IEEE Transactions on Information Forensics and Security, Volume 18, 2023, pp. 2368-2381
DOI:10.1109/TIFS.2023.3266705

Adversarial Reconfigurable Intelligent Surface Against Physical Layer Key Generation

Zhuangkun Wei, Bin Li, Weisi Guo

***Abstract*—The development of reconfigurable intelligent surfaces (RIS) has recently advanced the research of physical layer security (PLS). Beneficial impacts of RIS include but are not limited to offering a new degree-of-freedom (DoF) for key-less PLS optimization, and increasing channel randomness for physical layer secret key generation (PL-SKG). However, there is a lack of research studying how adversarial RIS can be used to attack and obtain legitimate secret keys generated by PL-SKG. In this work, we show an Eve-controlled adversarial RIS (Eve-RIS), by inserting into the legitimate channel a random and reciprocal channel, can partially reconstruct the secret keys from the legitimate PL-SKG process. To operationalize this concept, we design Eve-RIS schemes against two PL-SKG techniques used: (i) the CSI-based PL-SKG, and (ii) the two-way cross multiplication based PL-SKG. The channel probing at Eve-RIS is realized by compressed sensing designs with a small number of radio-frequency (RF) chains. Then, the optimal RIS phase is obtained by maximizing the Eve-RIS inserted deceiving channel. Our analysis and results show that even with a passive RIS, our proposed Eve-RIS can achieve a high key match rate with legitimate users, and is resistant to most of the current defensive approaches. This means the novel Eve-RIS provides a new eavesdropping threat on PL-SKG, which can spur new research areas to counter adversarial RIS attacks.**

***Index Terms*—Eavesdropping, Reconfigurable intelligent surface, Physical layer secret key, Wireless Communications.**

I. Introduction

Wireless communications are vulnerable to diverse attack vectors due to their broadcasting nature. Traditional cryptography techniques require high computational complexity and delays to ensure confidentiality, which makes them less attractive in real-time and lightweight systems [1]. To secure the wireless channels, a variety of physical layer security (PLS) techniques have been proposed and widely studied in the last decade.

A. Literature Review

PLS techniques can be categorized as key-less PLS and physical layer secret key generation (PL-SKG).

This work is supported by the Engineering and Physical Sciences Research Council [grant number: EP/V026763/1].
Zhuangkun Wei is with the School of Aerospace, Transport, and Manufacturing, Cranfield University, MK43 0AL, UK.
Weisi Guo is with the School of Aerospace, Transport, and Manufacturing, Cranfield University, MK43 0AL, UK, and also with the Alan Turing Institute, London, NW1 2DB, UK.
Bin Li is with the Department of Information Engineering, Beijing University of Posts and Telecommunications, Beijing, 100876, China.

1) Key-Less PLS: Key-less PLS tries to maintain the superiority of legitimate channels by maximizing the secrecy rate (via e.g., the beamforming vector [2], the trajectory of autonomous systems [3], the anti-jamming artificial noise [4], the spin modulation, etc.). The challenge lies in the high dependency on additional positioning data and the lack of guarantee of a feasible solution, especially when combined with real-world constraints.

2) PL-SKG: Another family is PL-SKG, which leverages the reciprocal channel randomness to generate shared secret keys [5]–[10]. Most of the PL-SKG schemes exploit the channel state information (CSI) as the common random feature, e.g., the received signal strength (RSS) [6], the channel phases [11], and the channel frequency response [12]. In these cases, two legitimate nodes (e.g., Alice and Bob) are required to send public pilot sequences to each other and pursue channel estimations to acquire these common CSI, which will then be passed to the quantization [13], [14], reconciliation [15] and privacy amplification [16] modules for key generation.

One challenge on PL-SKG is that the secret key rate cannot meet the industrial requirement due to insufficient channel randomness (e.g., RSS variations and small-scale channel scattering [17]), although optimization algorithms (e.g., power allocation [18]) can be used to improve the legitimate SKR. To address this, one-way based PL-SKG has been proposed by the works in [19]–[21], whereby one legitimate node (e.g., Alice) sends public pilots and Bob sends random signals. In this way, the common feature is Alice's received signals, which, at Bob's end, can be constructed by his channel estimation result and his sending random signals. As such, the feature randomness not only involves the random CSI but is enhanced by Bob's transmitted random signal, and thereby improving the SKR.

Inspired by the one-way randomness enhancement, the works in [22]–[25] further promote the SKR by leveraging the two-way random signals, whereby Alice and Bob send random pilots to each other and cross multiply their sent and received signals as the common feature (known as two-way cross multiplication method). In this view, the randomness of the common feature is further enhanced by two random spaces, and therefore leads to a higher SKR as opposed to one-way based and CSI-based PL-SKGs. Despite these advances, the improved SKR schemes are still not enough to approach the current Gbps levels of the transmission rate (i.e., making one information bit have one unique secret key for encryption), which renders as the main challenge to impede PL-SKG from civilian and commercial use.

3) When PLS meets RIS: Reconfigurable intelligent surface (RIS) has been recently proposed to change and adjust the

communication channels to improve the communication quality of services (QoS) [26]–[29]. In the context of PLS, RIS can (i) serve as a new degree-of-freedom (DoF) for optimizing the secrecy rate in key-less PLS [30], [31], and (ii) increase channel randomness by its phase controller for secret key generation [32]–[35]. To be specific, by randomly assigning the RIS phase in each channel estimation round, the reciprocal randomness of legitimate channels can be artificially enhanced, enabling a fast generation of the shared secret key. Based on this idea, [32] computes the SKR of RIS-secured low-entropy channel, and [36] further designs an optimal RIS phase set by maximizing the theoretical SKR.

The advance of RIS also provides new attack and eavesdropping potentials. This can be categorized as attackers that (i) destroy or (ii) maintain the channel reciprocity, where the former aims to ruin the legitimate PL-SKG, and the latter tries to obtain the legitimate secret keys. For example, the attacker in [37] controlled a RIS to damage PL-SKG at legitimate parties, by destroying the channel reciprocity via a fast change of RIS phase in the legitimate key generation process. However, the attackers that destroy the channel reciprocity cannot obtain legitimate secret keys without being detected. In this work, we focus on the second category, i.e., how a RIS can reconstruct the legitimate secret keys, by generating and inserting a deceivingly reciprocal and random channel into legitimate channels (named as Eve-RIS). Note that a similar idea, known as the secret key leakage attack by an adversarial RIS, has been proposed in the work [38]. However, the critical detailed implementation (e.g., how the RIS pursues channel estimation and how to optimize the deceiving channel) is missing. Besides, the countermeasure claimed in [38], (i.e., the two-way cross multiplication method in [17]) is actually what we are going to attack in this work (in Section III. B).

Indeed, an untrusted relay can pursue a similar man-in-the-middle (MITM) attack by generating and inserting a reciprocal channel to obtain legitimate secret keys. However, such a threat can be easily addressed by designing appropriate relay pilot transmission protocols [39], [40] (see Section IV. A for details). RIS, given its reflective property (unable to actively send pilots for protocol and authentication purposes), is naturally resistant to these countermeasures, and therefore paves a way to realize this MITM attack in a more concealed way. Also, given its ability to manipulate channels, we show (in Fig. 7) that a passive adversarial RIS (with 1600 elements) can achieve a comparable eavesdropping performance with the untrusted relay using 60dB amplifying gain. In this view, our proposed Eve-RIS provides a novel instance of MITM attacks, which is more energy-efficient and in a more concealed manner. A comprehensive comparison is provided in Section IV. A.

B. Contributions & Paper Structure

In this work, we aim to design an adversarial Eve-controlled RIS-based eavesdropping scheme. Eve-RIS aims to generate and insert a deceivingly random and reciprocal channel between Alice and Bob, so that their CSI-based secret key will be partially inferable to Eve. The main novel contributions are listed in the following.

(1) We show that an Eve-controlled RIS, by inserting a deceivingly random and reciprocal channel into legitimate channels, can partially reconstruct the PL-SKG based secret keys from legitimate users. The resulting theoretical key match rate between the proposed Eve-RIS and the legitimate users is deduced, whose geometry and qualitative properties provide insights for further designs and implementations.

(2) Operationalizing this, we design two eavesdropping schemes against (i) the CSI-based PL-SKG, and (ii) the two-way cross multiplication-based PL-SKG, respectively. Equipped with a small number of radio-frequency (RF) chains, compressed sensing-based baseband channel probing and feature extraction methods are designed for Eve-RIS. Then, the optimal RIS phase is obtained by maximizing the Eve-RIS inserted deceiving channel.

(3) We perform a comprehensive comparison between our designed Eve-RIS and other popular attackers, which are categorized by their ability to maintain or destroy the channel reciprocity. Specially, compared to the untrusted relay, the main difference is that the designed Eve-RIS is resistant to the countermeasures of the former. Then, from the implementation perspective, given the lack of receiving modules and the limited reflective gain compared with untrusted relays, the Eve-RIS design focuses on the optimal placement of the active RF chains to extract legitimate features, and the optimization of the deceivingly inserted channel to overcome the cascaded attenuation effect.

(4) We evaluate our proposed Eve-RIS via simulations. The results show that even with a passive RIS, our proposed Eve-RIS can achieve a high key match rate with legitimate users, and is resistant to most of the current defensive approaches. As such, our proposed Eve-RIS provides a new eavesdropping threat on PL-SKG, which should be seriously considered by further secret key designs to protect the confidentiality of wireless communications.

The rest of this work is structured as follows. In Section II, we describe the Eve-RIS model in the Alice-Bob scenario. In Section III, we elaborate on our design of Eve-RIS schemes against channel estimation based and two-way cross multiplication based PL-SKGs. In Section IV, we compare our designed Eve-RIS with other popular attackers, from the conceptual perspective. In Section V, we show our simulation results. We finally conclude this work in Section VI.

In this work, we use bold lower-case letters for vectors, and bold capital letters for matrices. We use $\|\cdot\|_2$ to denote the 2-norm, $\|\cdot\|_0$ to denote the 0-norm, and $diag(\cdot)$ to diagonalize a vector. $|\cdot|$ represents the absolute value of a complex value. We denote $mod(\cdot,\cdot)$ as the modulus operator and $\lfloor\cdot\rfloor$ is to truncate the argument. The matrix transpose, conjugate transpose, element-wise conjugate, Hadamard product operators and trace are denoted as $(\cdot)^T$, $(\cdot)^H$, $(\cdot)^*$, $\odot$ and $tr(\cdot)$. $\mathbb{E}(\cdot)$ and $\mathbb{D}(\cdot)$ represent the expectation and variance. $\mathcal{CN}(\mu, 2\sigma^2)$ is to represent the complex Gaussian distribution with mean as μ and variance as $2\sigma^2$.

II. System Model & Problem Formulation

In this work, two legitimate users Alice and Bob are to generate a shared secret key, leveraging the reciprocal channels

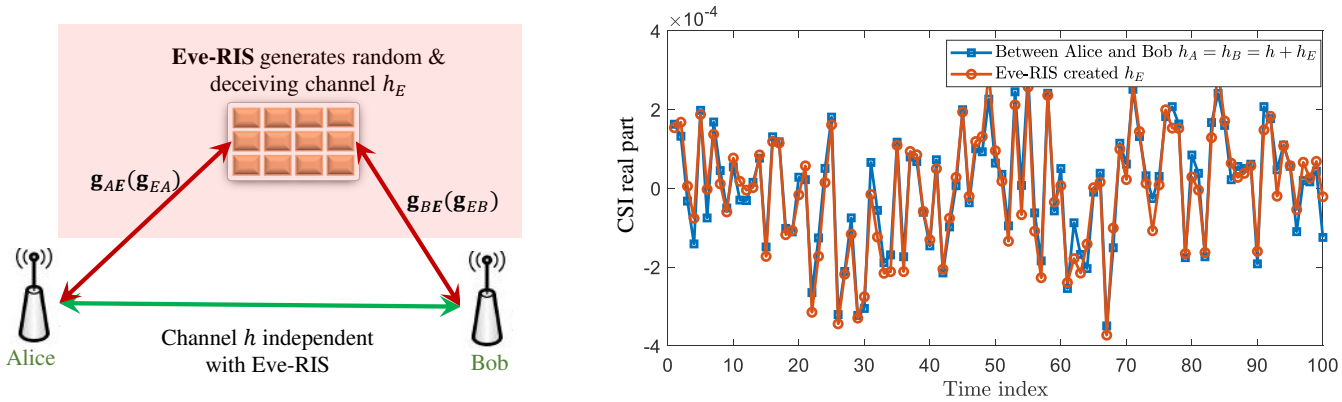

Fig. 1. Sketch of Eve-RIS: (a) the deployment of Eve-RIS to generate and insert a deceiving and random channel h_E, serving as a part of the legitimate channel between Alice and Bob, i.e., $h_A = h_B = h + h_E$, (b) illustration of $h_A = h_B \approx h_E$ by their real parts.

between them. Eve pursues eavesdropping by generating and inserting a random channel between Alice and Bob, which is achieved by a RIS, with uniform planar array (UPA) of size $M = M_x \times M_y$ (see Fig. 1(a)). Different from the general RIS design in [27], we refer to the hardware architecture in [41] to facilitate RIS's baseband channel estimation: A few of reflective elements are deployed with channel sensors, each of which connects to an RF chain for baseband measurements and signal processing.

The direct channel between Alice and Bob (irrelevant with Eve-RIS) is modeled as [42], [43]:

$$h \sim \mathcal{CN}(0, 2\sigma_h^2),\ 2\sigma_h^2 = C_0 d_{AB}^{-\alpha_N}, \tag{1}$$

where C_0 is the path loss at the reference distance (i.e., $1m$), d_{AB} is the LoS distance between Alice and Bob, and α_N is the NLoS path-loss exponent. Here, we omit the LoS part between Alice and Bob, since it cannot provide any randomness for PL-SKG, and can be easily removed by subtracting the means of channel probing results at Alice and Bob.

