Design of Physical Layer Challenge Response Authentication With RIS and Multiple-Antenna Devices

Anna Valeria Guglielmi, *Member, IEEE*, Laura Crosara, *Member, IEEE*, and Stefano Tomasin, *Senior Member, IEEE*

***Abstract*—This paper focuses on the challenge-response physical-layer authentication (CR-PLA) scheme where a reflecting intelligent surface (RIS) is under the control of a receiving base station (BS) (Bob) who aims at checking if received messages come from a legitimate user equipment (UE) Alice or from an impersonating device (Trudy). To this end, Bob sets a random configuration of the RIS which remains secret to the attacker, and verifies that the channel estimated on the received message corresponds to the set configuration. We design the probability distribution of RIS configurations chosen by the verifier to maximize average capacity while satisfying an upper bound on missed detection (MD) probability for a given false alarm (FA) probability. The balance of communication and security metrics demonstrated by the numerical results shows the effectiveness and potential of the CR-PLA scheme.**

***Index Terms*—Challenge response (CR) authentication, physical layer authentication (PLA), reflecting intelligent surface (RIS).**

I. Introduction

PHYSICAL layer authentication (PLA) is the problem of establishing whether a message comes from the declared sender or an attacker, using the physical-layer signals. TAG-based physical layer authentication (PLA) [1], includes the *association* phase, where verifier (Bob) estimates the channel from the authentic source (Alice), and the *identification* phase, wherein for each new message, Bob estimates the channel and compares it with that obtained in the first phase. The message is deemed authentic if the two channels are similar enough. More recently, challenge response (CR)-PLA has been introduced to exploit the possibility of modifying the wireless channel by controlling the phase shifts of the elements of a reflecting intelligent surface (RIS). When the RIS is under Bob's control, he can set a random configuration of the RIS which remains secret to the attacker, and verify that the channel estimated on the received message corresponds to the set configuration, [2].

Since the RIS is also used for communication purposes, the random perturbations introduced by Bob are around the communication-optimal configuration, and the literature has considered only a uniform distribution of the perturbation. However, such a choice is not optimal, and in this paper we aim at investigating the security performance obtained with a general probability density function (PDF) of the perturbation, still taking into account also the communication performance.

We focus on a cellular uplink scenario, where a base station (BS) verifies a user equipment (UE). The procedure operates in two steps. First, the UE transmits authenticated pilots via the RIS using configurations chosen by the BS. The BS uses these to model the channel, enabling it to predict the channel for any other RIS configuration. Then, the BS selects a random new BS configuration while the UE transmits a message. Upon receiving the message, the BS estimates the actual channel and compares it with the predicted channel for that configuration to confirm authenticity.

To evaluate communication performance, we consider the achievable rate of the UE-RIS-BS channel. For authentication, we consider a likelihood test (LT) and analyze the security performance in terms of false alarm (FA) and missed detection (MD) probabilities. We derive an attack strategy based on the least squares (LS) estimate of the cascade channel when the attacker has partial knowledge of the legitimate channel, assuming correlation between the attacker's channel and those of the legitimate UE and RIS. We assume the attacker can transmit the attack signal either directly to the BS or through the RIS.

We design the PDF of the randomly selected RIS phase shifts to maximize an upper bound on the average achievable rate subject to an upper bound on the MD probability for a target FA probability. Such optimization was not available in the literature. To this end, we identify two key statistical parameters (represented by two real numbers) that capture the effects of the PDF on both the communication and the security metrics, reducing the PDF design problem to the optimization of these two parameters, under other constraints. We derive approximate expressions for the FA and MD probabilities, as well as the achievable rate, in a multi-antenna scenario for both the UE and the BS.

