. Actually, the security level can be improved using artificial noise [27] and beamforming [64] with the sacrifice of additional complexity. The key generation example above can be replaced with such a sophisticated method.

E. SECRECY RATE OF THE PROPOSED SCHEME

In [65], Wang *et al.* defined the secrecy rate as

$$C_s = \max(0, I_B - I_E). \tag{40}$$

Here, I_B denotes the average mutual information (AMI) between Alice and Bob, while I_E denotes the AMI between Alice and Eve. Since the AMI for nonsquare differential coding is still unknown, we assume coherent detection at the receiver and calculate the AMI for the proposed time-varying codewords. The AMI I_B can be derived by extending the definition of [66] as follows:

$$I_{\mathrm{B}} = B - \frac{1}{2^B}\frac{1}{W-M}\sum_{i=M+1}^{W}\sum_{f=1}^{2^B} E_{\mathbf{H},\mathbf{V}}\left[\log_2 \sum_{g=1}^{2^B}\exp\left(\frac{\eta_{\mathrm{B}}[i,f,g]}{\sigma_v^2}\right)\right], \tag{41}$$

where we have

$$\eta_{\mathrm{B}}[i,f,g] = -\left\|\mathbf{H}\left(\mathbf{X}^{(f)} - \mathbf{X}^{(g)}\right)\exp\left(j\theta\left(x_i^t\right)\right)\mathbf{W}_1 + \mathbf{V}\right\|_{\mathrm{F}}^2 + \|\mathbf{V}\|_{\mathrm{F}}^2. \tag{42}$$

Similarly, the AMI I_E can be derived by

$$I_{\mathrm{E}} = B - \frac{1}{2^B}\frac{1}{W-M}\sum_{i=M+1}^{W}\sum_{f=1}^{2^B} E_{\mathbf{H},\mathbf{V}}\left[\log_2 \sum_{g=1}^{2^B}\exp\left(\frac{\eta_{\mathrm{E}}[i,f,g]}{\sigma_v^2}\right)\right], \tag{43}$$

where we have

$$\begin{aligned}\eta_{\mathrm{E}}[i,f,g] &= -\left\|\mathbf{H}\left(\mathbf{X}^{(f)}\exp\left(j\theta\left(x_i^t\right)\right) - \mathbf{X}^{(g)}\exp\left(j\theta\left(x_i^r\right)\right)\right)\mathbf{W}_1 + \mathbf{V}\right\|_{\mathrm{F}}^2 \\ &\quad + \left\|\mathbf{H}\left(\mathbf{X}^{(f)}\exp\left(j\theta\left(x_i^t\right)\right) - \mathbf{X}^{(f)}\exp\left(j\theta\left(x_i^r\right)\right)\right)\mathbf{W}_1 + \mathbf{V}\right\|_{\mathrm{F}}^2.\end{aligned} \tag{44}$$

F. LOW-COMPLEXITY CALIBRATION OF THE CHAOS SEQUENCE

The difference between x_0^t and x_0^r induces severe communication errors. For that reason, all the conventional chaos-based systems used a pre-shared key for initializing chaos sequences at the transmitter and receiver. To address this issue, we propose a novel calibration algorithm for the chaos sequence at the receiver. For each candidate $\mathbf{X}$ in (28), the corresponding $\mathbf{E}_1(i) = \exp(2\pi y_i^r Yj)\mathbf{W}_1$ in (28) is calibrated. Specifically, the detector tries to calibrate x_i^r and $y_i^r = 2\arcsin\sqrt{x_i^r}/\pi$, which might contain errors. The latter y_i^r can be calibrated by solving

$$\hat{y}_i^r = \arg\min_y \left\|\mathbf{Y}(i) - \exp(j2\pi yY)\hat{\mathbf{Y}}(i-1)\mathbf{X}\mathbf{W}_1\right\|_{\mathrm{F}}^2 \tag{45}$$

for a given $\mathbf{X}$. This optimization problem has multiple solutions because of the phase ambiguity induced by Y. Here, (45) is solvable by a low-complexity closed-form equation of

$$\hat{y}_i^r = \frac{\theta_y + \hat{n}\pi}{2\pi Y}, \tag{46}$$

where we have $\theta_y = \arctan(-\mathrm{Im}[\mathbf{a}]/\mathrm{Re}[\mathbf{a}])$, $\mathbf{a} = \mathrm{tr}[\mathbf{Y}(i)^{\mathrm{H}}\hat{\mathbf{Y}}(i-1)\mathbf{X}\mathbf{W}_1]$, and $\hat{n} = \lfloor 2y_i^r Y - \theta_y/\pi + 0.5\rfloor$. Finally, the receiver obtains the calibrated chaos sequences $\hat{y}_i^r$ and $\hat{x}_i^r = \sin^2 \hat{y}_i^r\pi/2$. After obtaining $\hat{y}_i^r$, the receiver updates $\mathbf{E}_1(i) = \exp(2\pi\hat{y}_i^r Yj)\mathbf{W}_1$ of (28).

The complexity of (45) is lower bounded by $\Omega(N_g N n \log n)$, where N_g is the search space size of y; it is set to a large value such as 10^3 or 10^4. By contrast, the complexity of (46) is negligible by virtue of its closed-form calculations. Therefore, the lower-bound for overall ML complexity with the calibration algorithm is the same as that of the conventional nonsquare differential decoding, as analyzed in Section IV-B.

In summary, the proposed detection process with the chaos calibration algorithm is outlined in Algorithm 1.

V. ATTACK ALGORITHM AND SECURITY ANALYSIS

We conceive an attack algorithm for the proposed scheme, for which we assume that Eve has PCSI [67] and infinite SNR, although Bob has no CSI and realistic SNR. In practice, it is a challenging task for Eve to obtain a precise estimate of CSI because the proposed scheme transmits no fixed reference symbol. If reference symbols are not available, then the receiver can exploit a sophisticated blind channel estimation method [68], [69]. However, this blind estimation is possible only if all the transmitted space-time

Algorithm 1 Proposed ML Detector With the Chaos Calibration Algorithm of Section IV-F