The channels from Alice and Bob to Eve-RIS are expressed as $\mathbf{g}_{aE} \sim \mathcal{CN}(\mathbf{g}_{aE}^{(\text{LoS})}, 2\mathbf{\Sigma}_{aE})$ $a \in \{A, B\}$, which are modeled as [44]:

$$\begin{aligned} \mathbf{g}_{aE} &= \mathbf{g}_{aE}^{(\text{LoS})} + \sum_{n=1}^{\iota} \frac{\rho_{aE,n}}{\sqrt{\iota}} \cdot \mathbf{u}\left(el_{aE,n}, az_{aE,n}\right) \\ \mathbf{g}_{aE}^{(\text{LoS})} &= \sqrt{C_0 \cdot d_{aE}^{-\alpha_L}} \cdot \mathbf{u}\left(el_{aE}^{(LoS)}, az_{aE}^{(LoS)}\right) \\ \rho_{aE,n} &\sim \mathcal{CN}(0, C_0 \cdot d_{aE}^{-\alpha_N}) \end{aligned} \tag{2}$$

In Eq. (2), d_{aE} is the LoS distance between a to Eve-RIS. ι is the number of NLoS Rayleigh paths and $\rho_{aE,n}$ is the gain for nth path. $\mathbf{u}(el, az) \triangleq [\exp(j\mathbf{a}(el, az)\mathbf{l}_1), \cdots, \exp(j\mathbf{a}(el, az)\mathbf{l}_M)]^T$, with $\mathbf{a}(el, az) \triangleq \frac{2\pi}{\lambda}[\sin(el)\cos(az), \sin(el)\sin(az), \cos(el)]$ and $\mathbf{l}_m \triangleq [0, mod(m-1, M_x)d, \lfloor (m-1)/M_y \rfloor d]^T$. $el_{aE,n}, az_{aE,n} \in [-\pi/2, \pi/2]$ are the half-space elevation and azimuth angles of nth path. For the structure of RIS, a square shape element is used with the size as $d \times d$, where $d = \lambda/8$ is set (i.e., less than half-wavelength $\lambda/2$ [44]–[46]).

With the modeling of the direct channels, the Eve-RIS generated channel, denoted as h_E, and its combined channels from Bob to Alice (Alice to Bob), denoted as h_A (h_B), can be expressed as follows:

$$\begin{aligned} h_E &= \mathbf{g}_{BE}^T \cdot diag(\mathbf{w}) \cdot \mathbf{g}_{AE} \sim \mathcal{CN}(0, 2\sigma_E^2), \\ h_A &= h_B = h + h_E. \end{aligned} \tag{3}$$

In Eq. (3), $\mathbf{w} = \sqrt{A_E}[\exp(j\theta_1), \cdots, \exp(j\theta_M)]^T$ is the phase vector of the Eve-RIS, with $A_E \in \mathbb{R}^+$ the amplifying gain. $\theta_m \in [0, 2\pi)$ with $m \in \{1, \cdots, M\}$ is the phase of mth RIS element. In this work, a random and an optimized RIS strategies will be studied: The random strategy is to select mth elemental phase of $\mathbf{w}$ identically and randomly over $[0, 2\pi)$, i.e., $\theta_m \in \mathcal{U}[0, 2\pi)$, while the optimized strategy is to derive an optimal (sub-optimal) $\mathbf{w}$ that improves the eavesdropping performance (will be detailed in Section III. C). Under these two strategies, the probability density distribution (PDF) of h_E can be approximated as:

$$h_E \sim \mathcal{CN}(\mu_E, 2\sigma_E^2), \tag{4}$$

for a large number of Eve-RIS elements (see Appendix A for details). In Eq. (4), μ_E and σ_E^2 are computed as:

$$\mu_E = \begin{cases} 0, & \text{random } \mathbf{w}, \\ \left(\mathbf{g}_{BE}^{(\text{LoS})}\right)^T \cdot diag(\mathbf{w}) \cdot \mathbf{g}_{AE}^{(\text{LoS})}, & \text{fixed } \mathbf{w}, \end{cases} \tag{5}$$

$$\sigma_E^2 = \begin{cases} 0.5A_E \cdot M \cdot C_0^2 \cdot d_{AE}^{-\alpha_L} \cdot d_{BE}^{-\alpha_L}, & \text{random } \mathbf{w}, \\ \mathbf{w}^H \cdot \mathbf{G} \cdot \mathbf{w}, & \text{fixed } \mathbf{w} \end{cases} \tag{6}$$

where $\mathbf{G} \triangleq 2\mathbf{\Sigma}_{AE} \odot \mathbf{\Sigma}_{BE} + diag(\mathbf{g}_{BE}^{(\text{LoS})})^* \mathbf{\Sigma}_{AE} diag(\mathbf{g}_{BE}^{(\text{LoS})}) + diag(\mathbf{g}_{AE}^{(\text{LoS})})^* \mathbf{\Sigma}_{BE} diag(\mathbf{g}_{AE}^{(\text{LoS})})$. The detailed deductions are provided in Appendix A. From Eq. (6), the variance σ_E^2 is determined by (i) the multiplication of variances of two sub-channels, i.e., $C_0 d_{AE}^{-\alpha_L} \cdot C_0 d_{BE}^{-\alpha_L}$, known as the cascaded channel attenuation, (ii) the number of RIS element M, and (iii) the RIS programmable phase vector $\mathbf{w}$. We will further show how these will be used to design our Eve-RIS schemes.

With the formulated model, the purpose of this work is to design how the Eve-RIS can eavesdropping the secret key between Alice and Bob. We will study two key generation cases: (i) SKG using channel estimation results, and (ii) SKG using the two-way method.

III. Designs of Eve-RIS against PL-SKG

We first give a sketch of the Eve-RIS scheme. As Alice and Bob use their reciprocal channels, i.e., $h_A = h_B$, for secret key generation, they do not know that h_A and h_B contains the Eve-RIS's deceiving channel h_E, i.e., $h_A = h_B = h + h_E$ (see Fig. 1(b)). In this view, a large variance of h_E, i.e., σ_E^2, will lead to a high correlation coefficient of h_E and h_A (h_B), i.e., $corr(h_E, h_A) = \sigma_E^2/\sqrt{\sigma_E^2 + \sigma_h^2}$, and therefore a high secret key match rate between Eve-RIS and Alice (Bob).

To be specific, we consider a general two-threshold quantization method, i.e., [14]

$$k_a = \begin{cases} 1, & z_a > \gamma_1^{(a)}, \\ 0, & z_b < \gamma_0^{(a)}, \end{cases} \quad a \in \{A, B, E\}, \tag{7}$$

where $\gamma_1^{(a)} = \mathbb{E}(z_a) + \beta\sqrt{\mathbb{D}(z_a)}$ and $\gamma_0^{(a)} = \mathbb{E}(z_a) - \beta\sqrt{\mathbb{D}(z_a)}$ are the upper and the lower quantization thresholds, with quantization threshold parameter $\beta \in [0, 0.5)$. In Eq. (7), z_a can be either $Re[h_a]$ or $Im[h_a]$, or the combination of $Re[h_a]$ and $Im[h_a]$. To simplify the further analysis, we assign $z_a = Re[h_a]$. The theoretical key match rate between Alice and Eve can be computed as:

$$\begin{aligned} &Pr\{k_A = k_E\} \\ =&\sqrt{\frac{2}{\pi}} \int_{\beta}^{+\infty} \Phi\left(-\beta\sqrt{\frac{\sigma_E^2}{\sigma_h^2} + 1} + \frac{\sigma_E}{\sigma_h}\zeta\right) \exp\left(-\frac{\zeta^2}{2}\right) d\zeta, \end{aligned} \tag{8}$$

where $\Phi(\cdot)$ is the cumulative distribution function of a normal distribution (see Appendix B for detailed deduction). It is noteworthy that a more compact version of Eq. (8) is not available, but it can still provide insights for the design and further implementation of Eve-RIS. One illustration of Eq. (8) is provided in Fig. 2.

First, the key match rate $\Pr\{k_A = k_E\}$ is monotonically increasing with the ratio between the variances of the Eve-RIS generated channel and the direct legitimate channel, i.e., σ_E^2/σ_h^2 (proved in Appendix B and shown in Fig. 2). Specially, we have,

$$\lim_{\sigma_E^2/\sigma_h^2 \to +\infty} Pr\{k_A = k_E\} = Pr\{k_A = k_B\} = 2\Phi(-\beta), \tag{9}$$

suggesting that in the wave-blockage scenario, the key match rate between Eve-RIS and legitimate users reaches the limit, i.e., $Pr\{k_A = k_E\} = Pr\{k_A = k_B\}$, which is determined by the quantization parameter β set by the legitimate users.

Second, $\Pr\{k_A = k_E\}$ at first increases faster and then gradually, with the increase of σ_E^2/σ_h^2 (shown in Fig. 2). This can be proved by the fact that its second-order derivative is less than 0 (see Appendix B for details). Such a phenomenon indicates the potential of the Eve-RIS, whereby a slight manipulation of its deceiving channel can make a huge legitimate secret key leakage.

Third, from the implementation view, Eq. (8) provides a theoretical mapping between the Eve-RIS's achievable key match rate and its controlling variable, i.e., σ_E^2. Such a mapping serves as the reference for the design of Eve-RIS to determine an appropriate σ_E^2, given the specific requirement of the eavesdropping key match rate, i.e., $Pr\{k_A = k_E\}$.

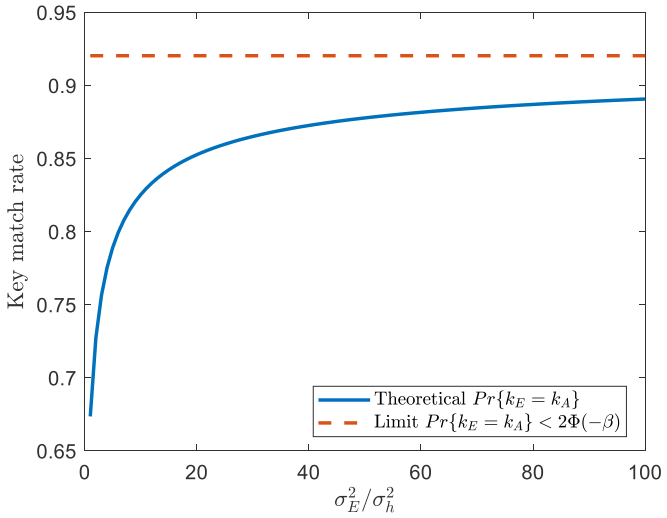

Fig. 2. Theoretical key match rate between proposed Eve-RIS and legitimate users, i.e., $Pr\{k_A = k_E\}$, which serves as a reference for the implementation of Eve-RIS to determine the variance of its generated deceiving channel σ_E^2.

Leveraging this, further detailed implementation to optimize σ_E^2 can be pursued (e.g., balancing the number of RIS elements that will be used and the active reflecting gain, as they both increase σ_E^2 from Eq. (6), and maximizing σ_E^2 by RIS phase).

We next elaborate on the details of our Eve-RIS to attack two popular PL-SKG schemes.

A. Eavesdropping CSI-based PL-SKG

1) PL-SKG using CSI Estimation: We first show the process of PL-SKG using the estimations of the legitimate reciprocal channel. In this case, Alice and Bob estimate the reciprocal channel in time-division duplex (TDD) mode, whereby in each channel estimation round, the channel between Alice and Bob remains unchanged. In odd and even time slots, Alice and Bob respectively send pilot sequence $\mathbf{x}_A, \mathbf{x}_B \in \mathbb{C}^{1\times L}$. Then, the channels estimated at Alice and Bob, denoted as $\hat{h}_A$ and $\hat{h}_B$, are [32]–[35]:

$$\begin{aligned} \hat{h}_A &= \frac{\mathbf{y}_A \cdot \mathbf{x}_B^H}{\|\mathbf{x}_B\|_2^2} = (h + h_E) + \hat{n}_A, \\ \hat{h}_B &= \frac{\mathbf{y}_B \cdot \mathbf{x}_A^H}{\|\mathbf{x}_A\|_2^2} = (h + h_E) + \hat{n}_B, \end{aligned} \tag{10}$$

In Eq. (10), $\mathbf{y}_A = h_A \cdot \mathbf{x}_B + \mathbf{n}_A$ and $\mathbf{y}_B = h_B \cdot \mathbf{x}_A + \mathbf{n}_B$ are the received signals at Alice and Bob, with $\mathbf{n}_A, \mathbf{n}_B \sim \mathcal{CN}(0, 2\sigma_n^2\mathbf{I}_L)$, the receiving noise components. $\hat{n}_A \sim \mathcal{CN}(0, 2\sigma_n^2/\|\mathbf{x}_B\|_2^2)$, $\hat{n}_B \sim \mathcal{CN}(0, 2\sigma_n^2/\|\mathbf{x}_A\|_2^2)$ are the estimating noises. As such, leveraging the common channel estimations, i.e., $\hat{h}_A$ and $\hat{h}_B$, secret keys can be generated from Alice and Bob, by replacing z_A and z_B with $\hat{z}_A = Re[\hat{h}_A]$ and $\hat{z}_B = Re[\hat{h}_B]$, respectively, in Eq. (7).

2) Eavesdropping design: The purpose of Eve-RIS is to generate and insert a deceivingly random channel h_E that contributes to part of h_A and h_B. To do so, for each channel legitimate estimation round of Alice and Bob, Eve-RIS assigns a RIS phase vector $\mathbf{w}$ by either random or optimized strategies. Here the random strategy is to select mth elemental phase of $\mathbf{w}$

identically and randomly over $[0, 2\pi)$, i.e., $\theta_m \in \mathcal{U}[0, 2\pi)$. The optimized strategy will be elaborated on in Section III. C. In order to hold randomness and reciprocity, such a $\mathbf{w}$ will remain unchanged during one Alice-Bob channel estimation round but will change independently for different channel estimation rounds.