Received 5 June 2025; revised 4 February 2026 and 20 March 2026; accepted 20 March 2026. Date of publication 23 March 2026; date of current version 9 April 2026. This work was supported in part by the Project ISP5G+ under Grant CUP D33C22001300002; in part by the Security and Rights in the Cyber Space (SERICS) Program under the National Recovery and Resilience Plan (NRRP) Ministero dell'Università e della Ricerca (MUR) Program funded by European Union - Next Generation EU (EU-NGEU) under Grant PE00000014; in part by European Commission through the Horizon Europe/Joint Undertaking (JU) Smart Networks and Services (SNS) Project ROBUST-6G under Grant 101139068; and in part by European Union under Italian NRRP Mission 4, Component 2, Investment 1.3, partnership on "Telecommunications of the Future" (Program "RESTART") under Grant CUP C93C22005250001 and Grant PE00000001. The associate editor coordinating the review of this article and approving it for publication was Dr. Hossein Pishro-Nik. *(Corresponding author: Anna Valeria Guglielmi.)*

Anna Valeria Guglielmi and Laura Crosara are with the Department of Information Engineering, Università degli Studi di Padova, 35131 Padua, Italy (e-mail: annavaleria.guglielmi@unipd.it; laura.crosara@unipd.it).

Stefano Tomasin is with the Department of Information Engineering and the Department of Mathematics, University of Padova, 35131 Padua, Italy (e-mail: stefano.tomasin@unipd.it).

Digital Object Identifier 10.1109/TIFS.2026.3676952

The main contributions of this paper are as follows:

1) Achievable Rate Analysis: We derive an upper bound on the average achievable rate for communication in RIS-based CR-PLA systems, specifically considering multi-antenna devices.
2) False Alarm Formulation: We analytically derive the FA probability for the proposed RIS-based authentication mechanism.
3) Security Assessment: We establish an upper bound for the MD probability under a specific threat model where the attacker relies on a LS estimate of the average Alice-RIS-Bob channel.
4) Optimization of RIS Configuration: We derive the optimal PDF for the random selection of RIS configurations, specifically designed to balance communication performance and security metrics.
5) Parametric Analysis: We investigate the impact of key system parameters—such as the correlation factor, the number of RIS elements, and the operating signal-to-noise ratio (SNR) on the PDF design and the resulting communication-security trade-off.

The paper is structured as follows. We first review related works and position our contributions in the next subsection. Section II presents the system model along with the CR-PLA framework. Section III formulates the authentication test, while Section IV defines the relevant communication and security metrics. The optimization of the RIS phase-shift distribution is addressed in Section V. We analyze the numerical results for the RIS-supported CR-PLA scheme in Section VI and offer concluding remarks in Section VII.

A. Related Works

PLA has been studied for several technologies, including orthogonal frequency division multiplexing (OFDM) and multiple-input multiple-output (MIMO) [3], [4], underwater acoustic communications [5], and using several techniques for the test, from Neyman-Pearson [6] to machine learning approaches [7]. For an overview on PLA see also [8] and [9]. PLA has been extensively studied over the years, with comprehensive overviews available in [8] and [9]. In this paper, we focus on challenge response PLA (CR-PLA) leveraging RIS for authentication.

RIS has been used to improve TAG-based PLA, where the message is authenticated by checking that it has been transmitted over the same channel as the previous message. In this case, the RIS can be optimized to enhance reception from a specific location and improve the authentication mechanism, [10]. Similarly, a TAG-based PLA method employing RIS in non-line-of-sight vehicular networks was introduced in [11], demonstrating an improvement in the authentication accuracy. In [12], a RIS-supported TAG-based PLA mechanism was presented, where channel gain and noise features are extracted and combined with a private key to generate a tagged signal. However, these schemes differ from our work as they rely on secret bits rather than physical-layer features for authentication. Moreover, in our analysis, we also consider the communication performance, specifically how the random selection of the RIS configuration in CR-PLA affects the data rate. Note that advanced RISs with sensing capabilities can further enhance channel knowledge and authentication performance [13], although our model assumes a passive RIS that only introduces a phase shift on impinging signals before passive reflection.

Here we also leverage the concept of CR-PLA, first introduced in [2]: the verifier randomly alters the propagation environment, serving as *challenge*, and uses the estimated channel as *response* for authentication. A different approach, also called CR-PLA, has been proposed in [14], where the channel as a *response* is replaced by a confidential message sent by the device under verification. This scheme has been further explored in [15], where artificial noise is introduced, and in [16], where actuators issue random transmissions that are detected by other sensors. In [14], [15], and [16], the challenge is embedded in the transmitted signal rather than the physical environment. In contrast, we introduce the challenge by randomly configuring RIS controlled by the receiver. CR-PLA has also been applied in scenarios involving drone swarms, where the movement of trusted drones acting as receivers presents the challenge [17].