Input: $\mathbf{Y}(i), \hat{\mathbf{Y}}(i-1), \hat{x}^r_{i-1}, \mathbf{W}_1, B, Y, N, \sigma_v^2$
Output: $\hat{\mathbf{b}}(i), \hat{\mathbf{Y}}(i), \hat{x}^r_i$
Initialization:
$\tau_{\min} = +\infty$ $\{\tau_{\min}$ is the minimum of (26)$\}$
$x^r_i = 4\hat{x}^r_{i-1}\left(1 - \hat{x}^r_{i-1}\right)$ {update x^r_i using (2)}
$y^r_i = 2\arcsin(\sqrt{x^r_i})/\pi$ {update y^r_i using (5)}
ML detection for $\hat{\mathbf{X}}(i) = \mathbf{X}^{(b_{\min})}$:
for $b = 0$ to 2^{B-1} **do**
  $\mathbf{a} = \mathrm{tr}\left[\mathbf{Y}(i)^{\mathrm{H}}\hat{\mathbf{Y}}(i-1)\mathbf{X}^{(b)}\mathbf{W}_1\right]$ {calibrate y^r_i}
  $\theta_y = \arctan(-\mathrm{Im}[\mathbf{a}]/\mathrm{Re}[\mathbf{a}])$
  $\hat{n} = \lfloor 2y^r_i Y - \theta_y/\pi + 0.5\rfloor$
  $\hat{y}^r_i = (\theta_y + \hat{n}\pi)/(2\pi Y)$ {obtain the calibrated $\hat{y}^r_i$}
  $\mathbf{E}_1(i) = \exp(2\pi\hat{y}^r_i Yj)\mathbf{W}_1$ {update $\mathbf{E}_1(i)$}
  $\mathbf{D} = \mathbf{Y}(i) - \hat{\mathbf{Y}}(i-1)\mathbf{X}^{(b)}\mathbf{E}_1(i)$
  $\tau = \|\mathbf{D}\|_{\mathrm{F}}^2$ {calculate ML detection norm using (26)}
  if $(\tau < \tau_{\min})$ **then**
    $\tau_{\min} = \tau$, $b_{\min} = b$, and $\mathbf{D}_{\min} = \mathbf{D}$
    $\hat{x}^r_i = \sin^2\left(\hat{y}^r_i\pi/2\right)$
    $\hat{\mathbf{E}}_1(i) = \mathbf{E}_1(i)$
  end if
end for
Finalization:
$\hat{\mathbf{b}}(i) = (b_{\min})_2$ {obtain the estimated bits}
$\alpha(i) = \min\left(N\cdot\sigma_v^2/\tau_{\min}, 0.99\right)$ {update $\hat{\mathbf{Y}}(i)$}
$\hat{\mathbf{Y}}(i) = \hat{\mathbf{Y}}(i-1)\mathbf{X}^{(b_{\min})} + (1-\alpha(i))\mathbf{D}_{\min}\hat{\mathbf{E}}_1^{\mathrm{H}}(i)$
return $\hat{\mathbf{b}}(i), \hat{\mathbf{Y}}(i), \hat{x}^r_i$

matrices are semi-unitary [68]. As described in an earlier report [68], results showed that 50 unknown unitary matrices were necessary to obtain a precise CSI for a 4×4 MIMO scenario. Another blind estimation method [69] required 2000 OFDM symbols, each of which had 64 subcarriers. It is unrealistic to apply these blind estimation approaches to high-mobility scenarios, where differential schemes work efficiently.

A. ATTACK ALGORITHM

Eve has PCSI $\mathbf{H}_{\mathrm{E}} \in \mathbb{C}^{N\times M}$, which is unrealistic, as described above. This fact enables coherent detection at Eve, although the transmitted matrix $\tilde{\mathbf{S}}(i)$ of (26) is differentially encoded. Later, the first M received symbols are denoted by $\bar{\mathbf{Y}}_{\mathrm{E}} = [\mathbf{Y}_{\mathrm{E}}(1)\ \mathbf{Y}_{\mathrm{E}}(2)\ \cdots\ \mathbf{Y}_{\mathrm{E}}(M)]$, which converges to $\mathbf{H}_{\mathrm{E}}\mathbf{E}(M)$ when SNR $\rightarrow +\infty$. The secret key x_0 can be estimated by solving

$$\hat{x}_0 = \arg\min_{x} g_1(x), \tag{47}$$

where we have

$$g_1(x) = \left\|\bar{\mathbf{Y}}_{\mathrm{E}} - \exp j\theta(x)\mathbf{H}_{\mathrm{E}}\mathbf{W}\right\|_{\mathrm{F}}^2 \tag{48}$$

and $\theta(x)$ defined in (37). However, this optimization problem returns multiple solutions because we have $|Y| > 1$. The phase ambiguity imposed by $|Y| > 1$ enhances security.

To address this challenge, Eve must perform high-complexity joint optimization over $i = 1, \ldots, W$ blocks. Letting $\mathbf{d} = [d_1, d_2, \ldots, d_D]$ be a set of integers indicating data-carrying matrices and letting $D = W - M$ be the number of these matrices, then as defined in (22), each integer in $\mathbf{d}$ ranges from 0 to 2^{B-1}. Accordingly, the number of possible patterns of $\mathbf{d}$ is calculable as 2^{BD}. Eve estimates $\hat{x}_0$ and $\hat{\mathbf{d}}$ simultaneously by solving

$$\hat{x}_0, \hat{\mathbf{d}} = \arg\min_{(x,\mathbf{d})} g_1(x) + \sum_{i'=1}^{D} g_2\left(i', x, \mathbf{d}\right), \tag{49}$$

where we have

$$g_2\left(i', x, \mathbf{d}\right) = \left\|\mathbf{Y}_{\mathrm{E}}M + i' - \exp j\theta f\left(i', x\right)\mathbf{H}_{\mathrm{E}}\mathbf{F}\left(i', \mathbf{d}\right)\right\|_{\mathrm{F}}^2 \tag{50}$$

and

$$\mathbf{F}\left(i', \mathbf{d}\right) = \underbrace{\mathbf{X}^{(d_1)}\mathbf{X}^{(d_2)}\cdots\mathbf{X}^{(d_{i'})}}_{i' \text{ matrices}}\mathbf{W}_1. \tag{51}$$

Note that $f(\cdot,\cdot)$ is defined in (6), $\mathbf{X}^{(\cdot)}$ is defined in (22), and $\mathbf{W}_1$ is defined in (33).

B. SECURITY ANALYSIS

The complexity of (49) is extremely high because of the global optimization imposed by $\hat{x}_0 \in [0, 1]$ and the large search space of $\hat{\mathbf{d}} \in \mathbb{Z}^D$.

1) Global optimization for continuous $\hat{x}_0$: Since the objective function of (49) is non-convex, Eve has to perform a global optimization for $\hat{x}_0$, such as the brute-force search with a step size of 10^{-64}, the dual annealing method [70] with a large number of iterations, and the differential evolution method [71]. The first derivative of (49) is given as

$$\begin{aligned}\frac{d}{dx}&\left[g_1(x) + \sum_{i'=1}^{D} g_2\left(i', x, \mathbf{d}\right)\right] \\ &= -2\cdot\frac{d}{dx}\mathrm{Re}\Big[\exp(j\theta(x))\mathrm{tr}\left(\bar{\mathbf{Y}}_{\mathrm{E}}^{\mathrm{H}}\mathbf{H}_{\mathrm{E}}\mathbf{W}\right) \\ &\qquad + \sum_{i'=1}^{D}\exp\left(j\theta\left(f\left(i', x\right)\right)\right) \\ &\qquad \times \mathrm{tr}\left(\mathbf{Y}_{\mathrm{E}}^{\mathrm{H}}(M + i')\mathbf{H}_{\mathrm{E}}\mathbf{F}\left(i', \mathbf{d}\right)\right)\Big],\end{aligned} \tag{52}$$

which contains a degree $D + 1$ polynomial function, and cannot be solved algebraically. For example, if we consider the $D = W - M = 80 - 4 = 76$ case, the first derivative $d/dx(\exp(j\theta(f(76, x))))$, where $f(76, x)$ is a degree 77 polynomial function, yields a lot of solutions, which makes this global optimization difficult.