Eve-RIS will pursue estimations of Alice to Eve-RIS, and Bob to Eve-RIS channels, i.e., $\mathbf{g}_{AE}$ and $\mathbf{g}_{BE}$, by their sending pilots. According to the RIS architecture in [41], the elements with active channel sensors receive the transmitted signals and feed them into the RF chains for further baseband channel estimation. Then, the compressed sensing-based channel estimation is pursued, due to the sparse representation of $\mathbf{g}_{AE}$ and $\mathbf{g}_{BE}$ in the beamspace dictionary, induced by the small number of Rayleigh scattering paths, i.e., $\iota \leq 5$ [47]. As such, the Eve-RIS RF chain received baseband signals from Alice and Bob in one channel estimation round are:

$$\begin{aligned} \mathbf{Y}_E^{(A)} &= \mathbf{C} \cdot \mathbf{D} \cdot \mathbf{s}_{AE} \cdot \mathbf{x}_A + \mathbf{N}_E^{(A)}, \\ \mathbf{Y}_E^{(B)} &= \mathbf{C} \cdot \mathbf{D} \cdot \mathbf{s}_{BE} \cdot \mathbf{x}_B + \mathbf{N}_E^{(B)}, \end{aligned} \tag{11}$$

where $\mathbf{N}_E^{(A)}$ and $\mathbf{N}_E^{(B)}$ are the noise components, whose elements are i.i.d complex Gaussian distributed with variance σ_n^2. In Eq. (11), $\mathbf{D} \in \mathbb{C}^{M \times D}$ is the dictionary designed by the spanning of the RIS beamspace:

$$\mathbf{D} = \left[\mathbf{u}\left(el_1, az_1\right), \cdots, \mathbf{u}\left(el_D, az_D\right)\right], \tag{12}$$

where the pairs $(el_1, az_1), \cdots, (el_D, az_D)$ evenly enumerate the joint space of azimuth and elevation angles, i.e., $[-\pi/2, \pi/2] \times [-\pi/2, \pi/2]$. With the design of the dictionary, we have $\mathbf{g}_{AE} = \mathbf{D} \cdot \mathbf{s}_{AE}$ and $\mathbf{g}_{BE} = \mathbf{D} \cdot \mathbf{s}_{BE}$ in Eq. (11), where $\mathbf{s}_{AE}, \mathbf{s}_{BE} \in \mathbb{C}^D$ are $\iota \leq 5$-sparse due to the fact that the number of their Rayleigh paths is less than 5 [47].

In Eq. (11), $\mathbf{C} \in \mathbb{R}^{C \times M}$ is the sensing matrix where each row only has one nonzero element, representing one entry of RIS element equipped with channel sensor and RF chain. It is noteworthy that the placement of RF chains affects the accuracy of channel estimation and the feature extraction of the legitimate nodes, which further influences the eavesdropping performance. The selection of $\mathbf{C}$ should ensure the restricted isometry property (RIP) [48], [49], i.e., the condition number of any 2ι columns of $\mathbf{CD}$ should be smaller than a threshold, which is an NP-hard problem. Here, a sub-optimal greedy strategy is proposed as:

$$\mathcal{C} \leftarrow \mathcal{C} \cup \{n\}, \ n = \underset{i}{\operatorname{argmin}}\, cond\left(\mathbf{D}_{\mathcal{C}+\{i\},:}\right), \tag{13}$$

where $\mathcal{C}$ is the set of the columns of the non-zero entries in $\mathbf{C}$, and $\mathbf{D}_{\mathcal{C},:}$ is the submatrix of $\mathbf{D}$ whose rows are selected by $\mathcal{C}$. Note that the selection in Eq. (13), i.e., the deployment of sensors and RF chains, is an off-line procedure, since $\mathbf{D}$ is fixed as the structure of RIS is determined, and its time and complexity consumption will not affect the real-time channel estimation. The number of channel sensors and RF chains, i.e., $|\mathcal{C}| = C$, is generally set as a little larger than $2 \times \iota$. In the context of channel estimation, this number can be further reduced by extending the spatial measurements via the large time-span pilots [47].

From the baseband measurements in Eq. (11), and the sparse representation by Eq. (12), the compressed sensing-based channel estimations at Eve-RIS are pursued by:

$$\begin{aligned} &\min_{\mathbf{s}_{AE}} \left\| \frac{\mathbf{Y}_E^{(A)} \cdot \mathbf{x}_A^H}{\|\mathbf{x}_A\|_2^2} - \mathbf{CD} \cdot \mathbf{s}_{AE} \right\|_2^2, \ s.t., \ \|\mathbf{s}_{AE}\|_0 \leq \iota \\ &\min_{\mathbf{s}_{BE}} \left\| \frac{\mathbf{Y}_E^{(B)} \cdot \mathbf{x}_B^H}{\|\mathbf{x}_B\|_2^2} - \mathbf{CD} \cdot \mathbf{s}_{BE} \right\|_2^2, \ s.t., \ \|\mathbf{s}_{BE}\|_0 \leq \iota \end{aligned} \tag{14}$$

To solve Eq. (14), one may loosen the constraints by 1-norm and then adopt the subgradient method, or use the orthogonal matching pursuit (OMP) algorithm [50].

By denoting the solutions of Eq. (14) as $\hat{\mathbf{s}}_{AE}$ and $\hat{\mathbf{s}}_{BE}$, the channel estimation results of $\mathbf{g}_{AE}$ and $\mathbf{g}_{BE}$ are:

$$\begin{aligned} \hat{\mathbf{g}}_{AE} &= \mathbf{D} \cdot \hat{\mathbf{s}}_{AE}, \\ \hat{\mathbf{g}}_{BE} &= \mathbf{D} \cdot \hat{\mathbf{s}}_{BE}. \end{aligned} \tag{15}$$

With the help of Eq. (15), Eve-RIS generated channel h_E can be estimated as:

$$\hat{h}_E = \hat{\mathbf{g}}_{BE}^T \cdot diag(\mathbf{w}) \cdot \hat{\mathbf{g}}_{AE}. \tag{16}$$

After the estimation of the Eve-RIS generated deceiving channel, i.e., $\hat{h}_E$ in Eq. (16), Eve can obtain the secret key by replacing z_E with $\hat{z}_E = Re[\hat{h}_E]$ of the quantization method in Eq. (7).

B. Eavesdropping PL-SKG using two-way method

Two-way PL-SKG leverages the random pilots sent from legitimate users to pursue channel randomization. This thereby prevents most of the untrusted relays and spoofing attackers that require exact channel estimations, e.g., [51]–[54]. In this part, we will show how Eve-RIS can obtain the legitimate secret keys generated by two-way PL-SKG.

1) PL-SKG using two-way method: In the two-way method, Alice and Bob send random pilots to each other in TDD mode, assigned as $q_A, q_B \in \mathbb{C}$. Here, several designs of the distribution of q_A and q_B have been made in the work [24], but in this work, the specific assignment will not affect our eavesdropping design. Then, Alice and Bob multiply their transmitted and received signals as their common features for further key quantization, i.e., [24], [55]

$$\begin{aligned} \hat{\phi}_A &= v_A \cdot q_A = (h + h_E) \cdot q_A \cdot q_B + \hat{\epsilon}_A \\ \hat{\phi}_B &= v_B \cdot q_B = (h + h_E) \cdot q_A \cdot q_B + \hat{\epsilon}_B, \end{aligned} \tag{17}$$

where $v_A = h_A \cdot q_B + n_A$ and $v_B = h_B \cdot q_A + n_B$ are the received signals at Alice and Bob, with $n_A, n_B \sim \mathcal{CN}(0, 2\sigma_n^2)$, the received noises. $\hat{\epsilon}_A = \epsilon_A \cdot q_A$ and $\hat{\epsilon}_B = \epsilon_B \cdot q_B$ are denoted for simplification. As such, by replacing z_A and z_B with $\hat{\varphi}_A = Re[\hat{\phi}_A]$ and $\hat{\varphi}_B = Re[\hat{\phi}_B]$ in Eq. (7), legitimate secret keys between Alice and Bob can be generated.

2) Eavesdropping design: The Eve-RIS design against two-way based PL-SKG shares similar parts to that against CSI-based PL-SKG, whereby a random or deliberately optimized phase vector $\mathbf{w}$ is assigned for each two-way key generation round. The difference is how Eve-RIS reconstructs the common feature, as exact channel estimations are unavailable due to the random pilots.

Given the equipped channel sensors and RF chains, the received baseband signals from Alice and Bob are:

$$\begin{aligned}\mathbf{r}_E^{(A)} &= \mathbf{C}\cdot\mathbf{D}\cdot\mathbf{s}_{AE}\cdot q_A + \boldsymbol{\varepsilon}_E^{(A)},\\ \mathbf{r}_E^{(B)} &= \mathbf{C}\cdot\mathbf{D}\cdot\mathbf{s}_{BE}\cdot q_B + \boldsymbol{\varepsilon}_E^{(B)},\end{aligned} \tag{18}$$

with $\boldsymbol{\varepsilon}_E^{(A)}, \boldsymbol{\varepsilon}_E^{(B)} \sim \mathcal{CN}(0, 2\sigma_n^2\mathbf{I}_C)$ the noise components. From Eq. (18), it is seen that an accurate estimation of $\mathbf{s}_{AE}$ ($\mathbf{s}_{BE}$) is unavailable, due to the involvement of the unknown random pilots q_A (q_B). However, it is noticed from Eq. (17) that with a large variance of Eve-RIS's deceiving channel, i.e., σ_E^2, the following approximation holds:

$$\begin{aligned}&(h+h_E)\cdot q_A\cdot q_B \approx h_E\cdot q_A\cdot q_B\\ &=(q_B\cdot\mathbf{s}_{BE})^T\cdot\mathbf{D}^T\cdot diag(\mathbf{w})\cdot\mathbf{D}\cdot(\mathbf{s}_{AE}\cdot q_A)\end{aligned} \tag{19}$$

This envisages us to estimate the combined $\mathbf{s}_{AE}\cdot q_A$ ($\mathbf{s}_{BE}\cdot q_B$) from Eq. (18). Given the sparse representations, i.e., $\|\mathbf{s}_{AE}\cdot q_A\|_0 = \|\mathbf{s}_{AE}\|_0 = \iota < 5$ and $\|\mathbf{s}_{BE}\cdot q_B\|_0 = \|\mathbf{s}_{BE}\|_0 = \iota < 5$, the compressed sensing problem can be built as:

$$\begin{aligned}&\min_{\mathbf{s}_{AE}\cdot q_A}\left\|\mathbf{r}_E^{(A)} - \mathbf{CD}\cdot(\mathbf{s}_{AE}\cdot q_A)\right\|_2^2,\ s.t.,\ \|\mathbf{s}_{AE}\cdot q_A\|_0 \le \iota,\\ &\min_{\mathbf{s}_{BE}\cdot q_B}\left\|\mathbf{r}_E^{(B)} - \mathbf{CD}\cdot(\mathbf{s}_{BE}\cdot q_B)\right\|_2^2,\ s.t.,\ \|\mathbf{s}_{BE}\cdot q_B\|_0 \le \iota,\end{aligned} \tag{20}$$

where OMP is used to find the optimal (sub-optimal) solutions, denoted as $\widehat{\mathbf{s}_{AE}\cdot q_A}$ and $\widehat{\mathbf{s}_{BE}\cdot q_B}$.

Then, Eve-RIS can reconstruct part of $\hat{\phi}_A$ (or $\hat{\phi}_B$) as:

$$\hat{\phi}_E = \left(\widehat{\mathbf{s}_{BE}\cdot q_B}\right)^T\cdot\mathbf{D}^T\cdot diag(\mathbf{w})\cdot\mathbf{D}\cdot\left(\widehat{\mathbf{s}_{AE}\cdot q_A}\right) \tag{21}$$

After the derivation of the feature $\hat{\phi}_E$, Eve-RIS can regenerate the legitimate secret key of Alice and Bob. This is achieved by computing $\hat{\varphi}_E = Re[\hat{\phi}_E]$ and replacing z_E in Eq. (7), given the shared information between $\hat{\phi}_E$ and $\hat{\phi}_A$, i.e., $h_E\cdot q_A\cdot q_B$.

C. Eve-RIS Phase Optimization

From the key match rate analysis of the proposed Eve-RIS designs in Eq. (8), a better eavesdropping performance can be achieved by maximizing the variance of the Eve-RIS generated deceiving channel, i.e., σ_E^2. This is done by finding the optimized RIS phase vector $\mathbf{w}$. According to the expression of σ_E^2 in Eq. (6), the optimization problem is formulated as:

$$\begin{aligned}&\max_{\mathbf{w}}\ \mathbf{w}^H\mathbf{G}\mathbf{w},\\ &s.t.,\ \mathbf{w}^H\mathbf{E}_m\mathbf{w} = A_E,\ m=1,\cdots,M\end{aligned} \tag{22}$$

where $\mathbf{E}_m \triangleq \mathbf{e}_m\cdot\mathbf{e}_m^H$, and $\mathbf{e}_m$ denotes a unit-norm vector whose mth element is 1.

We first replace the bound constraints in Eq. (22) with an equivalent convex constraints, i.e.,

$$\begin{aligned}&\max_{\mathbf{w}}\ \mathbf{w}^H\mathbf{G}\mathbf{w},\\ &s.t.,\ \mathbf{w}^H\mathbf{E}_m\mathbf{w} \le A_E,\ m=1,\cdots,M.\end{aligned} \tag{23}$$

The equivalency proof is provided in Appendix C.