In a previous work, a CR-PLA scheme that takes advantage of the reconfigurability of an RIS under the control of the verifier [18] has been proposed, considering single-antenna devices and uniform random distribution of phase shifts. The results demonstrate reduced overhead compared to higher-layer cryptographic methods, along with a clear trade-off between authentication accuracy and communication performance. The same system model has also been used to explore attack strategies [19] and to optimize the randomness introduced by the RIS [20] in the single-antenna scenario. In contrast to [18], [19], and [20], in this paper we focus on the design of CR-PLA considering multi-antenna devices and looking for the distribution of the RIS phase shifts that maximizes the average capacity while satisfying an upper bound on the MD probability for a desired FA probability.

Finally, while this paper focuses on authentication, RIS technology can also enhance other security aspects, such as confidentiality [21], [22], [23], [24], [25], [26] and anti-jamming [27]. Indeed, recent works have explored RIS-assisted semantic anti-jamming for integrated networks [28] and advanced transceiver designs to improve multi-functional network performance [29]. However, while physical-layer security (PLS) and anti-jamming aim at ensuring confidentiality and availability, this paper specifically addresses PLA, leveraging RIS reconfigurability to provide a robust identity verification mechanism.

Notation: The operator $\text{vec}(\boldsymbol{X})$ converts any matrix $\boldsymbol{X}$ with C columns into a column vector, by indexing the matrix by rows (top to bottom and left to right), i.e., $(\text{vec}(\boldsymbol{X}))_i = (\boldsymbol{X})_{\lfloor \frac{i}{C} \rfloor, i \bmod C}$. Vector and matrix indices start from 0; thus, column vector $\boldsymbol{x}$ with N entries is $\boldsymbol{x} = [x_0, \ldots, x_{N-1}]^T$. $\mathbb{1}_{\text{cond}}$ is the indicator function, taking value 1 when condition cond is verified and 0 otherwise.

II. System Model

We consider the uplink cellular network scenario depicted in Fig. 1. A legitimate receiver Bob, the BS, authenticates messages from a legitimate transmitter Alice, the UE. We

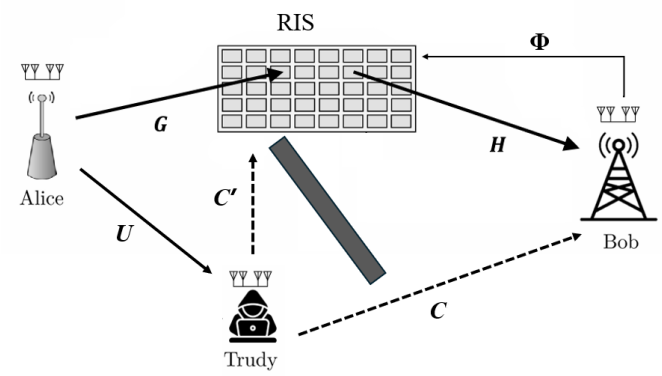

Fig. 1. Communication scenario.

account for the presence of an attacker device, Trudy, who aims to impersonate Alice by forging messages and transmitting them to Bob. We assume that Alice, Bob, and Trudy are equipped with uniform linear arrays (ULAs) of N_{A}, N_{B}, and N_{T} antennas, respectively.

We further assume that the communication between Alice and Bob is supported by an RIS with N reflecting elements, each acting as a receive and transmit antenna. In particular, the n-th element, $n = 0, 1, \ldots, N-1$, of the RIS introduces a phase shift $\phi_n = e^{j\theta_n}$ on the equivalent baseband signal and has unitary gain. We define vector $\boldsymbol{\phi} = [e^{j\theta_0}, e^{j\theta_1}, \ldots, e^{j\theta_{N-1}}]^{\mathrm{T}}$ as the *configuration* that collects the phase shifts introduced by the RIS on each element, and the $N \times N$ diagonal matrix