2) Brute-force search for discrete $\mathbf{d}$: To solve the optimization problem of (49) and to resolve the phase ambiguity, Eve must perform a brute-force combinatorial search for estimation of $\mathbf{d}$. The search space for $\mathbf{d}$ is calculable as $2^{B(W-M)} = 2^{BD}$, which increases exponentially with the transmission rate $R = B$ and the number of transmit antennas M. For example, in the

$(M, R, D) = (4, 1, 2)$ case, we have $2^2 = 4$ patterns: $\mathbf{d} = [d_1, d_2] = [0, 0], [0, 1], [1, 0], [0, 1]$. Each set determines the combination of differentially encoded symbols as $\mathbf{X}^{(0)}\mathbf{X}^{(0)}, \mathbf{X}^{(0)}\mathbf{X}^{(1)}, \mathbf{X}^{(1)}\mathbf{X}^{(0)}, \mathbf{X}^{(1)}\mathbf{X}^{(1)}$. To maximize the effective transmission rate, $R^{\mathrm{eff}} = (1 - M/W) \cdot R = (1 - M/W) \cdot R$, the frame length is set to a large value such as $W = 20M$, $100M$, or $1000M$. In each case, the search space becomes $2^{BD} = 2^{R(W-M)} = 2^{19RM}, 2^{99RM}$, and 2^{999RM}. Here, Eve must prepare 2^{304}, 2^{1584}, and 2^{15984} candidates. Because the National Security Agency in the USA recommended the key length of 256 bits as the advanced encryption standard,[9] the search space of the proposed scheme is sufficiently large.

In summary, the overall complexity of (49) is lower bounded by $\Omega(2^{RD} N_g D^2 M^2 N n \log n)$, where N_g denotes the maximum iteration limit of the global optimization method. Although we considered the simplest and the worst real-valued key generation method in Section IV-D, the minimum security level achieved by the proposed system is sufficiently high.

VI. PERFORMANCE COMPARISONS

We investigated the performance of the proposed scheme in terms of the secrecy rate and bit error ratio (BER). Specifically, the proposed scheme having the time-varying chaos basis of (35) was considered, where the novel detector of (28) and the chaos calibration algorithm of (46) were used. Additionally, we considered two conventional chaos-based MIMO schemes: MIMO-CSK [14] described in Section III-A and C-MIMO [10]–[13] described in Section III-C. Here, PCSI at the legitimate receiver was assumed to benefit these conventional schemes. We also considered the classic differential star-QAM (SQAM) [72] and the conventional nonsquare DUC (N-DUC) [33], both of which worked efficiently without CSI. As a reference, the massive MIMO (M-MIMO) cryptography method of [73] was considered, although it required PCSI at both the transmitter and receiver. It is noteworthy that the M-MIMO cryptography is designed particularly for large-scale scenarios, but is also beneficial for small-scale scenarios by virtue of its linear precoding [73].

For our simulations, we assumed the ideal Rayleigh fading channel model as described in Section II. The numbers of transmit and receive antennas were, respectively, $M = 4$ and $N = 4$. The transmission rate was $R = 4$ [bit/symbol]. The frame length was $W = 20 \cdot M = 80$. The conventional MIMO-CSK scheme used (12) with two 256-QAM symbols, whereas the conventional C-MIMO scheme used (20) with $(M, T, N_s) = (4, 1, 100)$ and $(4, 2, 100)$. The M-MIMO cryptography method used 16-QAM symbols. The proposed scheme used (35) with $Y = 8$ to generate a time-varying chaos basis. The data-carrying unitary matrix of (21) was generated by $[u_1, u_2, u_3, u_4] = [1, 1, 1, 1]$ for $R = B = 2$, $[1, 3, 5, 7]$ for $R = 4$, $[1, 21, 24, 25]$ for $R = 6$, and $[1, 35, 41, 119]$ for $R = 8$.

9. https://apps.nsa.gov/iaarchive/programs/iad-initiatives/cnsa-suite.cfm

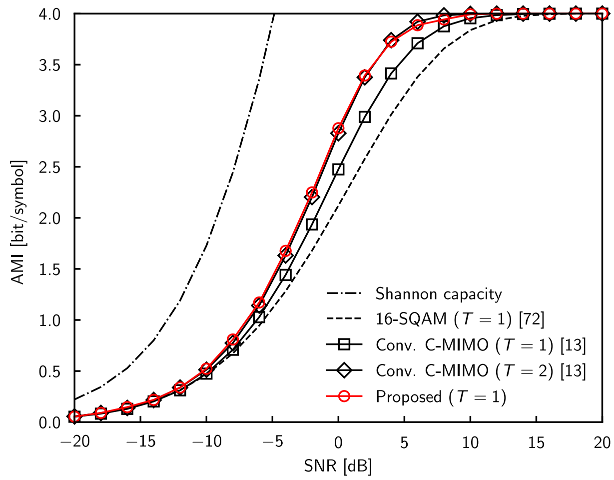


FIGURE 4. AMI comparison for which we considered the frame length of $W = M + 1 = 5$ and the transmission rate of $R = B = 4$ [bit/symbol].

First, we investigated the performance of the proposed scheme in channel-coded scenarios. Fig. 4 portrays an AMI comparison where we considered our proposed scheme and four other reference curves: the Shannon capacity, 16–SQAM [72], and the C-MIMO scheme [10] having $T = 1$ and 2. Here, we calculated AMI in the same manner as [44]. Because the constrained AMI of the nonsquare differential coding is not yet known, we assumed PCSI at the legitimate receiver and used the frame length of $W = M + 1 = 5$, which implies that the effects of differential encoding and decoding were not considered. Additionally, we assumed that Alice's and Bob's initial keys (x_0^t, x_0^r) were identical. As shown in Fig. 4, our proposed scheme having $T = 1$ outperformed the C-MIMO scheme having $T = 1$ and achieved the same AMI as the C-MIMO scheme having $T = 2$ that required additional decoding complexity. We observed the same trend in the $(M, R) = (8, 8)$ case. In fact, the C-MIMO has the advantage of Gaussian-distributed symbols, which are difficult for Eve to perceive, whereas the proposed scheme generates constant-envelope symbols.

Following Fig. 4, we investigated the achievable secrecy rate of the proposed scheme in Fig. 5, where the transmission rate was increased from $R = 2$ to 8 [bit/symbol] and where other simulation parameters were identical to those used in Fig. 4. Here, we calculated the secrecy rate defined by an earlier study [65], which was introduced in Section IV-E. We assumed that Bob's and Eve's channel matrices were independent, and assumed that both SNRs were identical. As shown in Fig. 5, the secrecy rate improved upon increasing the transmission rate from $R = 2$ to 8 monotonically. The corresponding ratios of the information leaked to Eve were, respectively, 19.83%, 20.39%, 12.37%, and 9.61% of the transmitted bits. This observation suggests that Eve is unable to decode all the private information correctly when we use the powerful channel coding technique with the coding rates of 1/2, 2/3, 3/4, or 5/6, for example.