For Eq. (23), we use semidefinite relaxation (SDR) technique to find a sub-optimal solution [56]. To be specific, by denoting $\mathbf{W} \triangleq \mathbf{w}\mathbf{w}^H$, we have $\mathbf{W}$ as a positive semidefinite one rank matrix (i.e., $\mathbf{W}\succeq 0$, and $rank(\mathbf{W}) = 1$). Eq. (23) then can be transformed as:

$$\max_{\mathbf{W}}\ tr(\mathbf{G}\mathbf{W}), \tag{24a}$$

$$s.t.,\ W_{m,m} \le A_E,\ m=1,\cdots,M \tag{24b}$$

$$\mathbf{W}\succeq 0, \tag{24c}$$

$$rank(\mathbf{W}) = 1, \tag{24d}$$

where $W_{m,m}$ is the (m,m)th element of matrix $\mathbf{W}$. In Eq. (24), the optimization problem defined by Eqs. (24a)-(24c), i.e., excluding the rank-one constraint, is convex (i.e., a linear objective function to $\mathbf{W}$ with convex constraints), and therefore can be solved by CVX. Leveraging the CVX result $\mathbf{W}$, the sub-optimal rank-one solution is derived by the following randomization process [56]. We first sample N vectors as $\boldsymbol{\varpi}_n \sim \mathcal{CN}(0, \mathbf{I}_M),\ n=1,\cdots,N$, and construct $\boldsymbol{\omega}_n = \sqrt{A_E}\exp(j\cdot\arg(\boldsymbol{\Gamma}\boldsymbol{\Lambda}^{\frac{1}{2}}\boldsymbol{\varpi}_n))$, with $\boldsymbol{\Gamma}\boldsymbol{\Lambda}\boldsymbol{\Gamma}^H = \mathbf{W}$ the eigen-decomposition, and $\arg(\cdot)$ and $\exp(\cdot)$ the elemental-wise angle extraction and exponential operator for normalization. Then, the sub-optimal $\mathbf{w}$ can be obtained by

$$\mathbf{w} = \underset{n=1,\cdots,N}{\operatorname{argmax}}\ \boldsymbol{\omega}_n^H\mathbf{G}\boldsymbol{\omega}_n. \tag{25}$$

It has been demonstrated by [57] that the SDR approach with a large randomization number N guarantees an $\pi/4$-approximation of the optimal objective value of Eq. (24).

IV. Difference from Current Eavesdroppers

In this section, we compare our designed Eve-RIS with other popular attackers, from the conceptual perspectives. Here, we categorize the attackers by whether the channel reciprocity between Alice and Bob is maintained or destroyed.

A. Attackers Maintaining Channel Reciprocity

Attackers that maintain the channel reciprocity, aim at partially estimating secret keys, by generating and inserting a reciprocal part to Alice's and Bob's channel probing results. In a mathematical view, the channel estimation results at Alice, Bob, and the attacker, denoted as $\hat{\psi}_A$, $\hat{\psi}_B$ and $\hat{\psi}_E$, are expressed as:

$$\begin{aligned}\hat{\psi}_A &= h + \psi_E + \hat{n}_A,\\ \hat{\psi}_B &= h + \psi_E + \hat{n}_B,\\ \hat{\psi}_E &= \psi_E + \hat{n}_E,\end{aligned} \tag{26}$$

where ψ_E is the inserted part from the attackers. Here, we compare our proposed Eve-RIS with other three popular types of attackers maintaining channel reciprocity, i.e.,

$$\psi_E = \begin{cases} h_E & \text{Eve-RIS (proposed)}\\ \tilde{\mathbf{g}}_{BE}^T\cdot diag(\tilde{\mathbf{w}})\cdot\tilde{\mathbf{g}}_{AE} & \text{Untrusted relay in [39], [40]}\\ 1_{\|\tilde{\mathbf{g}}_{AE}\|_2\approx\|\tilde{\mathbf{g}}_{BE}\|_2}\cdot p & \text{Spoofing in [52]–[54]}\\ \tilde{\mathbf{g}}_{BE}^T\cdot\mathbf{p} = \tilde{\mathbf{g}}_{AE}^T\cdot\mathbf{p} & \text{Spoofing in [51]}\end{cases} \tag{27}$$

where $\tilde{\mathbf{g}}_{AE}$ and $\tilde{\mathbf{g}}_{BE}$ are the channels from Alice and Bob to the attacker, respectively (note that these channels are different from those of Alice and Bob to RIS, given the structural difference between RIS and other devices).

1) Attackers with Physical Reciprocity: Attackers that can physically maintain the channel reciprocity include our proposed Eve-RIS, and untrusted relays in [39], [40], whose combined channels (e.g., Alice to Eve-RIS/relay to Bob, and Bob to Eve-RIS/relay to Alice channels) naturally maintain the reciprocal property. Similar to the proposed Eve-RIS, the untrusted relay can assign its transit vector $\tilde{\mathbf{w}}$ in Eq. (27) to insert a deceiving channel ψ_E, which enables to obtain the legitimate secret keys.

The main differences between our proposed Eve-RIS and untrusted relays are in the following two aspects. First, untrusted relay attack is vulnerable to the existing relay transmission protocols [39], [40]. Such protocols request relays to send a pilot in the first place (e.g., phase 1 in [39], and step 1 in [40]). Then, leveraging the estimation of channels from relays to legitimate users, a more secured PL-SKG can be designed to generate secret keys with no leakage, which therefore compromises the threat from the untrusted relay and its performed MITM channel insertion attack. By contrast, the defensive approaches designed for untrusted relays cannot be implanted on our proposed Eve-RIS, as is less practical to assume a reflective surface to actively send pilots for protocol and authentication purposes. As such, the proposed Eve-RIS provides a better way to realize the MITM channel insertion attack, which therefore demonstrates an arising new eavesdropping threat that requires further countermeasure designs specific to Eve-RIS.

Second, from the implementation view, both the untrusted relay and our proposed Eve-RIS suffer from severe cascaded channel attenuation. For the untrusted relay, the variance of its inserted channel in Eq. (27) is computed as $\|\tilde{\mathbf{w}}\|_2^2 C_0^2 d_{AE}^{-\alpha_L} d_{BE}^{-\alpha_L}$. This suggests that the only way to counter the cascaded attenuation is to increase the relaying (amplifying) gain, i.e., $\|\tilde{\mathbf{w}}\|_2^2$ (e.g., in Fig. 7, a 60dB gain is required for a 25m attack). Such an amplifying gain is impractical even with active RIS elements. To approach a comparable channel attenuation compensation and eavesdropping performance with the untrusted relay, we leverage the advantages of RIS, i.e., can be equipped with more reflective elements, and its ability to optimize channel constrained by the limited reflective gain, given the eigenvalues gap of the subchannels' covariance matrix [44]. By doing so, the designed Eve-RIS, with 1600 passive elements, reaches equivalent performance with untrusted relays using 60dB gain (see Fig. 7). This comparable eavesdropping performance demonstrates that our proposed Eve-RIS provides a substitute pathway to implement MITM attack, which is not affected by the defensive relay transmission protocols that compromise the untrusted relay.

2) Attackers with Probabilistic Reciprocity: Attackers that have a probability to maintain channel reciprocity include the works in [52]–[54]. In these works, the attackers firstly estimate the RSS-based channels from Alice and Bob to them, i.e., $\|\tilde{\mathbf{g}}_{AE}\|_2^2$ and $\|\tilde{\mathbf{g}}_{BE}\|_2^2$, via Alice's and Bob's sending pilots. Then, the attacker waits for the chance of RSS-based channels reaching reciprocity, i.e., $\|\tilde{\mathbf{g}}_{AE}\|_2^2 \approx \|\tilde{\mathbf{g}}_{BE}\|_2^2$, and send spoofing signal to both Alice and Bob (Eq. (4) in [53]). As such, the RSS-based channel probing results at Alice and Bob contain the attacker-generated reciprocal parts, i.e., $p = \|\tilde{\mathbf{g}}_{AE}\|_2^2 \approx \|\tilde{\mathbf{g}}_{BE}\|_2^2$, which can be used by the attacker to partially obtain the legitimate secret keys.

The differences between our proposed Eve-RIS and these attackers are in the two aspects. First, the attackers with probabilistic reciprocity are harshly restricted by an "attack opportunity", i.e., $Pr\{\|\tilde{\mathbf{g}}_{AE}\|_2^2 - \|\tilde{\mathbf{g}}_{BE}\|_2^2 < \epsilon\}$, only within which the attack can be pursued. This suggests the inability to eavesdropping the secret keys over a long duration period. By contrast, our proposed Eve-RIS can physically maintain the channel reciprocity, thereby enabling it to pursue the estimation of the legitimate secret keys continuously, without the limitation of "attack opportunity".

Second, the measuring of insert opportunity requires the exact channel estimation of $\|\tilde{\mathbf{g}}_{AE}\|_2$ and $\|\tilde{\mathbf{g}}_{BE}\|_2$. This suggests the vulnerability of these attackers to channel randomization, i.e., the two-way cross multiplication PL-SKG in Eq. (5) of [54], by sending random pilots to disable the channel estimations at the attacker. By comparison, one distinguish of our proposed Eve-RIS is its ability to defeat the two-way PL-SKG (in Section III. B), which thereby constitutes the difference from the defensive point of view.

3) Attackers with Created Reciprocity: In Eq. (27), the spoofing scheme in [51] is to create and send artificial reciprocal signals to Alice and Bob, i.e., $\tilde{\mathbf{g}}_{BE}^T \cdot \mathbf{p}$ ($\tilde{\mathbf{g}}_{AE}^T \cdot \mathbf{p}$), respectively in the Alice's and Bob's pilot sending time slots. Here, the precoding $\mathbf{p}$ is designed by the attacker to maintain the channel reciprocity, i.e., $\tilde{\mathbf{g}}_{BE}^T \cdot \mathbf{p} = \tilde{\mathbf{g}}_{AE}^T \cdot \mathbf{p}$.

Given the aforementioned process, the differences between our proposed Eve-RIS and this attack can be categorized as two aspects. First, in the spoofing scheme [51], the precoding design, i.e., $\mathbf{p}$, requires exact estimations of channels from Alice and Bob to them, i.e., $\tilde{\mathbf{g}}_{BE}^T$ and $\tilde{\mathbf{g}}_{AE}$. This suggests its vulnerability if Alice and Bob use random pilots, i.e., the two-way based PL-SKG, which will ruin the attacker's channel probing process. This thereby provides a difference from the defensive perspective, as our designed Eve-RIS can defeat the two-way based PL-SKG (shown in Section III. B).

Second, even if ordinary pilots are used by Alice and Bob, the attacker in [51] is still difficult to obtain channel estimations of $\tilde{\mathbf{g}}_{AE}$ and $\tilde{\mathbf{g}}_{BE}$ in time to design its precoding $\mathbf{p}$. Consider the real-time scenario, where the channels from Alice and Bob to the spoofing Eve, i.e., $\tilde{\mathbf{g}}_{AE}$ and $\tilde{\mathbf{g}}_{BE}$, change independently for two consecutive channel estimation round (one channel estimation round contains an odd and an even time slot for Alice and Bob sending pilots in the TDD mode for channel estimations). In this view, when the spoofing Eve sends pilots in the same odd time slot as Alice, the spoofing Eve is hard to obtain the Bob-to-Eve channel of the current channel estimation round, i.e., $\tilde{\mathbf{g}}_{BE}$, since currently there is no pilot from Bob in the odd time slot (but will be in the following even time slot). As such, the precoding $\mathbf{p}$ to maintain channel reciprocity, i.e., $\psi_E = \tilde{\mathbf{g}}_{BE}^T \cdot \mathbf{p} = \tilde{\mathbf{g}}_{AE}^T \cdot \mathbf{p}$, cannot be generated in time for the attack in current channel estimation round. This, therefore, hinders the spoofing attack in [51] to obtain the legitimate secret keys relying on the fast time-varying and independent CSI.

B. Attackers Destroying Channel Reciprocity

Attackers that destroy the channel reciprocity include a wide range, e.g., pilot spoofing in [58], [59], and jamming [37]. Here, we only compare with the pilot spoofing, as the jamming attackers are not designed to obtain legitimate secret keys. In the existence of a pilot spoofing Eve, the channel probing results at Alice and Bob are not reciprocal, i.e., [58], [59]

$$\begin{aligned}\hat{\psi}_A &= h + \tilde{g}_{AE}\frac{\sqrt{E_s}}{\|\mathbf{x}_B\|_2} + \hat{n}_A,\\ \hat{\psi}_B &= h + \hat{n}_B,\\ \hat{\psi}_E^{(A)} &= \tilde{g}_{AE} + \hat{n}_E^{(A)},\end{aligned} \tag{28}$$

where $\tilde{g}_{AE} \sim \mathcal{CN}(\tilde{g}_{AE}^{(\text{LoS})}, 2\sigma_{AE}^2)$ is the single-antenna based channels from Alice and Bob to spoofing Eve. $E_S/\|\mathbf{x}_B\|_2^2$ is the spoofing gain of the legitimate pilots. Here, different from the spoofing in [51] that aims to maintain channel reciprocity, the spoofing Eve in Eq. (28) aims to pretend as one of the legitimate users (e.g., Bob) and generates secret keys with another (e.g., Alice). This is achieved by increasing the spoofing gain $E_s/\|\mathbf{x}_B\|_2^2$, so $\tilde{g}_{AE}$ will dominate the channel probing result at Alice, making $\hat{\psi}_A$ and $\hat{\psi}_E^{(A)}$ alike.

The main difference between our designed Eve-RIS and the spoofing Eve in [58], [59] is whether the legitimate channel is still reciprocal. In spoofing Eve scenarios, the channels between Alice and Bob are not reciprocal, due to the participation of the spoofing activity, i.e., $\hat{\psi}_A \neq \hat{\psi}_B$. In this view, Alice and Bob can compare their channel estimation results to determine whether a spoofing Eve exists. Compared to the spoofing Eve, the channels between Alice and Bob under Eve-RIS are still reciprocal, i.e., $h_A = h_B = h + h_E$. This, to some extent, helps conceal the Eve-RIS, as Alice and Bob cannot detect a considerable difference from their channel estimation results.

V. Simulation Results

In this section, we evaluate our designed Eve-RIS schemes. The model configuration is provided in the following. In a 3D space, Alice and Bob are located at $(0, 0, 0)$, $(0, 50, 0)$, with unit m. Eve (either the Eve-RIS or the peering attackers) is located at $(0, 10, 5)$, unless other specifications. The direct channels from Alice and Bob to Eve-RIS are modeled in Eq. (2) according to [44], where a square structure of RIS is considered, i.e., $M_x = M_y$. Here, the referenced path loss is set as $C_0 = -30$dB at the reference distance (i.e., 1m), and the LoS and NLoS path loss exponents are $\alpha_L = 2$ and $\alpha_N = 3$. The number of paths is $\iota = 5$ [42], where the first path is the LoS path and the rest 4 paths are NLoS paths with random half-space elevation and azimuth angles independently and randomly distributed over $\mathcal{U}[-\pi/2, \pi/2]$. For the PL-SKG using channel estimation results, the pilot sequences are set as publicly known with $\|\mathbf{x}_A\|_2^2 = \|\mathbf{x}_B\|_2^2 = 0.1$W. For the two-way PL-SKG method, we assign $q_A, q_B \sim \mathcal{CN}(0, 1\text{W})$. The variance of the receiving noise is assigned as $\sigma_n^2 = -110$dBW.