$$\boldsymbol{\Phi} = \mathrm{diag}\{\boldsymbol{\phi}\}. \tag{1}$$

Bob controls the RIS and sets its configuration $\boldsymbol{\Phi}$ using a secure dedicated channel (between Bob and the RIS) not accessible to Trudy. We assume that the communication between Alice and Bob only happens through the RIS without any additional direct link (because of obstruction). We denote the baseband equivalent matrix for the channel from Alice to the RIS as $\boldsymbol{G} \in \mathbb{C}^{N \times N_{\mathrm{A}}}$, the channel from the RIS to Bob as $\boldsymbol{H} \in \mathbb{C}^{N_{\mathrm{B}} \times N}$. Thus, the resulting Alice-RIS-Bob cascaded channel is

$$\boldsymbol{Q}_{\mathrm{AIB}} = \boldsymbol{H}\boldsymbol{\Phi}\boldsymbol{G}. \tag{2}$$

We assume that Trudy can choose to perform her attack either by transmitting directly to Bob through a full-rank channel $\boldsymbol{C}$, or through the RIS via the channel $\boldsymbol{C}'$. Both channel matrices are assumed to be full-rank and known to Trudy. Note that all channels are time-invariant and reciprocal, while the RIS configuration identified by matrix $\boldsymbol{\Phi}$ is under Bob's control and can be changed over time, making the cascaded channel controllable. We consider correlated Rayleigh fading channels; thus, all entries of the channel matrices $\boldsymbol{H}$ and $\boldsymbol{G}$ are independent zero-mean complex Gaussian with unitary power. However, most of the obtained derivations could be applied to a more general channel model.

We assume that Trudy has as many antennas as the RIS elements ($N_{\mathrm{T}} = N$) and channels ($\boldsymbol{U}$ and $\boldsymbol{C}$ from Fig. 1) are correlated to the channels $\boldsymbol{G}$ and $\boldsymbol{H}$, respectively. Specifically, Trudy is assumed to have partial information on $\boldsymbol{G}$ through $\boldsymbol{U}$, that is [18]

$$\boldsymbol{U} = \rho\boldsymbol{G} + \sqrt{1-\rho^2}\boldsymbol{D}, \tag{3}$$

where $\rho \in [0, 1]$ is the *correlation factor* between channels $\boldsymbol{U}$ and $\boldsymbol{G}$; the entries of $\boldsymbol{D}$ are circularly symmetric complex Gaussian (CSCG) independent of those of both $\boldsymbol{G}$ and $\boldsymbol{H}$.

A. Communication-Optimal RIS Configuration

Since the RIS also serves for communication purposes (beyond security purposes) between Alice and Bob, Bob should configure it with this functionality in mind.

We define the RIS configuration that maximizes the spectral efficiency (referred to as *communication optimal RIS configuration*) as

$$\overline{\boldsymbol{\Phi}} = \mathrm{diag}\{[e^{j\bar{\theta}_0}, \ldots, e^{j\bar{\theta}_{N-1}}]^T\}, \tag{4}$$

where $\bar{\theta}_n$, $n = 0, \ldots, N-1$, are the communication-optimal phase shifts of the RIS. Recently, several works in the literature propose solutions for the optimization of the RISs configuration (see the survey [30] and recent papers [31], [32], [33]). As RIS optimization is beyond the scope of this work, we refer to these papers for the computation of $\{\bar{\theta}_n\}$.

B. CR-PLA Mechanism

In the following, we will mainly focus on the scenario where Trudy transmits directly to Bob. This choice is justified by the fact that the through RIS attack represents a degenerate case for the CR-PLA mechanism (as we will discuss later). In this scenario, the random phase challenges affect the legitimate and forged signals identically, causing the performance to revert to that of a conventional TAG-based PLA. Therefore, the direct attack provides a more significant scenario to evaluate the security gains enabled by the proposed authentication scheme.