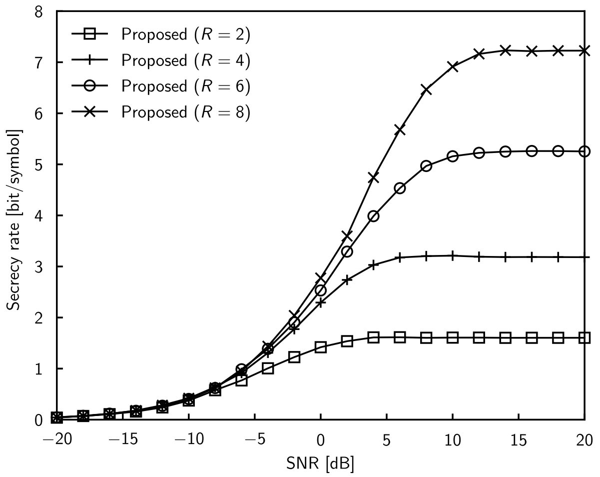


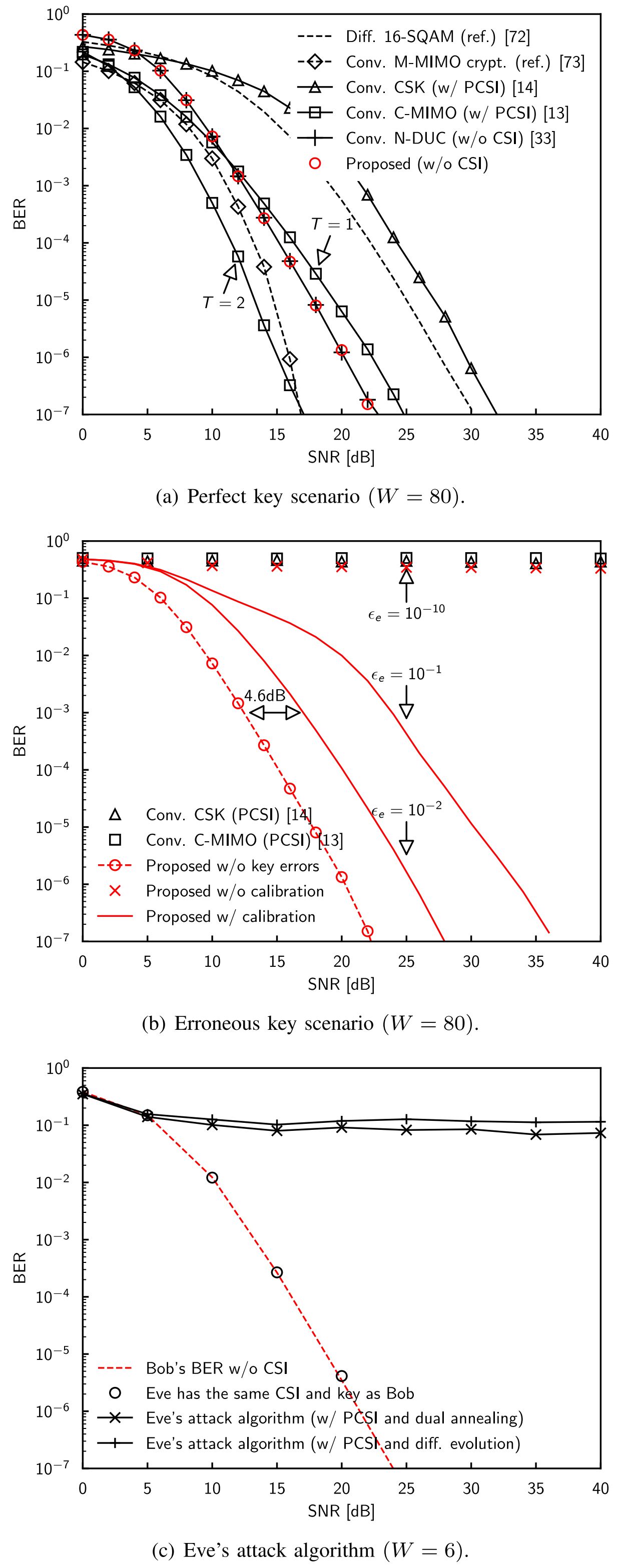
FIGURE 5. **Secrecy rate comparison upon increasing the transmission rate $R = B$ [bit/symbol], where other parameters are the same as those used in Fig. 4.**

Second, as shown in the panels of Fig. 6, we investigated the BER performance of the proposed scheme. Here, we considered (a) perfect key and (b) erroneous key scenarios. Additionally, we investigated the performance of Eve's attack algorithm as shown in Fig. 6(c).

For Fig. 6(a), an ideal condition was assumed: Alice and Bob had the same generated key $x_0^t = x_0^r$. Only the conventional schemes of MIMO-CSK [14], C-MIMO [10], and M-MIMO cryptography [73] had PCSI. Other schemes, including our proposal, had no CSI. As shown in Fig. 6(a), the proposed scheme with the time-varying chaos basis achieved the same performance as that of the conventional N-DUC scheme of [33]. This finding implies that the use of chaos basis induces no performance penalty. The conventional MIMO-CSK scheme having PCSI exhibited worse performance than the classic differential SQAM. Actually, this is true because MIMO-CSK relies on the diagonal matrix (13) composed of a chaos sequence. Since this diagonal matrix is not unitary, the minimum Euclidean distance of the resultant space-time matrix becomes a small value.[10] The conventional C-MIMO scheme having $T = 1$ and 2 achieved competitive performances. Specifically, the C-MIMO scheme having $T = 2$ achieved the best performance in the sacrifice of complexity, as analyzed in Section III-C, and outperformed M-MIMO cryptography, which required PCSI at the transmitter. In the $T = 1$ case, our proposed noncoherent scheme exhibited a similar trend to that of the coherent C-MIMO scheme having PCSI. This is particularly noteworthy because a noncoherent system generally exhibits the well-known 3 [dB] loss, unlike its coherent counterpart. Because the C-MIMO symbols follow a Gaussian distribution, which is a good property for improving security, the resultant BER might become a little worse despite having PCSI.

In Fig. 6(b), Alice and Bob obtained a secret key from the wireless channel, as described in Section IV-D. Both

10. Note that the original MIMO-CSK scheme was conceived for spread spectrum communications [14].



(a) Perfect key scenario ($W = 80$).

(b) Erroneous key scenario ($W = 80$).

(c) Eve's attack algorithm ($W = 6$).

FIGURE 6. **BER comparisons for which we considered three scenarios, where the transmission rate was $R = 4$ [bit/symbol]. (a) Perfect key scenario ($W = 80$). (b) Erroneous key scenario ($W = 80$). (c) Eve's attack algorithm ($W = 6$).**

extracted keys mutually differed. The difference was determined by the model of (39), where the error metric of $\epsilon_e = 10^{-1}, 10^{-2}$, and 10^{-10} were considered. The proposed scheme remains free from the estimation of a full channel matrix $\mathbf{H}(i) \in \mathbb{C}^{N\times M}$, but it requires an RSS value to generate a secret key x_0^t and x_0^r, as described in Section IV-D. As shown in Fig. 6(b), the proposed scheme without the calibration algorithm exhibited an error floor where the key contained the small error of $\epsilon_e = 10^{-10}$, which were similar to other conventional schemes. By contrast, the proposed scheme with the calibration algorithm achieved practical BER performance, even though we considered the high errors of $\epsilon_e = 10^{-1}$ and 10^{-2}. The SNR gap separating the perfect key scenario was 4.6 [dB] at BER $= 10^{-3}$.