For the designed Eve-RIS, the number of channel sensors and RF chains for compressed sensing based channel feature extraction is set as $C = 20$. We examine different groups of amplifier gain A_E and the number of Eve-RIS elements M, where A_E ranges from 0dB (passive RIS) to 30dB and $M = M_x \times M_y$ are selected from $\{64, 100, 400, 1600\}$. It is noteworthy that with the increase of the reflective elements, there may be other issues spanning from theoretical analysis to practical manufacturing. Here, the maximal number of reflective elements is assigned as 1600, which is the same as the study in [44].

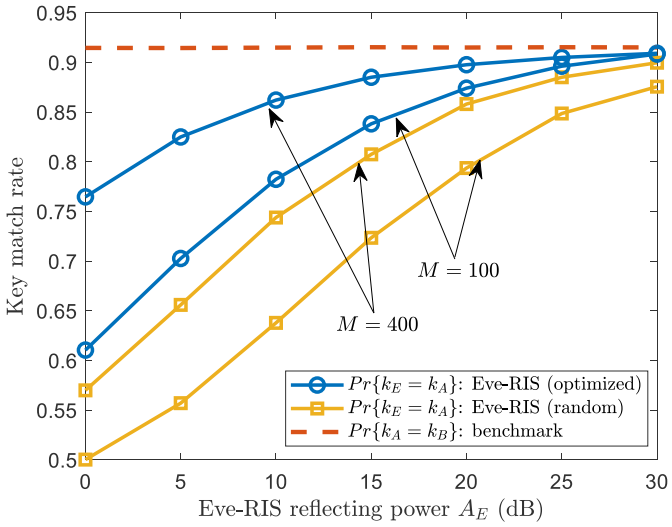

Fig. 3. Proposed Eve-RIS against CSI-based PL-SKG: Key match rate v.s. Eve-RIS amplifying gain.

A. Performance of Eve-RIS against CSI-based PL-SKG

1) Key match rate analysis: We first evaluate the key match rate between Alice and our proposed Eve-RIS, when attacking the CSI-based PL-SKG. For Fig. 3, the quantization threshold parameter is as $\beta = 0.1$. Further results for different β are shown by Figs. 4. In Fig. 3, the x-coordinate represents the amplifying gain of the Eve-RIS, i.e., A_E, while the y-coordinate gives the key match rate.

It is first seen that with the increase of the Eve-RIS amplifying gain A_E, the key match rate between Alice and Eve, i.e., $Pr\{k_E = k_A\}$, grows. When $A_E > 20$dB, $Pr\{k_E = k_A\}$ even approaches to key match rate of Alice and Bob, i.e., $Pr\{k_E = k_A\} \approx Pr\{k_A = k_B\} = 0.92$. Second, with the same Eve-RIS amplifying gain, a larger number of Eve-RIS elements, i.e., M, leads to a higher $Pr\{k_E = k_A\}$. For example, with the optimized Eve-RIS phase, when $A_E = 10$dB, $Pr\{k_E = k_A\}$ increases from 0.75 to 0.85 as M grows from 100 to 400. The reason behind these two observations is that both the amplifying gain and the number of Eve-RIS elements determine the variance of its generated deceiving channel, i.e., $\sigma_E^2 \propto A_E \cdot M$ given by Eq. (6), which, if increased, will increase $Pr\{k_E = k_A\}$ as deduced and analyzed by the theoretical key match rate in Eq. (8).

Then, it is observed that Eve-RIS using an optimized RIS phase has a larger key match rate with the legitimate user. For example, an increase of $Pr\{k_A = K_E\}$ from 0.65 to 0.85 at $A_E = 5$dB is obtained by the optimal RIS phase. This is attributed to the RIS's ability to manipulate channels, which

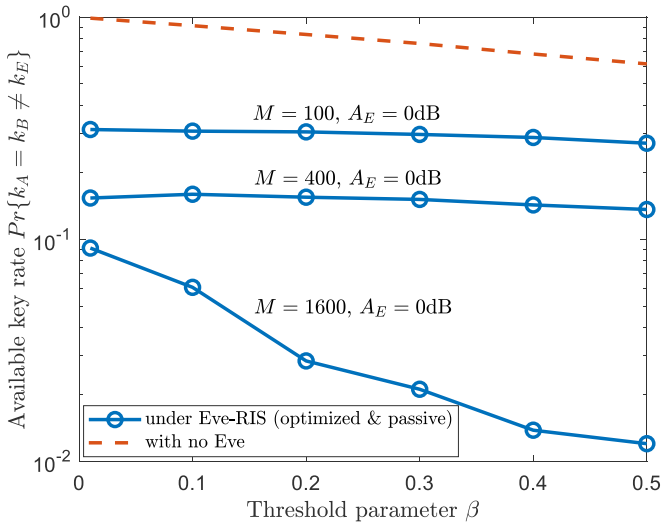

Fig. 4. Proposed Eve-RIS against CSI-based PL-SKG: Legitimate available key rate v.s. threshold parameter.

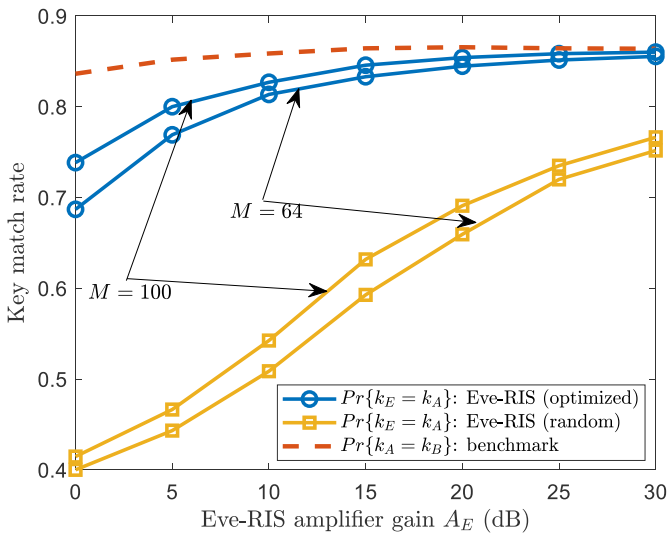

Fig. 5. Proposed Eve-RIS against two-way based PL-SKG: Key match rate v.s. Eve-RIS amplifying gain.

maximizes the variance of its generated deceiving channel and thereby achieves a better eavesdropping key match rate.

2) Legitimate key available rate analysis: We next define and test the available key rate between Alice and Bob under our designed Eve-RIS, i.e., $Pr\{k_A = k_B \neq k_E\}$. In Fig. 4, the x-coordinate is the quantization threshold parameter β while the y-coordinate is the available key rate, i.e., $Pr\{k_A = k_B \neq k_E\}$. It is first seen that with the increase of the quantization threshold β, all the available key rates between Alice and Bob with and without our designed Eve-RIS decrease. This is because a larger β leads to a larger upper quantization threshold γ_1 and a smaller lower quantization threshold γ_0, which reduces the total number of keys.

Second, we show in Fig. 4 that the optimized Eve-RIS with passive reflecting elements can drastically decrease the legitimate available key rate between Alice and Bob, i.e., $Pr\{k_A = k_B \neq k_E\}$. For instance, given a fixed threshold parameter as $\beta = 0.2$, $Pr\{k_A = k_B \neq k_E\}$ decreases from 0.80 (no Eve) to 0.3 (passive Eve-RIS with $M = 100$ elements). Such a legitimate key rate can be further reduced to 0.03 by the Eve-RIS using a larger number of elements (i.e., $M = 1600$). This is because the variance of the Eve-RIS generated deceiving channel, i.e., σ_E^2, can be increased by enhancing not only the reflecting gain A_E but also the number of RIS elements M, shown by Eq. (6). This further demonstrates the eavesdropping potential of our designed Eve-RIS: even a passive RIS can achieve a threatening secret key leakage attack.

B. Performance of Eve-RIS against two-way based PL-SKG

We next evaluate the eavesdropping performance of our proposed Eve-RIS, when attacking two-way based PL-SKG. Similar results to Eve-RIS against CSI-based PL-SKG (i.e., Figs. 3-4) can be seen in Figs. 5-6. The proposed Eve-RIS with optimized phase has better eavesdropping ability (i.e., high key match rate $Pr\{k_A = K_E\}$ and small legitimate key rate $Pr\{k_A = k_B \neq k_E\}$), with the enhancement of either

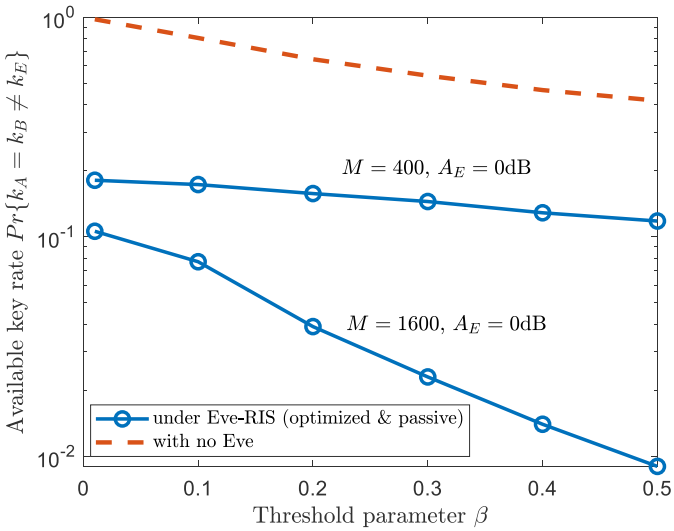

Fig. 6. Proposed Eve-RIS against two-way based PL-SKG: Available key rate v.s. quantization threshold parameter.

the number of RIS reflective elements M or the reflecting gain A_E. Furthermore, it is observed from Fig. 6 that a passive RIS (i.e., $M = 1600$, $A_E = 0$dB) can also provide a promising attack effect against the two-way based PL-SKG, with a reduction of the legitimate key rate to 10^{-2}.

The reason that our designed Eve-RIS can obtain the two-way PL-SKG based secret keys is different from the one against CSI-based PL-SKG, as the exact channel probing at Eve-RIS is unavailable due to the randomized channels by the two-way random pilots. Notably, neither the legitimate users nor the Eve-RIS relies on the original channels for key generation, but the features of randomized channels, i.e., $q_B(h + h_E)q_A$. For Eve-RIS, there is no information loss from the randomized Eve-RIS combined channel: the part of legitimate feature composed by this channel can be reconstructed with Eve-RIS's received signals, i.e., $q_B h_E q_A =$

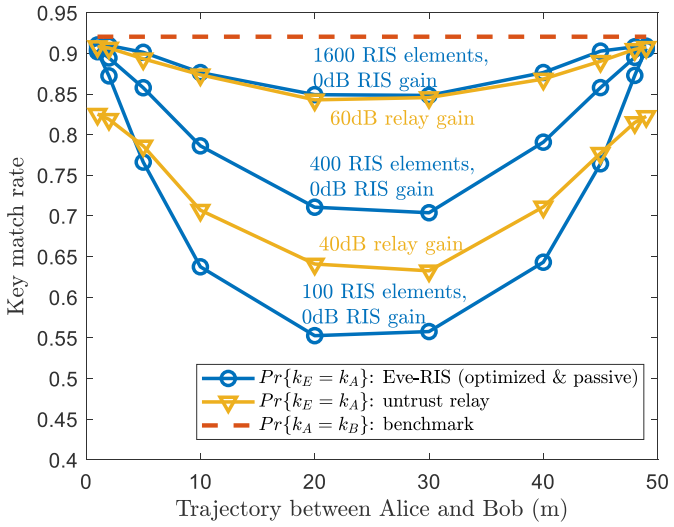

Fig. 7. Comparison between proposed Eve-RIS and untrusted relays when attacking CSI-based PL-SKG. The x-coordinate represents the trajectory of the attacker from $(0, 1, 0)$ to $(0, 49, 0)$ (Alice and Bob are located at $(0, 0, 0)$ and $(0, 50, 0)$ with unit m), and y-coordinate is the key match rate. A comparable key match rate with the untrusted relay using 60dB gain can be seen by the Eve-RIS equipped with 1600 passive elements.

$(q_B \mathbf{g}_{BE})^T \cdot diag(\mathbf{w}) \cdot (\mathbf{g}_{AE} q_A)$, which, therefore, guarantees the leakage attack of legitimate secret keys.

C. Comparison with Existing Attackers

The comparison between our proposed Eve-RIS and other popular attackers mentioned in Section IV is pursued in this part. For a fair comparison, only untrusted relay method is selected, since other schemes (e.g., [51]–[54]) cannot maintain an unconditional channel reciprocity as stated in Section IV.

We evaluate the key match rate between Eve (proposed Eve-RIS and the untrusted relay) and the legitimate user, concerning different Eve positions. In Fig. 7, the x-coordinate represents the trajectory of Eve from $(0, 1, 0)$ to $(0, 49, 0)$ (Alice and Bob are located at $(0, 0, 0)$ and $(0, 50, 0)$ with unit m), and y-coordinate is the key match rate. As shown in Fig. 7, 3 optimized Eve-RIS equipped with passive reflective elements are tested, where the number of RIS elements is selected from $M \in \{100, 400, 1600\}$.

It is first seen that the key match rate $Pr\{k_A = K_E\}$ follows a symmetrical pattern over the trajectory: as Eve-RIS moves from Alice to Bob, $Pr\{k_A = K_E\}$ drops at first and then grows back after it passes the middle point, i.e., $(0, 25, 0)$. This is because when Eve-RIS is close to one legitimate user, the cascaded channel attenuation $C_0^2 d_{AE}^{-\alpha_L} d_{BE}^{-\alpha_L}$ is smaller than when Eve-RIS is at the middle point, which therefore gives a larger variance of Eve-RIS's generated deceiving channel to obtain the legitimate secret keys.