The CR-PLA mechanism introduced in [2] works as follows. In the *association* phase, Alice transmits several authenticated pilot signals, and Bob estimates the cascaded channel $\boldsymbol{Q}_{\mathrm{AIB}}$ for various RIS configurations. We assume that this phase is long enough to let Bob obtain estimates $\overline{\boldsymbol{Q}}_{\mathrm{AIB}}(\boldsymbol{\Phi})$ of the cascaded channel for *any* RIS configuration $\boldsymbol{\Phi}$

$$\overline{\boldsymbol{Q}}_{\mathrm{AIB}}(\boldsymbol{\Phi}) = \boldsymbol{Q}_{\mathrm{AIB}}(\boldsymbol{\Phi}) + \overline{\boldsymbol{W}}, \tag{5}$$

where $\overline{\boldsymbol{W}}$ is the estimation error at Bob, modeled as an additive white Gaussian noise (AWGN) matrix with independent entries, each having zero mean and power σ_{B}^2.

In the *verification* phase, Bob sets a random configuration $\boldsymbol{\Phi}'$ of the RIS (that constitutes the *challenge*). Specifically, Bob modifies the RIS configuration around the communication-optimal one $\overline{\boldsymbol{\Phi}}$. In particular, $\boldsymbol{\Phi}'$ is a diagonal matrix such that

$$[\boldsymbol{\Phi}']_{n,n} = e^{j\theta'_n} = e^{j(\bar{\theta}_n + \epsilon_n)}, \quad n = 0, \ldots, N-1, \tag{6}$$

where $\epsilon_n \in (-\pi, \pi]$ is a random variable. We assume that ϵ_n, $n = 0, \ldots, N-1$, are generated by Bob as independent and identically distributed (i.i.d.) for all the RIS elements. Their PDF $p_\epsilon(a)$, better specified in the following, is assumed to be an even function of the phase a. Whenever Bob receives a message, he estimates the cascaded channel $\boldsymbol{R}$ and checks if it is consistent with the expected channel $\overline{\boldsymbol{Q}}_{\mathrm{AIB}}(\boldsymbol{\Phi}')$ as discussed in Section III. We assume that all parties know the pilot signals used for channel estimation.

Under legitimate conditions when Alice transmits, given the configuration $\boldsymbol{\Phi}'$, Bob obtains the channel estimate $\boldsymbol{R} = \hat{\boldsymbol{Q}}_{\text{AIB}}$ with

$$\hat{\boldsymbol{Q}}_{\text{AIB}} = \boldsymbol{Q}_{\text{AIB}}(\boldsymbol{\Phi}') + \boldsymbol{W}_{\text{B}}\,, \tag{7}$$

where $\boldsymbol{W}_{\text{B}}$ denotes the estimation error at Bob, modeled as AWGN with zero mean matrix, with independent entries, each having power σ_{B}^2. Note that Trudy does not know the current RIS configuration; she does not know the channel condition when Alice is transmitting. This will enable the authentication process.

C. Attacker Model

The purpose of Trudy is to impersonate Alice, i.e., to pass the authentication check at the legitimate receiver. Trudy is assumed to perfectly know the RIS-Bob channel $\boldsymbol{H}$. Although this is a generous assumption to Trudy, it can be motivated by assuming that Trudy is located near Bob, making it more reasonable for her to properly estimate the RIS-Bob channel. Finally, we assume that Trudy knows both the pilot transmitted by Alice and the PDF of noise at the legitimate receiver.

By transmitting precoded pilot signals, Trudy induces the channel estimate $\boldsymbol{V}_0$ (apart from the presence of noise) at Bob. Thus, under attack, Bob estimates the channel $\boldsymbol{R} = \boldsymbol{V}$, with

$$\boldsymbol{V} = \boldsymbol{V}_0 + \boldsymbol{W}_{\text{B}}\,, \tag{8}$$

with $\boldsymbol{W}_{\text{B}}$ independent of $\boldsymbol{V}_0$, with statistics as in (7). We consider the Trudy-RIS channel matrix $\boldsymbol{C}$ to be full-rank, which can easily be obtained when Trudy has many antennas. In these conditions, Trudy can precode the pilot signals transmitted to Bob with a suitable precoding matrix $\boldsymbol{B} = \boldsymbol{C}^{-1}\boldsymbol{V}$, so that the equivalent channel (the cascaded of the precoder and $\boldsymbol{C}$) seen by Bob when Trudy transmits directly to Bob is $\boldsymbol{V}$.