In Fig. 6(c), we investigated the performance of Eve's attack algorithm described in Section V. In Figs. 6(a) and (b), we considered a realistic frame length $W = 20 \cdot M = 80$. By contrast, in Fig. 6(c), we changed the frame length $W = 80$ to 6 to enable the brute-force search at Eve. The corresponding search space was reduced from $2^{4(80-4)} = 2^{304}$ to $2^{4(6-4)} = 2^8$. In this unrealistic setup, the effective transmission rate was reduced from 3.80 to 1.33 [bit/symbol]. Since the brute-force search with a step size of 10^{-64} is infeasible, we used the dual annealing method [70] and the differential evolution method [71].[11] As shown in Fig. 6(c), Eve was able retrieve the same information as Bob when she had the same CSI as him. When Eve had no knowledge of CSI between Alice and Bob, and used the attack algorithm with the PCSI between Alice and Eve, she also succeeded in decoding 92.67% of information. Here, the performance of the dual annealing outperformed that of the differential evolution method. In summary, the proposed scheme offers limited security when Eve has the same CSI as Bob or has PCSI between her and Alice. The proposed scheme does not transmit fixed reference symbols and semi-unitary space-time matrices, as described in Section V. Consequently, it is difficult for Eve to obtain accurate CSI, especially when we consider high-mobility scenarios.

In Fig. 6(c), we observed a high error floor of BER $= 7.33 \cdot 10^{-2}$ when Eve had PCSI and SNR $\rightarrow +\infty$. To elucidate characteristics of this error floor, in Fig. 7, we calculated Eve's detection norm (49) at SNR $= 100$ [dB], where all the $2^8 = 256$ patterns for $\mathbf{d}$ were considered. Ideally, the detection norm (49) converges to zero at a high SNR. However, as shown in Fig. 7, we observed many local optimal solutions, which revealed that the optimization of (49) was not straightforward. This ambiguity resulted from the high complexity of (49) and its numerical errors. Specifically, the $|Y| > 1$ setup yields multiple solutions and induces the phase ambiguity for Eve. As a result, Eve selected the global optimal solution and obtained incorrect bits, whereas Bob selected a sub-optimal solution and obtained the correct bits. This promising result can be expected in general.

Finally, in Fig. 8, we investigated the effects of Eve's channel correlation and the security gap, which is defined

11. Specifically, we used the corresponding functions scipy.optimize.dual_annealing and scipy.optimize.differential_evolution with the maximum iteration limit of $N_g = 100$.

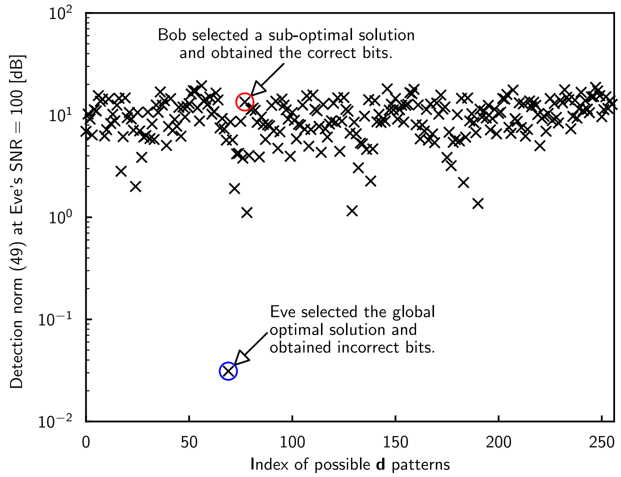


FIGURE 7. **Detection norms of Eve's attack algorithm** (49) **for all the possible d patterns, where the parameters were the same as those used in Fig. 6(c), the size of search space for d was $2^8 = 256$, and the dual annealing method [70] was used.**

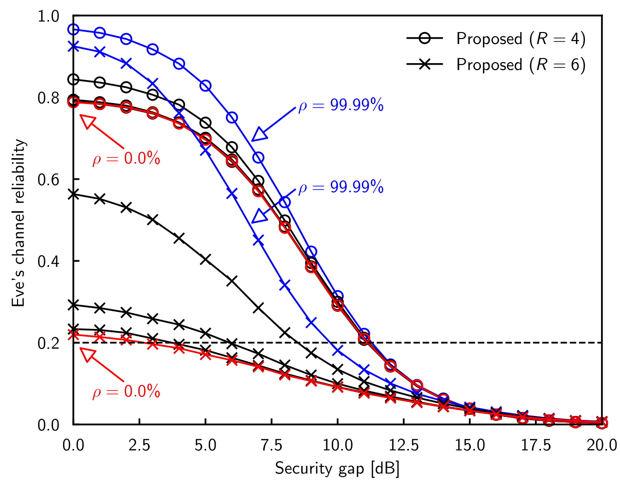


FIGURE 8. **Eve's channel reliability upon increasing the channel correlation ρ and the security gap. The simulation parameters were the same as those used in Fig. 6(a).**

by $\mathrm{SNR_{Bob}} - \mathrm{SNR_{Eve}}$ in dB [74], [75]. The simulation parameters were fundamentally the same as those used for Fig. 6(a), except for the correlation ρ between the Alice–Bob and Alice–Eve channel matrices. Specifically, Bob's channel model is defined in (1), whereas Eve's channel model is defined in (2). Here, the correlation coefficient was $\rho = 99.99\%, 99.9\%, 99.0\%, 90.0\%$, and 0.0%. Because the BER curves were difficult to differentiate, we showed the channel reliability instead. The channel reliability is defined as $1-2\cdot$BER. It is useful to calculate the channel capacity as demonstrated in [76]. In the same manner as [75], we define the BER ≥ 0.4 region as safe. The channel reliability must be less than 0.2. As presented in Fig. 8, the proposed scheme having $R = 4$ exhibited high channel reliability at Eve and required the security gap of about 11.1 [dB] to reach the safe region. By contrast, the proposed scheme having $R = 6$ required the security gap of about 3.0 [dB] in the $\rho \leq 90.0\%$

case. Its performance is comparable to those of conventional PLS methods [74], [75].

VII. CONCLUSION

In this paper, we proposed the chaos-based differential MIMO to alleviate the channel estimation overhead that would help the eavesdropper obtain an exact full channel matrix. Uniquely, the proposed scheme extracts an initial condition of the pure chaos sequence from wireless nature, which is the first attempt in the literature. Due to the sensitivity to initial conditions, conventional chaos-based communication systems must exchange a common secret key in advance with no exception. In our work, the extracted noisy key is used to generate an artificially time-varying unitary matrix, which obfuscates private data symbols. The keys extracted respectively for each transmitter and receiver might become different. To address this mismatch issue, we then proposed the low-complexity calibration algorithm for the chaos sequence at the receiver. Additionally, we conceived the brute-force attack algorithm for the proposed scheme. Our security analysis revealed that, because of its differential encoding structure, this attack algorithm was much more complex than the existing standard encryption method. Despite the fact that the proposed scheme requires no channel estimation, it outperformed the representative chaos-based scheme that had perfect channel estimates. It was found that the proposed calibration algorithm worked properly even if the extracted key contained non-negligible errors. However, our proposed system was unable to provide security if the eavesdropper had similar channel coefficients to a legitimate receiver, i.e., $\rho \geq 99.99\%$, or if the frame length was extremely short, such as $W \leq M+2$. Based on our analysis, we conclude that differential encoding can achieve practical physical layer security in wireless communications.

ACKNOWLEDGMENT

The authors are indebted to the Editor and the anonymous reviewers for their invaluable suggestions and comments, which further improved this paper.

REFERENCES

[1] P. Barsocchi, G. Oligeri, and F. Potorti, "Measurement-based frame error model for simulating outdoor Wi-Fi networks," *IEEE Trans. Wireless Commun.*, vol. 8, no. 3, pp. 1154–1158, Mar. 2009.