Then, it is observed that as the number of RIS elements M increases (e.g., from 100 to 1600), $Pr\{k_A = K_E\}$ grows rapidly (e.g., from 0.55 to 0.85). This is because the variance of Eve-RIS generated and inserted channel can be increased by the enhancement of either the number of RIS elements M or the RIS's reflecting gain A_E, shown in Eq. (6). This thereby shows the eavesdropping potential of our proposed Eve-RIS, which, even equipped with passive reflective elements, can achieve a threatening secret key leakage attack.

Third, a comparison with the untrusted relay is shown in Fig. 7, where the proposed Eve-RIS equipped with $M = 1600$ passive RIS elements reaches a comparable $Kr\{k_A = k_E\}$ to the untrusted relay with 60dB relaying gain. Recalling that the attacks from both the proposed Eve-RIS and the untrusted relay suffer from the cascaded channel attenuation, which means the variances of their inserted deceiving channels should compensate for $C_0^2 d_{AE}^{-\alpha_L} d_{BE}^{-\alpha_L}$. In contrast to the untrusted relay that can only use amplifying gain for such compensation, our proposed Eve-RIS can leverage by (i) equipped with more RIS elements (as deduced in Eq. (6)), and (ii) the optimization of RIS phase to maximize the variance of its inserted channel. Also, given the resistance to the defensive approaches dealing with untrusted relays (analyzed in Section IV. A 1)), our proposed Eve-RIS provides a new instance to implement the MITM channel insertion attack, which requires further specific countermeasure designs.

D. Discussion of Results and potential Countermeasures

We here discuss the implementation and the potential countermeasures to our proposed Eve-RIS. To ensure a large inserted channel variance for secret key leakage attack, the implementation of the proposed Eve-RIS has two DoF, i.e., by either using a large number of RIS elements or increasing the amplifying gain. For example, when a structural constraint on RIS elements number is applied (e.g., $M = 100$), the active RIS structure designed in [29] may be adopted, which can provide a 20dB amplifying gain (with corresponding $Pr\{k_A = k_E\} = 0.86$ according to Fig. 3). On the other hand, such a level of $Pr\{k_A = k_E\}$ can also be achieved by a passive RIS with $M = 1600$ elements [44] (Fig. 7).

Then an open discussion on the potential countermeasures is provided. As studied in Section III, current PL-SKG based secret keys can be attacked and obtained by the proposed Eve-RIS. In this view, key-based PLS seems less attractive as a countermeasure or at least needs to be redesigned in the future works. Also, the defensive approaches for untrusted relays that rely on relay transmission protocols are not suitable, given the inability of RIS to actively send signals via its reflective elements. Then, from key-less PLS, one potential method may be the beamforming of legitimate users to minimize the variance of Eve-RIS inserted channels. For this way, one should also consider the Eve-RIS anti-beamforming ability, given that the number of RIS elements is much larger than the number of antennas of Alice and Bob.

VI. Conclusion

In this paper, we demonstrated that the advance in RIS for securing the wireless communications is a double-edged sword. On the one hand, recent work has shown that RIS can improve the channel randomness and secrecy rate of PLS [30]–[35]. On the other hand, our work here has shown that the presence of an adversarial Eve-controlled RIS (Eve-RIS) has the potential to reconstruct the PL-SKG based secret keys between Alice and Bob. We showed how the Eve-RIS can

achieve this by generating and inserting a deceivingly random and reciprocal channel. As a result, current PL-SKGs with channel estimation and two-way cross-multiplication methods can be eavesdropped by our designed Eve-RIS scheme.

Analysis and simulation results demonstrated the high key match rate obtained by our Eve-RIS with legitimate users, the low detectability as opposed to the spoofing Eve, and the resistance to most of the existing defensive approaches to untrusted relays. As such, our proposed Eve-RIS provides a new eavesdropping threat on PL-SKG and should be seriously considered by further PL-SKG designs and security works in securing wireless communications.

APPENDIX A
DEDUCTION OF EQS. (4)-(6)

We first re-write the expression of h_E from Eq. (3), i.e.,

$$h_E = \sum_{m=1}^{M} w_m \cdot g_{AE,m} \cdot g_{BE,m}, \tag{29}$$

where $g_{AE,m}$ and $g_{BE,m}$ are the mth element of $\mathbf{g}_{AE}$ and $\mathbf{g}_{BE}$, respectively. In Eq. (29), h_E is the summation of M random variables with weak dependence, since $g_{AE,m} \cdot g_{BE,m}$ can only be independent with $g_{AE,n} \cdot g_{BE,n}$ when nth RIS element is half-wavelength away from mth RIS element [44]. As such, given the central limit theorem under weak dependence (Theorem 27.5 in [60]), with a large number of RIS elements, e.g., $M > 50$, h_E can be approximated as complex Gaussian distribution, as shown in Eq. (4).

The mean and variance of h_E can be computed by re-writing h_E as the combinations of LoS and NLoS components, i.e.,

$$\begin{aligned} h_E &= \left[\mathbf{g}_{BE}^{(\text{LoS})} + \mathbf{g}_{BE}^{(\text{NLoS})}\right]^T \cdot diag(\mathbf{w}) \cdot \left[\mathbf{g}_{AE}^{(\text{LoS})} + \mathbf{g}_{AE}^{(\text{NLoS})}\right] \\ &= \left(\mathbf{g}_{BE}^{(\text{LoS})}\right)^T diag(\mathbf{w}) \mathbf{g}_{AE}^{(\text{LoS})} + \left(\mathbf{g}_{BE}^{(\text{NLoS})}\right)^T diag(\mathbf{w}) \mathbf{g}_{AE}^{(\text{LoS})} \\ &\quad + \left(\mathbf{g}_{BE}^{(\text{LoS})}\right)^T diag(\mathbf{w}) \mathbf{g}_{AE}^{(\text{NLoS})} + \left(\mathbf{g}_{BE}^{(\text{NLoS})}\right)^T diag(\mathbf{w}) \mathbf{g}_{AE}^{(\text{NLoS})} \end{aligned} \tag{30}$$

where $\mathbf{g}_{aE}^{(\text{LoS})}$ and $\mathbf{g}_{aE}^{(\text{NLoS})}$ ($a \in \{A, B\}$) represent the corresponding LoS and NLoS components, with relation $\mathbf{g}_{aE} = \mathbf{g}_{aE}^{(\text{LoS})} + \mathbf{g}_{aE}^{(\text{NLoS})}$. Thus, the mean and variance of h_E depends on whether the RIS phase $\mathbf{w}$ is random or fixed (optimized).

A. When RIS phase $\mathbf{w}$ is random

In this case, all 4 terms in Eq. (30) are random variables, determined by either or combinations of the independent random phase, i.e., $\mathbf{w}$, and NLoS channel components, i.e., $\mathbf{g}_{AE}^{(\text{NLoS})}$ and $\mathbf{g}_{BE}^{(\text{NLoS})}$. As such, the mean can be computed as:

$$\begin{aligned} \mathbb{E}(h_E) = &\left[\mathbf{g}_{BE}^{(\text{LoS})} + \mathbb{E}\left(\mathbf{g}_{BE}^{(\text{NLoS})}\right)\right]^T diag\left(\mathbb{E}(\mathbf{w})\right) \\ &\cdot \left[\mathbf{g}_{AE}^{(\text{LoS})} + \mathbb{E}\left(\mathbf{g}_{AE}^{(\text{NLoS})}\right)\right] = 0, \end{aligned} \tag{31}$$

since (i) $\mathbb{E}(\mathbf{w}) = \mathbf{0}$ when all M RIS phases of $\mathbf{w}$ are independently and identically distributed over $\mathcal{U}[0, 2\pi)$, and (ii) $\mathbb{E}[(\mathbf{g}_{AE}^{(\text{NLoS})})^T \mathbf{g}_{BE}^{(\text{NLoS})}] = \mathbb{E}(\mathbf{g}_{AE}^{(\text{NLoS})})^T \mathbb{E}(\mathbf{g}_{BE}^{(\text{NLoS})})$, given the in-dependency of $\mathbf{g}_{AE}^{(\text{NLoS})}$ and $\mathbf{g}_{BE}^{(\text{NLoS})}$ (as Alice and Bob are generally more than half-wavelength far from each other).

The variance of h_E is then computed by taking Eq. (29) and Eq. (31) into the its definition, i.e.,

$$\begin{aligned} &\mathbb{D}(h_E) = \mathbb{E}\left(h_E^* h_E\right) - \mathbb{E}\left(h_E^*\right)\mathbb{E}\left(h_E\right) = \mathbb{E}\left(h_E^* h_E\right) \\ =&\mathbb{E}\left(\left(\sum_{m=1}^{M} w_m^* g_{AE,m}^* g_{BE,m}^*\right) \cdot \left(\sum_{m'=1}^{M} w_{m'} g_{AE,m'}^* g_{BE,m'}^*\right)\right) \\ =&\sum_{m=1}^{M}\sum_{m'=1}^{M} \mathbb{E}\left(w_m^* w_{m'}\right) \mathbb{E}\left(g_{AE,m}^* g_{AE,m'}\right) \mathbb{E}\left(g_{BE,m}^* g_{BE,m'}\right) \\ \overset{(a)}{=}&\sum_{m=1}^{M} \mathbb{E}\left(|w_m|^2\right) \mathbb{E}\left(|g_{AE,m}|^2\right) \mathbb{E}\left(|g_{BE,m}|^2\right) \\ =&A_E \sum_{m=1}^{M} \left(2\Sigma_{AE,m,m} + \left|g_{AE,m}^{(\text{LoS})}\right|^2\right)\left(2\Sigma_{BE,m,m} + \left|g_{BE,m}^{(\text{LoS})}\right|^2\right) \\ \overset{(b)}{\approx}&A_E \cdot M \cdot C_0^2 \cdot d_{AE}^{-\alpha_L} \cdot d_{BE}^{-\alpha_L} \end{aligned} \tag{32}$$

where $\Sigma_{AE,m,m}$ is the (m,m)th element of matrix $\mathbf{\Sigma}_{AE}$, and $\Sigma_{BE,m,m}$ is the (m,m)th element of matrix $\mathbf{\Sigma}_{BE}$. In Eq. (32), (a) is due to $\mathbb{E}(w_m^* w_{m'}) = 0$, given the independent random phase assignment for different RIS elements, i.e., $m \neq m'$. The approximation in (b) is because the energy of LoS component is greatly larger than that of NLoS component. From Eq. (32), σ_E^2 in Eq. (6) can be computed by dividing 2.

B. When RIS phase $\mathbf{w}$ is fixed (optimized)

From Eq. (30), if $\mathbf{w}$ is fixed, then the mean of h_E is contributed by the cascaded LoS components, i.e.,

$$\begin{aligned} \mathbb{E}(h_E) =& \left[\mathbf{g}_{BE}^{(\text{LoS})} + \mathbb{E}\left(\mathbf{g}_{BE}^{(\text{NLoS})}\right)\right]^T diag\left(\mathbf{w}\right) \left[\mathbf{g}_{AE}^{(\text{LoS})} + \mathbb{E}\left(\mathbf{g}_{AE}^{(\text{NLoS})}\right)\right] \\ =& \left(\mathbf{g}_{BE}^{(\text{LoS})}\right)^T \cdot diag(\mathbf{w}) \cdot \mathbf{g}_{AE}^{(\text{LoS})}, \end{aligned} \tag{33}$$

which thereby completes the computation of $\boldsymbol{\mu}_E$ in Eq. (5).

For the variance of h_E, only the last 3 terms in Eq. (30) are random variables, determined by the NLoS channel components, i.e., $\mathbf{g}_{AE}^{(\text{NLoS})}$ and $\mathbf{g}_{BE}^{(\text{NLoS})}$, As such, the variance is:

$$
\begin{aligned}
\mathbb{D}(h_E)=&\mathbb{D}\Big\{\Big[\Big(\mathbf{g}_{AE}^{(\mathrm{NLoS})}\Big)^T diag\left(\mathbf{g}_{BE}^{(\mathrm{LoS})}\right)+\Big(\mathbf{g}_{BE}^{(\mathrm{NLoS})}\Big)^T diag\left(\mathbf{g}_{AE}^{(\mathrm{LoS})}\right)\\
&+\left(\mathbf{g}_{AE}^{(\mathrm{NLoS})}\odot\mathbf{g}_{BE}^{(\mathrm{NLoS})}\right)^T\Big]\mathbf{w}\Big\}\\
=&\mathbf{w}^H\mathbb{E}\Bigg\{\Bigg[diag\left(\mathbf{g}_{BE}^{(\mathrm{LoS})}\right)^*\Big(\mathbf{g}_{AE}^{(\mathrm{NLoS})}\Big)^*+diag\left(\mathbf{g}_{AE}^{(\mathrm{LoS})}\right)^*\Big(\mathbf{g}_{BE}^{(\mathrm{NLoS})}\Big)^*\\
&+\left(\mathbf{g}_{BE}^{(\mathrm{NLoS})}\odot\mathbf{g}_{AE}^{(\mathrm{NLoS})}\right)^*\Bigg]\cdot\Bigg[\Big(\mathbf{g}_{AE}^{(\mathrm{NLoS})}\Big)^T diag\left(\mathbf{g}_{BE}^{(\mathrm{LoS})}\right)\\
&+\Big(\mathbf{g}_{BE}^{(\mathrm{NLoS})}\Big)^T diag\left(\mathbf{g}_{AE}^{(\mathrm{LoS})}\right)+\left(\mathbf{g}_{AE}^{(\mathrm{NLoS})}\odot\mathbf{g}_{BE}^{(\mathrm{NLoS})}\right)^T\Bigg]\Bigg\}\mathbf{w}\\
=&\mathbf{w}^H\Bigg\{diag\left(\mathbf{g}_{BE}^{(\mathrm{LoS})}\right)^*\mathbb{E}\left[\Big(\mathbf{g}_{AE}^{(\mathrm{NLoS})}\Big)^*\Big(\mathbf{g}_{AE}^{(\mathrm{NLoS})}\Big)^T\right]diag\left(\mathbf{g}_{BE}^{(\mathrm{LoS})}\right)\\
&+diag\left(\mathbf{g}_{AE}^{(\mathrm{LoS})}\right)^*\mathbb{E}\left[\Big(\mathbf{g}_{BE}^{(\mathrm{NLoS})}\Big)^*\Big(\mathbf{g}_{BE}^{(\mathrm{NLoS})}\Big)^T\right]diag\left(\mathbf{g}_{AE}^{(\mathrm{LoS})}\right)\\
&+\mathbb{E}\left[\Big(\mathbf{g}_{BE}^{(\mathrm{NLoS})}\odot\mathbf{g}_{AE}^{(\mathrm{NLoS})}\Big)^*\cdot\Big(\mathbf{g}_{AE}^{(\mathrm{NLoS})}\odot\mathbf{g}_{BE}^{(\mathrm{NLoS})}\Big)^T\right]\Bigg\}\mathbf{w}\\
=&\mathbf{w}^H\Bigg\{2diag\left(\mathbf{g}_{BE}^{(\mathrm{LoS})}\right)^*\mathbf{\Sigma}_{AE}diag\left(\mathbf{g}_{BE}^{(\mathrm{LoS})}\right)\\
&+2diag\left(\mathbf{g}_{AE}^{(\mathrm{LoS})}\right)^*\mathbf{\Sigma}_{BE}diag\left(\mathbf{g}_{AE}^{(\mathrm{LoS})}\right)+4\mathbf{\Sigma}_{AE}\odot\mathbf{\Sigma}_{BE}\Bigg\}\mathbf{w},
\end{aligned}
\tag{34}
$$

which, divided by 2, gives the result in Eq. (6) with fixed $\mathbf{w}$.

Appendix B
Deduction of Theoretical Key Match Rate

The theoretical key match rate in Eq. (8) is deduced by:

$$
\begin{aligned}
&Pr\{k_A=k_E\}\\
=&Pr\left(z_A>\gamma_1^{(\mathrm{A})},z_E>\gamma_1^{(\mathrm{E})}\right)+Pr\left(z_A<\gamma_0^{(\mathrm{A})},z_E<\gamma_0^{(\mathrm{E})}\right)\\
=&\iint\limits_{\substack{\{\upsilon>\gamma_1^{(\mathrm{A})},\zeta>\gamma_1^{(\mathrm{E})}\}\\ \bigcup\{\upsilon<\gamma_0^{(\mathrm{A})},\zeta<\gamma_0^{(\mathrm{E})}\}}} p_{z_A|z_E}(\upsilon|\zeta)\cdot p_{z_E}(\zeta)d\upsilon d\zeta\\
=&\iint\limits_{\substack{\{\zeta+\xi>\gamma_1^{(\mathrm{A})},\zeta>\gamma_1^{(\mathrm{E})}\}\\ \bigcup\{\zeta+\xi<\gamma_0^{(\mathrm{A})},\zeta<\gamma_0^{(\mathrm{E})}\}}} p_z(\xi)\cdot p_{z_E}(\zeta)d\xi d\zeta\\
=&\iint\limits_{\substack{\{\zeta+\xi>\gamma_1^{(\mathrm{A})},\zeta>\gamma_1^{(\mathrm{E})}\}\\ \bigcup\{\zeta+\xi<\gamma_0^{(\mathrm{A})},\zeta<\gamma_0^{(\mathrm{E})}\}}} \mathcal{N}(\xi,0,\sigma_h^2)\cdot\mathcal{N}(\zeta,0,\sigma_E^2)d\xi d\zeta\\
=&\frac{1}{\sqrt{2\pi}}\int_{\frac{\gamma_1^{(\mathrm{E})}}{\sigma_E}}^{+\infty}\Phi\left(\frac{-\gamma_1^{(\mathrm{A})}+\sigma_E\zeta}{\sigma_h}\right)\exp\left(-\frac{\zeta^2}{2}\right)d\zeta\\
&+\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\frac{\gamma_0^{(\mathrm{E})}}{\sigma_E}}\Phi\left(\frac{\gamma_0^{(\mathrm{A})}-\sigma_E\zeta}{\sigma_h}\right)\exp\left(-\frac{\zeta^2}{2}\right)d\zeta\\
=&\sqrt{\frac{2}{\pi}}\int_{\beta}^{+\infty}\Phi\left(-\beta\sqrt{\frac{\sigma_E^2}{\sigma_h^2}+1}+\frac{\sigma_E}{\sigma_h}\zeta\right)\exp\left(-\frac{\zeta^2}{2}\right)d\zeta
\end{aligned}
\tag{35}
$$

where $p_{z_A|z_E}(\upsilon|\zeta)$ is the PDF of z_A conditioned on z_E, and $p_{z_E}(\zeta)$ is the PDF of z_E. $\mathcal{N}(\xi,0,\sigma^2)$ is the real Gaussian PDF of ξ with 0 expectation and σ^2 as variance.

With $x\triangleq\sigma_E/\sigma_h$, the first-order derivative showing its monotonically increasing property is:

$$
\begin{aligned}
&\frac{\partial Pr\{k_A=k_E\}}{\partial x}\\
=&\frac{1}{\pi}\int_{\beta}^{+\infty}\exp\left(-\frac{\left(x\zeta-\beta\sqrt{x^2+1}\right)^2+\zeta^2}{2}\right)\left(\zeta-\frac{\beta}{\sqrt{1+\frac{1}{x^2}}}\right)d\zeta\overset{(a)}{>}0
\end{aligned}
\tag{36}
$$

where (a) is because (i) $\exp(-((x\zeta-\beta\sqrt{x^2+1})^2+\zeta^2)/2)>0$, and (ii) $\zeta-\beta/\sqrt{1+1/x^2}$ takes its minimum value 0 when $x\to+\infty$ and $\zeta=\beta$.

Next, we prove $Pr\{k_A=k_E\}$ at first increases quickly and then gradually, with respect to σ_E^2/σ_h^2. This is done by evaluating its second-order derivative. As we denote $f_1(x,\zeta)\triangleq\exp(-((x\zeta-\beta\sqrt{x^2+1})^2+\zeta^2)/2)$ and $f_2(x,\zeta)\triangleq(\zeta-\beta/\sqrt{1+1/x^2})$, the second-order derivative is:

$$
\begin{aligned}
&\frac{\partial^2 Pr\{k_A=k_E\}}{\partial x^2}=\frac{1}{\pi}\int_{\beta}^{+\infty}\frac{\partial\left[f_1(x,\zeta)\cdot f_2(x,\zeta)\right]}{\partial x}d\zeta\\
=&\frac{1}{\pi}\int_{\beta}^{+\infty}-f_1(x,\zeta)f_2^2(x,\zeta)\Big(x\zeta-\beta\sqrt{x^2+1}\Big)-\frac{\beta f_1(x,\zeta)}{(1+x^2)^{\frac{3}{2}}}d\zeta<0.
\end{aligned}
\tag{37}
$$

Appendix C
Equivalency proof between Eqs. (22)-(23)

We assume that an optimal $\mathbf{w}_{opt}$ that maximizes the objective function $\mathbf{w}_{opt}^H\mathbf{G}\mathbf{w}_{opt}$, has a set of indices, i.e.,

$$
\mathcal{S}=\left\{i|\mathbf{w}_{opt}^H\mathbf{E}_i\mathbf{w}_{opt}=|w_{opt,i}|^2<A_E,\forall i\in\{1,\cdots,M\}\right\},
\tag{38}
$$

whose powers are less than A_E. Then, we create a better $\mathbf{w}$ with larger objective value, and ensure all $|w_i|^2=A_E$. This is done by:

$$
\mathbf{w}=\mathbf{w}_{opt}+\boldsymbol{\varrho}\odot\left(\mathbf{G}\mathbf{w}_{opt}\right),
\tag{39}
$$

$$
\varrho_i=\begin{cases}\frac{\sqrt{\varsigma^2+(A_E-|w_{opt,i}|^2)|\mathbf{G}_{i,:}\mathbf{w}_{opt}|^2}-\varsigma}{|\mathbf{G}_{i,:}\mathbf{w}_{opt}|^2} & i\in\mathcal{S}\\ 0 & i\notin\mathcal{S}\end{cases}
\tag{40}
$$

where $\varsigma\triangleq Re[w_{opt,i}^*\mathbf{G}_{i,:}\mathbf{w}_{opt}]$, and $\mathbf{G}_{i,:}$ is the ith row of $\mathbf{G}$. Here, for the elements in $\mathbf{w}_{opt}$ already having full power A_E, i.e., $\forall i\notin\mathcal{S}$, we assign $\varrho_i=0$ to make $w_i=w_{opt,i}$ and therefore $|w_i|^2=\mathbf{w}^H\mathbf{E}_i\mathbf{w}=\mathbf{w}_{opt}^H\mathbf{E}_i\mathbf{w}_{opt}=|w_{opt,i}|^2=A_E$. For the elements in $\mathbf{w}_{opt}$ with less power than A_E, i.e., $\forall i\in\mathcal{S}$, ϱ_i is assigned as the positive solution to function $|w_{opt,i}+\varrho_i\mathbf{G}_{i,:}\mathbf{w}_{opt}|^2=A_E$ (positive solution exists given $A_E-|w_{opt,i}|^2>0$). This therefore makes $|w_i|^2=|w_{opt,i}+\varrho_i\mathbf{G}_{i,:}\mathbf{w}_{opt}|^2=A_E$, i.e., $\mathbf{w}^H\mathbf{E}_i\mathbf{w}=A_E$.

Next, we show the constructed $\mathbf{w}$ has a larger objective value, i.e., $\mathbf{w}^H\mathbf{G}\mathbf{w}>\mathbf{w}_{opt}^H\mathbf{G}\mathbf{w}_{opt}$. This is because:

$$
\begin{aligned}
\mathbf{w}^H\mathbf{G}\mathbf{w}=&\mathbf{w}_{opt}^H\mathbf{G}\mathbf{w}_{opt}+\left[\boldsymbol{\varrho}\odot\left(\mathbf{G}\mathbf{w}_{opt}\right)\right]^H\mathbf{G}\left[\boldsymbol{\varrho}\odot\left(\mathbf{G}\mathbf{w}_{opt}\right)\right]\\
&+2Re\left\{\left[\boldsymbol{\varrho}\odot\left(\mathbf{G}\mathbf{w}_{opt}\right)\right]^H\mathbf{G}\mathbf{w}_{opt}\right\}\\
\overset{(a)}{>}&\mathbf{w}_{opt}^H\mathbf{G}\mathbf{w}_{opt}+2Re\left\{\left[\boldsymbol{\varrho}\odot\left(\mathbf{G}\mathbf{w}_{opt}\right)\right]^H\mathbf{G}\mathbf{w}_{opt}\right\}\overset{(b)}{>}\mathbf{w}_{opt}^H\mathbf{G}\mathbf{w}_{opt}
\end{aligned}
\tag{41}
$$

where (a) is because $\mathbf{G}$ (the covariance matrix) is positive definite. (b) is because $[\boldsymbol{\varrho}\odot(\mathbf{G}\mathbf{w}_{opt})]^H\mathbf{G}\mathbf{w}_{opt}=$

$(\mathbf{G}\mathbf{w}_{opt})^H diag(\boldsymbol{\varrho})\mathbf{G}\mathbf{w}_{opt} = \boldsymbol{\varrho}^T diag(\mathbf{G}\mathbf{w}_{opt})^* \mathbf{G}\mathbf{w}_{opt} = \boldsymbol{\varrho}^T[(\mathbf{G}\mathbf{w}_{opt})^* \odot (\mathbf{G}\mathbf{w}_{opt})] > 0$, given elements in $\boldsymbol{\varrho}$ and in $(\mathbf{G}\mathbf{w}_{opt})^* \odot (\mathbf{G}\mathbf{w}_{opt})$ are real and no less than 0. As such, the maximal value of the objective function under constraints from Eq. (23), takes at the bound of Eq. (22), therefore making Eq. (22) and Eq. (23) equivalent.

REFERENCES

[1] H.-M. Wang, X. Zhang, and J.-C. Jiang, "UAV-Involved Wireless Physical-Layer Secure Communications: Overview and Research Directions," *IEEE Wireless Communications*, vol. 26, no. 5, pp. 32–39, 2019.

[2] W. Wang, X. Liu, J. Tang, N. Zhao, Y. Chen, Z. Ding, and X. Wang, "Beamforming and Jamming Optimization for IRS-Aided Secure NOMA Networks," *IEEE Transactions on Wireless Communications*, vol. 21, no. 3, pp. 1557–1569, 2022.

[3] X. Pang, N. Zhao, J. Tang, C. Wu, D. Niyato, and K.-K. Wong, "IRS-Assisted Secure UAV Transmission via Joint Trajectory and Beamforming Design," *IEEE Transactions on Communications*, vol. 70, no. 2, pp. 1140–1152, 2022.

[4] Y. Zhou, P. L. Yeoh, H. Chen, Y. Li, R. Schober, L. Zhuo, and B. Vucetic, "Improving Physical Layer Security via a UAV Friendly Jammer for Unknown Eavesdropper Location," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 11, pp. 11 280–11 284, 2018.

[5] A. Mukherjee, S. A. A. Fakoorian, J. Huang, and A. L. Swindlehurst, "Principles of Physical Layer Security in Multiuser Wireless Networks: A Survey," *IEEE Communications Surveys Tutorials*, vol. 16, no. 3, pp. 1550–1573, 2014.

[6] J. Zhang, T. Q. Duong, A. Marshall, and R. Woods, "Key Generation From Wireless Channels: A Review," *IEEE Access*, vol. 4, pp. 614–626, 2016.

[7] C. Ye, A. Reznik, and Y. Shah, "Extracting Secrecy from Jointly Gaussian Random Variables," in *2006 IEEE International Symposium on Information Theory*, 2006, pp. 2593–2597.