[2] F. Armknecht, J. Girao, A. Matos, and R. L. Aguiar, "Who said that? Privacy at link layer," in *Proc. IEEE INFOCOM*, Barcelona, Spain, May 2007, pp. 2521–2525.

[3] J. Noh, J. Kim, and S. Cho, "Secure authentication and four-way handshake scheme for protected individual communication in public Wi-Fi networks," *IEEE Access*, vol. 6, pp. 16539–16548, 2018.

[4] N. Ishikawa, Y. Ohishi, and K. Maeda, "Nulls in the air: Passive and low-complexity QoS estimation method for a large-scale Wi-Fi network based on null function data frames," *IEEE Access*, vol. 7, pp. 28581–28591, 2019.

[5] R. L. Rivest, A. Shamir, and L. Adleman, "A method for obtaining digital signatures and public-key cryptosystems," *Commun. ACM*, vol. 21, no. 2, pp. 120–126, 1978.

[6] P. W. Shor, "Algorithms for quantum computation: Discrete logarithms and factoring," in *Proc. Annu.Symp. Found. Comput. Sci.*, Santa Fe, NM, USA, Nov. 1994, pp. 124–134.

[7] P. Botsinis *et al.*, "Quantum search algorithms for wireless communications," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 2, pp. 1209–1242, 2nd Quart., 2019.

[8] J. M. Hamamreh, H. M. Furqan, and H. Arslan, "Classifications and applications of physical layer security techniques for confidentiality: A comprehensive survey," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 2, pp. 1773–1828, 2nd Quart., 2019.

[9] H. Dedieu, M. P. Kennedy, and M. Hasler, "Chaos shift keying: Modulation and demodulation of a chaotic carrier using self-synchronizing Chua's circuits," *IEEE Trans. Circuits Syst. II, Analog Digit. Signal Process.*, vol. 40, no. 10, pp. 634–642, Oct. 1993.

[10] E. Okamoto, "A chaos MIMO transmission scheme for channel coding and physical-layer security," *IEICE Trans. Commun.*, vol. E95-B, no. 4, pp. 1384–1392, 2012.

[11] E. Okamoto and Y. Inaba, "A chaos MIMO transmission scheme using turbo principle for secure channel-coded transmission," *IEICE Trans. Commun.*, vol. E98B, no. 8, pp. 1482–1491, 2015.

[12] E. Okamoto and N. Horiike, "Application of MAP decoding for chaos MIMO scheme to improve error rate performance," *IEICE Commun. Exp.*, vol. 5, no. 10, pp. 365–370, 2016.

[13] E. Okamoto and N. Horiike, "Performance improvement of chaos MIMO scheme using advanced stochastic characteristics," *IEICE Commun. Exp.*, vol. 1, no. 10, pp. 371–377, 2016.

[14] G. Kaddoum, M. Vu, and F. Gagnon, "On the performance of chaos shift keying in MIMO communications systems," in *Proc. IEEE Wireless Commun. Netw. Conf.*, Mar. 2011, pp. 1432–1437.

[15] J. Zhang, A. Marshall, R. Woods, and T. Q. Duong, "Design of an OFDM physical layer encryption scheme," *IEEE Trans. Veh. Technol.*, vol. 66, no. 3, pp. 2114–2127, Mar. 2017.

[16] J. M. Hamamreh, E. Basar, and H. Arslan, "OFDM-subcarrier index selection for enhancing security and reliability of 5G URLLC services," *IEEE Access*, vol. 5, pp. 25863–25875, 2017.

[17] W. Trappe, "The challenges facing physical layer security," *IEEE Commun. Mag.*, vol. 53, no. 6, pp. 16–20, Jun. 2015.

[18] A. Mukherjee, S. A. A. Fakoorian, J. Huang, and A. L. Swindlehurst, "Principles of physical layer security in multiuser wireless networks: A survey," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 3, pp. 1550–1573, 3rd Quart., 2014.

[19] J. Zhang, G. Li, A. Marshall, A. Hu, and L. Hanzo, "A new frontier for IoT security emerging from three decades of key generation relying on wireless channels," *IEEE Access*, vol. 8, pp. 138406–138446, 2020.

[20] U. Maurer and S. Wolf, "Secret-key agreement over unauthenticated public channels–Part I: Definitions and a completeness result," *IEEE Trans. Inf. Theory*, vol. 49, no. 4, pp. 822–831, Apr. 2003.

[21] U. Maurer and S. Wolf, "Secret-key agreement over unauthenticated public channels–Part II: The simulatability condition," *IEEE Trans. Inf. Theory*, vol. 49, no. 4, pp. 832–838, Apr. 2003.

[22] U. Maurer and S. Wolf, "Secret-key agreement over unauthenticated public channels–Part III: Privacy amplification," *IEEE Trans. Inf. Theory*, vol. 49, no. 4, pp. 839–851, Apr. 2003.

[23] P. Huang and X. Wang, "Fast secret key generation in static wireless networks: A virtual channel approach," in *Proc. IEEE INFOCOM*, Turin, Italy, Apr. 2013, pp. 2292–2300.

[24] K. Zeng, "Physical layer key generation in wireless networks: Challenges and opportunities," *IEEE Commun. Mag.*, vol. 53, no. 6, pp. 33–39, Jun. 2015.

[25] J. Tang, H. Wen, K. Zeng, R.-F. Liao, F. Pan, and L. Hu, "Light-weight physical layer enhanced security schemes for 5G wireless networks," *IEEE Netw.*, vol. 33, no. 5, pp. 126–133, Sep./Oct. 2019.

[26] Z. Rezki, M. Zorgui, B. Alomair, and M.-S. Alouini, "Secret key agreement: Fundamental limits and practical challenges," *IEEE Wireless Commun.*, vol. 24, no. 3, pp. 72–79, Jun. 2017.

[27] D. Chen, X. Mao, Z. Qin, Z. Qin, P. Yang, and Y. Liu, "SmokeGrenade: A key generation protocol with artificial interference in wireless networks," in *Proc. IEEE Int. Conf. Mobile Ad-Hoc Sens. Syst.*, Oct. 2013, pp. 200–208.

[28] K. Zeng, D. Wu, A. Chan, and P. Mohapatra, "Exploiting multiple-antenna diversity for shared secret key generation in wireless networks," in *Proc. IEEE INFOCOM*, Mar. 2010, pp. 1837–1845.

[29] J. W. Wallace and R. K. Sharma, "Automatic secret keys from reciprocal MIMO wireless channels: Measurement and analysis," *IEEE Trans. Inf. Forensics Security*, vol. 5, pp. 381–392, 2010.

[30] B. M. Hochwald and W. Sweldens, "Differential unitary space-time modulation," *IEEE Trans. Commun.*, vol. 48, no. 12, pp. 2041–2052, Dec. 2000.
[31] C. Xu *et al.*, "Sixty years of coherent versus non-coherent tradeoffs and the road from 5G to wireless futures," *IEEE Access*, vol. 7, pp. 178246–178299, 2019.
[32] N. Ishikawa and S. Sugiura, "Rectangular differential spatial modulation for open-loop noncoherent massive-MIMO downlink," *IEEE Trans. Wireless Commun.*, vol. 16, no. 3, pp. 1908–1920, Mar. 2017.
[33] N. Ishikawa, R. Rajashekar, C. Xu, S. Sugiura, and L. Hanzo, "Differential space-time coding dispensing with channel estimation approaches the performance of its coherent counterpart in the open-loop massive MIMO-OFDM downlink," *IEEE Trans. Commun.*, vol. 66, no. 12, pp. 6190–6204, Dec. 2018.
[34] N. Ishikawa *et al.*, "Differential-detection aided large-scale generalized spatial modulation is capable of operating in high-mobility millimeter-wave channels," *IEEE J. Sel. Topics Signal Process.*, vol. 13, no. 6, pp. 1360–1374, Oct. 2019.
[35] C. Wu, Y. Xiao, L. Xiao, P. Yang, X. Lei, and W. Xiang, "Space-time block coded rectangular differential spatial modulation: System design and performance analysis," *IEEE Trans. Commun.*, vol. 67, no. 9, pp. 6586–6597, Sep. 2019.
[36] L. Xiao *et al.*, "Differentially-encoded rectangular spatial modulation approaches the performance of its coherent counterpart," *IEEE Trans. Commun.*, vol. 68, no. 12, pp. 7593–7607, Dec. 2020.
[37] C. Xu, R. Rajashekar, N. Ishikawa, S. Sugiura, and L. Hanzo, "Single-RF index shift keying aided differential space-time block coding," *IEEE Trans. Signal Process.*, vol. 66, no. 3, pp. 773–788, Feb. 2018.
[38] C. Xu *et al.*, "Finite-cardinality single-RF differential space-time modulation for improving the diversity-throughput tradeoff," *IEEE Trans. Commun.*, vol. 67, no. 1, pp. 318–335, Jan. 2019.
[39] R. Rajashekar, C. Xu, N. Ishikawa, S. Sugiura, K. V. S. Hari, and L. Hanzo, "Algebraic differential spatial modulation is capable of approaching the performance of its coherent counterpart," *IEEE Trans. Commun.*, vol. 65, no. 10, pp. 4260–4273, Oct. 2017.
[40] G. Kolumban, B. Vizvki, W. Schwarz, and A. Abel, "Differential chaos shift keying: A robust coding for chaotic communication," in *Proc. Int. Workshop Nonlinear Dyn. Electron. Syst.*, Jun. 1996, pp. 1–6.
[41] G. Kaddoum, M. Vu, and F. Gagnon, "Performance analysis of differential chaotic shift keying communications in MIMO systems," in *Proc. IEEE Int. Symp. Circuits Syst.*, May 2011, pp. 1580–1583.
[42] B. Chen, L. Zhang, and H. Lu, "High security differential chaos-based modulation with channel scrambling for WDM-aided VLC system," *IEEE Photon. J.*, vol. 8, no. 5, pp. 1–13, Oct. 2016.
[43] W. Hu, L. Wang, and G. Kaddoum, "Design and performance analysis of differentially spatial modulated chaos shift keying modulation system," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 64, no. 11, pp. 1302–1306, Nov. 2017.
[44] N. Ishikawa, S. Sugiura, and L. Hanzo, "50 years of permutation, spatial and index modulation: From classic RF to visible light communications and data storage," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 3, pp. 1905–1938, 3rd Quart., 2018.
[45] D. E. Knuth, "Big Omicron and big Omega and big Theta," *ACM SIGACT News*, vol. 8, no. 2, pp. 18–24, 1976.
[46] E. Cavus and B. Daneshrad, "A very low-complexity space-time block decoder (STBD) ASIC for wireless systems," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 53, no. 1, pp. 60–69, Jan. 2006.
[47] R. P. Brent and P. Zimmermann, *Modern Computer Arithmetic*. Cambridge, U.K.: Cambridge Univ. Press, 2010.
[48] J. C. Sprott, *Chaos and Time-Series Analysis*. New York, NY, USA: Oxford Univ. Press, 2003.
[49] J. Von Neumann, *Various Techniques Used in Connection With Random Digits* (National Bureau of Standards Applied Mathematics Series), vol. 12. Washington, DC, USA: Nat. Bureau Stand., 1951, pp. 36–38.
[50] O. E. Rössler, "An equation for continuous chaos," *Phys. Lett. A*, vol. 57, no. 5, pp. 397–398, 1976.
[51] J. J. Collins, M. Fanciulli, R. G. Hohlfeld, D. C. Finch, G. V. H. Sandri, and E. S. Shtatland, "A random number generator based on the logit transform of the logistic variable," *Comput. Phys.*, vol. 6, no. 6, pp. 630–632, 1992.
[52] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 8, pp. 1451–1458, Oct. 1998.
[53] P. Wolniansky, G. Foschini, G. Golden, and R. Valenzuela, "V-BLAST: An architecture for realizing very high data rates over the rich-scattering wireless channel," in *Proc. Int. Symp. Signals Syst. Electron.*, Pisa, Italy, Oct. 1998, pp. 295–300.
[54] Y. Zhu and H. Jafarkhani, "Differential modulation based on quasi-orthogonal codes," *IEEE Trans. Wireless Commun.*, vol. 4, no. 6, pp. 3018–3030, Nov. 2005.
[55] S. Wang and X. Wang, "M-DCSK-based chaotic communications in MIMO multipath channels with no channel state information," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 57, no. 12, pp. 1001–1005, Dec. 2010.
[56] P. Chen, L. Wang, and F. C. M. Lau, "One analog STBC-DCSK transmission scheme not requiring channel state information," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 60, no. 4, pp. 1027–1037, Apr. 2013.
[57] N. Ishikawa, "IMToolkit: An open-source index modulation toolkit for reproducible research based on massively parallel algorithms," *IEEE Access*, vol. 7, pp. 93830–93846, 2019.
[58] N. Ishikawa and S. Sugiura, "Unified differential spatial modulation," *IEEE Wireless Commun. Lett.*, vol. 3, no. 4, pp. 337–340, Aug. 2014.
[59] R. Rajashekar, N. Ishikawa, S. Sugiura, K. V. S. Hari, and L. Hanzo, "Full-diversity dispersion matrices from algebraic field extensions for differential spatial modulation," *IEEE Trans. Veh. Technol.*, vol. 66, no. 1, pp. 385–394, Jan. 2017.
[60] T. Allen, J. Cheng, and N. Al-Dhahir, "Secure space-time block coding without transmitter CSI," *IEEE Wireless Commun. Lett.*, vol. 3, no. 6, pp. 573–576, Dec. 2014.
[61] Z. Ji *et al.*, "Wireless secret key generation for distributed antenna systems: A joint space-time-frequency perspective," *IEEE Internet Things J.*, early access, May 27, 2021, doi: 10.1109/JIOT.2021.3084361.
[62] S. Althunibat, V. Sucasas, and J. Rodriguez, "A physical-layer security scheme by phase-based adaptive modulation," *IEEE Trans. Veh. Technol.*, vol. 66, no. 11, pp. 9931–9942, Nov. 2017.
[63] J. B. Perazzone, P. L. Yu, B. M. Sadler, and R. S. Blum, "Artificial noise-aided MIMO physical layer authentication with imperfect CSI," *IEEE Trans. Inf. Forensics Security*, vol. 16, pp. 2173–2185, 2021.
[64] Z. Kong, S. Yang, D. Wang, and L. Hanzo, "Robust beamforming and jamming for enhancing the physical layer security of full duplex radios," *IEEE Trans. Inf. Forensics Security*, vol. 14, pp. 3151–3159, 2019.
[65] L. Wang, S. Bashar, Y. Wei, and R. Li, "Secrecy enhancement analysis against unknown eavesdropping in spatial modulation," *IEEE Commun. Lett.*, vol. 19, no. 8, pp. 1351–1354, Aug. 2015.
[66] S. X. Ng and L. Hanzo, "On the MIMO channel capacity of multidimensional signal sets," *IEEE Trans. Veh. Technol.*, vol. 55, no. 2, pp. 528–536, Mar. 2006.
[67] H. M. Wang, T. Zheng, and X. G. Xia, "Secure MISO wiretap channels with multiantenna passive eavesdropper: Artificial noise vs. artificial fast fading," *IEEE Trans. Wireless Commun.*, vol. 14, no. 1, pp. 94–106, Jan. 2015.
[68] S. Shahbazpanahi, A. B. Gershman, and J. H. Manton, "Closed-form blind MIMO channel estimation for orthogonal space-time block codes," *IEEE Trans. Signal Process.*, vol. 53, no. 12, pp. 4506–4517, Dec. 2005.
[69] C. Shin, R. W. Heath, and E. J. Powers, "Blind channel estimation for MIMO-OFDM systems," *IEEE Trans. Veh. Technol.*, vol. 56, no. 2, pp. 670–685, Mar. 2007.
[70] Y. Xiang, S. Gubian, B. Suomela, and J. Hoeng, "Generalized simulated annealing for global optimization: The GenSA package," *R J.*, vol. 5, no. 1, pp. 13–29, 2013.
[71] J. Lampinen, "A constraint handling approach for the differential evolution algorithm," in *Proc. Congr. Evol. Comput.*, 2002, pp. 1468–1473.
[72] W. Webb, L. Hanzo, and R. Steele, "Bandwidth efficient QAM schemes for Rayleigh fading channels," *IEE Proc.*, vol. 138, no. 3, pp. 169–175, 1991.
[73] T. R. Dean and A. J. Goldsmith, "Physical-layer cryptography through massive MIMO," *IEEE Trans. Inf. Theory*, vol. 63, no. 8, pp. 5419–5436, Aug. 2017.
[74] D. Klinc, J. Ha, S. W. McLaughlin, J. Barros, and B.-J. Kwak, "LDPC codes for the gaussian wiretap channel," *IEEE Trans. Inf. Forensics Security*, vol. 6, pp. 532–540, 2011.