[8] C. Ye, S. Mathur, A. Reznik, Y. Shah, W. Trappe, and N. B. Mandayam, "Information-Theoretically Secret Key Generation for Fading Wireless Channels," *IEEE Transactions on Information Forensics and Security*, vol. 5, no. 2, pp. 240–254, 2010.

[9] X. Wu, Y. Song, C. Zhao, and X. You, "Secrecy extraction from correlated fading channels: An upper bound," in *2009 International Conference on Wireless Communications & Signal Processing*, 2009, pp. 1–3.

[10] J. Li, P. Wang, L. Jiao, Z. Yan, K. Zeng, and Y. Yang, "Security analysis of triangle channel-based physical layer key generation in wireless backscatter communications," *IEEE Transactions on Information Forensics and Security*, vol. 18, pp. 948–964, 2023.

[11] Y. Peng, P. Wang, W. Xiang, and Y. Li, "Secret Key Generation Based on Estimated Channel State Information for TDD-OFDM Systems Over Fading Channels," *IEEE Transactions on Wireless Communications*, vol. 16, no. 8, pp. 5176–5186, 2017.

[12] K. Moara-Nkwe, Q. Shi, G. M. Lee, and M. H. Eiza, "A Novel Physical Layer Secure Key Generation and Refreshment Scheme for Wireless Sensor Networks," *IEEE Access*, vol. 6, pp. 11 374–11 387, 2018.

[13] S. N. Premnath, S. Jana, J. Croft, P. L. Gowda, M. Clark, S. K. Kasera, N. Patwari, and S. V. Krishnamurthy, "Secret Key Extraction from Wireless Signal Strength in Real Environments," *IEEE Transactions on Mobile Computing*, vol. 12, no. 5, pp. 917–930, 2013.

[14] S. Mathur, W. Trappe, N. Mandayam, C. Ye, and A. Reznik, "Radio-Telepathy: Extracting a Secret Key from an Unauthenticated Wireless Channel," in *Proceedings of the 14th ACM international conference on Mobile computing and networking*, 2008, pp. 128–139.

[15] G. Brassard and L. Salvail, "Secret-Key Reconciliation by Public Discussion," in *Advances in Cryptology — EUROCRYPT '93*, T. Helleseth, Ed. Berlin, Heidelberg: Springer Berlin Heidelberg, 1994, pp. 410–423.

[16] R. Impagliazzo, L. A. Levin, and M. Luby, "Pseudo-Random Generation from One-Way Functions," in *Proceedings of the twenty-first annual ACM symposium on Theory of computing*, 1989, pp. 12–24.

[17] N. Aldaghri and H. Mahdavifar, "Physical Layer Secret Key Generation in Static Environments," *IEEE Transactions on Information Forensics and Security*, vol. 15, pp. 2692–2705, 2020.

[18] X. Lu, J. Lei, and W. Li, "Secret Key Generation based on Signal Power Allocation Optimisation," *IET Communications*, vol. 16, no. 14, pp. 1724–1730, 2022.

[19] G. Li, A. Hu, J. Zhang, and B. Xiao, "Security Analysis of a Novel Artificial Randomness Approach for Fast Key Generation," in *GLOBECOM 2017 - 2017 IEEE Global Communications Conference*, 2017, pp. 1–6.

[20] Y. Lou, L. Jin, Z. Zhong, K. Huang, and S. Zhang, "Secret Key Generation Scheme based on MIMO Received Signal Spaces," *Scientia Sinica Informationis*, vol. 47, no. 3, pp. 362–373, 2017.

[21] H. Taha and E. Alsusa, "Secret Key Exchange Using Private Random Precoding in MIMO FDD and TDD Systems," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 6, pp. 4823–4833, 2017.

[22] A. Khisti, "Secret-Key Agreement over Non-Coherent Block-Fading Channels with Public Discussion," *IEEE Transactions on Information Theory*, vol. 62, no. 12, pp. 7164–7178, 2016.

[23] S. Sharifian, F. Lin, and R. Safavi-Naini, "Secret key agreement using a virtual wiretap channel," in *IEEE INFOCOM 2017 - IEEE Conference on Computer Communications*, 2017, pp. 1–9.

[24] S. Zhang, L. Jin, Y. Lou, and Z. Zhong, "Secret Key Generation based on Two-Way Randomness for TDD-SISO System," *China Communications*, vol. 15, no. 7, pp. 202–216, 2018.

[25] G. Wunder, R. Fritschek, and K. Reaz, "RECiP: Wireless Channel Reciprocity Restoration Method for Varying Transmission Power," in *2016 IEEE 27th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, 2016, pp. 1–5.

[26] Q. Wu, S. Zhang, B. Zheng, C. You, and R. Zhang, "Intelligent Reflecting Surface-Aided Wireless Communications: A Tutorial," *IEEE Transactions on Communications*, vol. 69, no. 5, pp. 3313–3351, 2021.

[27] Q. Wu and R. Zhang, "Towards smart and reconfigurable environment: Intelligent reflecting surface aided wireless network," *IEEE Communications Magazine*, vol. 58, no. 1, pp. 106–112, 2020.

[28] ——, "Intelligent Reflecting Surface Enhanced Wireless Network via Joint Active and Passive Beamforming," *IEEE Transactions on Wireless Communications*, vol. 18, no. 11, pp. 5394–5409, 2019.

[29] Q. Ma, L. Chen, H. B. Jing, Q. R. Hong, H. Y. Cui, Y. Liu, L. Li, and T. J. Cui, "Controllable and Programmable Nonreciprocity based on Detachable Digital Coding Metasurface," *Advanced Optical Materials*, vol. 7, no. 24, p. 1901285, 2019.

[30] W. Jiang, B. Chen, J. Zhao, Z. Xiong, and Z. Ding, "Joint Active and Passive Beamforming Design for the IRS-Assisted MIMOME-OFDM Secure Communications," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 10, pp. 10 369–10 381, 2021.

[31] S. Hong, C. Pan, H. Ren, K. Wang, and A. Nallanathan, "Artificial-Noise-Aided Secure MIMO Wireless Communications via Intelligent Reflecting Surface," *IEEE Transactions on Communications*, vol. 68, no. 12, pp. 7851–7866, 2020.

[32] X. Hu, L. Jin, K. Huang, X. Sun, Y. Zhou, and J. Qu, "Intelligent Reflecting Surface-Assisted Secret Key Generation With Discrete Phase Shifts in Static Environment," *IEEE Wireless Communications Letters*, vol. 10, no. 9, pp. 1867–1870, 2021.

[33] X. Lu, J. Lei, Y. Shi, and W. Li, "Intelligent Reflecting Surface Assisted Secret Key Generation," *IEEE Signal Processing Letters*, vol. 28, pp. 1036–1040, 2021.

[34] Z. Ji, P. L. Yeoh, G. Chen, C. Pan, Y. Zhang, Z. He, H. Yin, and Y. Li, "Random Shifting Intelligent Reflecting Surface for OTP Encrypted Data Transmission," *IEEE Wireless Communications Letters*, vol. 10, no. 6, pp. 1192–1196, 2021.

[35] P. Staat, H. Elders-Boll, M. Heinrichs, R. Kronberger, C. Zenger, and C. Paar, "Intelligent Reflecting Surface-Assisted Wireless Key Generation for Low-Entropy Environments," in *2021 IEEE 32nd Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, 2021, pp. 745–751.

[36] Z. Ji, P. L. Yeoh, D. Zhang, G. Chen, Y. Zhang, Z. He, H. Yin, and Y. li, "Secret Key Generation for Intelligent Reflecting Surface Assisted Wireless Communication Networks," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 1, pp. 1030–1034, 2021.

[37] L. Hu, G. Li, H. Luo, and A. Hu, "On the RIS Manipulating Attack and Its Countermeasures in Physical-layer Key Generation," in *2021 IEEE 94th Vehicular Technology Conference (VTC2021-Fall)*, 2021, pp. 1–5.

[38] G. Li, L. Hu, P. Staat, H. Elders-Boll, C. Zenger, C. Paar, and A. Hu, "Reconfigurable Intelligent Surface for Physical Layer Key Generation: Constructive or Destructive?" *IEEE Wireless Communications*, vol. 29, no. 4, pp. 146–153, 2022.

[39] C. D. T. Thai, J. Lee, and T. Q. Quek, "Physical-Layer Secret Key Generation with Colluding Untrusted Relays," *IEEE Transactions on Wireless Communications*, vol. 15, no. 2, pp. 1517–1530, 2015.

[40] M. Letafati, A. Kuhestani, D. W. K. Ng, and H. Behroozi, "A New Frequency Hopping-Aided Secure Communication in the Presence of an Adversary Jammer and an Untrusted Relay," in *2020 IEEE International*

Conference on Communications Workshops (ICC Workshops). IEEE, 2020, pp. 1–7.

[41] A. Taha, M. Alrabeiah, and A. Alkhateeb, "Enabling Large Intelligent Surfaces with Compressive Sensing and Deep Learning," *IEEE Access*, vol. 9, pp. 44 304–44 321, 2021.

[42] A. Alkhateeb, O. El Ayach, G. Leus, and R. W. Heath, "Channel Estimation and Hybrid Precoding for Millimeter Wave Cellular Systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 831–846, 2014.

[43] C. You and R. Zhang, "Hybrid Offline-Online Design for UAV-Enabled Data Harvesting in Probabilistic LoS Channels," *IEEE Transactions on Wireless Communications*, vol. 19, no. 6, pp. 3753–3768, 2020.

[44] E. Björnson and L. Sanguinetti, "Rayleigh Fading Modeling and Channel Hardening for Reconfigurable Intelligent Surfaces," *IEEE Wireless Communications Letters*, vol. 10, no. 4, pp. 830–834, 2021.

[45] O. Tsilipakos, A. C. Tasolamprou, A. Pitilakis, F. Liu, X. Wang, M. S. Mirmoosa, D. C. Tzarouchis, S. Abadal, H. Taghvaee, C. Liaskos *et al.*, "Toward Intelligent Metasurfaces: The Progress from Globally Tunable Metasurfaces to Software-Defined Metasurfaces with an Embedded Network of Controllers," *Advanced optical materials*, vol. 8, no. 17, p. 2000783, 2020.

[46] O. Özdogan, E. Björnson, and E. G. Larsson, "Intelligent Reflecting Surfaces: Physics, Propagation, and Pathloss Modeling," *IEEE Wireless Communications Letters*, vol. 9, no. 5, pp. 581–585, 2020.

[47] G. C. Alexandropoulos and E. Vlachos, "A Hardware Architecture for Reconfigurable Intelligent Surfaces with Minimal Active Elements for Explicit Channel Estimation," in *ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2020, pp. 9175–9179.

[48] E. J. Candes and T. Tao, "Decoding by Linear Programming," *IEEE Transactions on Information Theory*, vol. 51, no. 12, pp. 4203–4215, 2005.

[49] E. J. Candes, "The Restricted Isometry Property and its Implications for Compressed Sensing," *Comptes rendus mathematique*, vol. 346, no. 9-10, pp. 589–592, 2008.

[50] T. Zhang, "Sparse Recovery with Orthogonal Matching Pursuit under RIP," *IEEE Transactions on Information Theory*, vol. 57, no. 9, pp. 6215–6221, 2011.

[51] M. Letafati, H. Behroozi, B. H. Khalaj, and E. A. Jorswieck, "Deep Learning for Hardware-Impaired Wireless Secret Key Generation with Man-in-the-Middle Attacks," in *2021 IEEE Global Communications Conference (GLOBECOM)*. IEEE, 2021, pp. 1–6.

[52] S. Eberz, M. Strohmeier, M. Wilhelm, and I. Martinovic, "A Practical Man-in-the-Middle Attack on Signal-based Key Generation Protocols," in *European symposium on research in computer security*. Springer, 2012, pp. 235–252.

[53] Y. Pan, Z. Xu, M. Li, and L. Lazos, "Man-in-the-Middle Attack Resistant Secret Key Generation via Channel Randomization," in *Proceedings of the Twenty-second International Symposium on Theory, Algorithmic Foundations, and Protocol Design for Mobile Networks and Mobile Computing*, 2021, pp. 231–240.

[54] R. Jin and K. Zeng, "Manipulative Attack against Physical Layer Key Agreement and Countermeasure," *IEEE Transactions on Dependable and Secure Computing*, vol. 18, no. 1, pp. 475–489, 2021.

[55] L. Jin, S. Zhang, Y. Lou, X. Xu, and Z. Zhong, "Secret Key Generation With Cross Multiplication of Two-Way Random Signals," *IEEE Access*, vol. 7, pp. 113 065–113 080, 2019.

[56] Z. Luo, W. Ma, A. M. So, Y. Ye, and S. Zhang, "Semidefinite relaxation of quadratic optimization problems," *IEEE Signal Processing Magazine*, vol. 27, no. 3, pp. 20–34, 2010.

[57] A. So, J. Zhang, and Y. Ye, "On approximating complex quadratic optimization problems via semidefinite programming relaxations," *Mathematical Programming*, vol. 110, no. 1, pp. 93–110, 2007.

[58] X. Zhou, B. Maham, and A. Hjorungnes, "Pilot Contamination for Active Eavesdropping," *IEEE Transactions on Wireless Communications*, vol. 11, no. 3, pp. 903–907, 2012.

[59] S. Im, H. Jeon, J. Choi, and J. Ha, "Secret Key Agreement with Large Antenna Arrays under the Pilot Contamination Attack," *IEEE Transactions on Wireless Communications*, vol. 14, no. 12, pp. 6579–6594, 2015.

[60] P. Billingsley, "Probability and Measure. 3rd Wiley," *New York*, 1995.

2023-04-12

Adversarial reconfigurable intelligent surface against physical layer key generation

Wei, Zhuangkun

IEEE

Wei Z, Li B, Guo W. (2023) Adversarial reconfigurable intelligent surface against physical layer key generation. IEEE Transactions on Information Forensics and Security, Volume 18, 2023, pp. 2368-2381

https://doi.org/10.1109/TIFS.2023.3266705