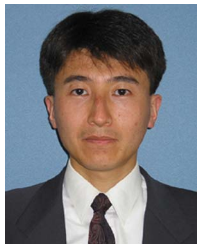
[75] M. Baldi, M. Bianchi, and F. Chiaraluce, "Coding with scrambling, concatenation, and HARQ for the AWGN wire-tap channel: A security gap analysis," *IEEE Trans. Inf. Forensics Security*, vol. 7, pp. 883–894, 2012.

[76] F. Yilmaz, "On the relationships between average channel capacity, average bit error rate, outage probability, and outage capacity over additive white Gaussian noise channels," *IEEE Trans. Commun.*, vol. 68, no. 5, pp. 2763–2776, May 2020.

NAOKI ISHIKAWA (Member, IEEE) was born in Kanagawa, Japan, in 1991. He received the B.E., M.E., and Ph.D. degrees from the Tokyo University of Agriculture and Technology, Tokyo, Japan, in 2014, 2015, and 2017, respectively. In 2015, he was an Academic Visitor with the School of Electronics and Computer Science, University of Southampton, U.K. From 2016 to 2017, he was a Research Fellow with the Japan Society for the Promotion of Science. From 2017 to 2020, he was an Assistant Professor with the Graduate School of Information Sciences, Hiroshima City University, Japan. Since 2020, he has been an Associate Professor with the Faculty of Engineering, Yokohama National University, Japan. His research interests include massive MIMO, physical layer security, and quantum speedup for wireless communications. He received the Yasujiro Niwa Outstanding Paper Award from Tokyo Denki University in 2018, the Telecom System Technology student Award (Honorable mention) from Telecommunications Advancement Foundation of Japan in 2014, and the Outstanding Paper Award for Young C&C Researchers from NEC C&C Foundation in 2014. He was certified as an Exemplary Reviewer of IEEE TRANSACTIONS ON COMMUNICATIONS 2017.

JEHAD M. HAMAMREH (Member, IEEE) received the Ph.D. degree in telecommunication engineering and cyber systems from Istanbul Medipol University, Turkey, in 2018.

He is the Founder and the Director of WISLAB, and a Professor with the Electrical and Electronics Engineering Department, Antalya Bilim University. He worked as a Researcher with the Department of Electrical and Computer Engineering, Texas A&M University. He is the Inventor of more than over 20 patents and authored more than over 75 peer-reviewed scientific papers along with several book chapters. His current research interests include wireless physical and MAC layers security, orthogonal frequency-division multiplexing and multiple-input multiple-output systems, advanced waveforms design, multidimensional modulation techniques, and orthogonal/non-orthogonal multiple access schemes for future wireless systems. His innovative patented works won the gold, silver, and bronze medals by numerous international invention contests and fairs. He is a serial referee for various scientific journals as well as a TPC member for several international conferences. He is an Editor at Researcherstore, *RS Open Journal on Innovative Communication Technologies*, and *Frontiers in Communications and Networks*.



EIJI OKAMOTO (Member, IEEE) received the B.E., M.S., and Ph.D. degrees in electrical engineering from Kyoto University in 1993, 1995, and 2003, respectively. In 1995, he joined the Communications Research Laboratory, Japan. He is currently an Associate Professor with the Nagoya Institute of Technology. In 2004, he was a Guest Researcher with Simon Fraser University. His current research interests are in the areas of wireless technologies, satellite communication, and mobile communication systems. He received the Young Researchers' Award in 1999 from IEICE, and the FUNAI Information Technology Award for Young Researchers in 2008.

CHAO XU (Senior Member, IEEE) received the M.Sc. degree (with Distinction) in radio frequency communication systems and the Ph.D. degree in wireless communications from the University of Southampton, U.K., in 2009 and 2015, respectively, where he is currently a Research Fellow. His research interests include index modulation, reduced-complexity MIMO design, noncoherent detection, and turbo detection. He was awarded the Best M.Sc. student in Broadband and Mobile Communication Networks by the IEEE Communications Society (U.K. and Republic of Ireland Chapter) in 2009. He also received 2012 Chinese Government Award for Outstanding Self-Financed Student Abroad and 2017 Dean's Award, Faculty of Physical Sciences and Engineering, University of Southampton.

LIXIA XIAO (Member, IEEE) received the B.E., M.E., and Ph.D. degrees from UESTC in 2010, 2013, and 2017, respectively. From 2016 to 2017, she was a Visiting Student with the School of Electronics and Computer Science, University of Southampton. From 2018 to 2019, she has been a Research Fellow with the Department of Electrical Electronic Engineering, University of Surrey. She is currently a Full Professor with the Wuhan National Laboratory for Optoelectronics, Huazhong University of Science and Technology. In particular, she is very interested in signal detection and performance analysis of wireless communication systems. Her research interests include wireless communications and communication theory.