Remark: Trudy's perfect knowledge of $\boldsymbol{H}$ simplifies the security analysis, although it may not always be realistic. Still, this represents a conservative worst-case scenario from the defender's standpoint, thus providing bounds on the achievable security level. For an imperfect estimate $\hat{\boldsymbol{H}}$ of $\boldsymbol{H}$, the forged cascaded channel would be affected by error due to this misalignment. Note also that we assume a static channel model, where channel realizations remain constant over the authentication interval, so that once an estimate of $\boldsymbol{H}$ is obtained, it remains valid for the duration of the attack.

III. Authentication Test

Bob aims to figure out whether the estimated channel $\boldsymbol{R}$ is authentic or forged by using his knowledge of $\overline{\boldsymbol{Q}}_{\text{AIB}}(\boldsymbol{\Phi}')$.

From the previously described model, the channel estimate at Bob is

$$\boldsymbol{R} = \begin{cases} \hat{\boldsymbol{Q}}_{\text{AIB}}(\boldsymbol{\Phi}') & \text{if Alice is transmitting } (b=0)\,, \\ \boldsymbol{V} & \text{if Trudy is transmitting } (b=1)\,, \end{cases} \tag{9}$$

with b indicating the effective legitimate/attack state. Bob performs an authentication test, wherein, given $\overline{\boldsymbol{Q}}_{\text{AIB}}(\boldsymbol{\Phi}')$ and $\boldsymbol{R}$, he decides between two hypotheses: $\hat{b}=0$ if the message is from Alice, and $\hat{b}=1$ if the message is from Trudy. Therefore, the authentication procedure outputs a Boolean value $\hat{b}$, which is an estimate of b, and correct verification is achieved if $\hat{b}=b$.

Since Trudy can perform a variety of attacks $\boldsymbol{V}_0$, we consider that the receiver employs an LT, which is appropriate in case of unknown $\boldsymbol{V}_0$ statistics. Let $p_{\hat{\boldsymbol{Q}}_{\text{AIB}}|\mathcal{H}_0}$ be the PDF of $\hat{\boldsymbol{Q}}_{\text{AIB}}$ in state $\hat{b}=0$. The generalized log-likelihood function is

$$\Psi = \log p_{\hat{\boldsymbol{Q}}_{\text{AIB}}|\mathcal{H}_0}(\boldsymbol{R}). \tag{10}$$

According to LT, Ψ is then compared with respect to a threshold τ, and the authentication procedure outputs

$$\hat{b} = \begin{cases} 0 & \Psi < \tau\,, \\ 1 & \Psi \geq \tau\,. \end{cases} \tag{11}$$

In state $\hat{b}=0$, and conditioned on the configuration $\boldsymbol{\Phi}'$ chosen by Bob, we can rewrite $\boldsymbol{R}$ from (5) and (7) as

$$\boldsymbol{R} = \overline{\boldsymbol{Q}}_{\text{AIB}}(\boldsymbol{\Phi}') + \boldsymbol{W} = \overline{\boldsymbol{Q}}_{\text{AIB}}(\boldsymbol{\Phi}') + \boldsymbol{W}_{\text{B}} - \overline{\boldsymbol{W}}\,, \tag{12}$$

where $\boldsymbol{W} = \boldsymbol{W}_{\text{B}} - \overline{\boldsymbol{W}}$ represents the overall noise matrix with independent entries, each with zero mean and variance $\sigma^2 = 2\sigma_{\text{B}}^2$. Thus, neglecting irrelevant constants, (10) becomes

$$\Psi = \frac{2}{\sigma^2}\|\boldsymbol{R} - \overline{\boldsymbol{Q}}_{\text{AIB}}(\boldsymbol{\Phi}')\|^2 = \frac{2}{\sigma^2}\|\boldsymbol{\delta}\|^2, \tag{13}$$

with

$$\boldsymbol{\delta} = \begin{cases} \boldsymbol{W}\,, & \text{in state } \hat{b}=0 \\ \boldsymbol{V}_0 - \boldsymbol{H}\boldsymbol{\Phi}'\boldsymbol{G} + \boldsymbol{W}\,, & \text{in state } \hat{b}=1. \end{cases} \tag{14}$$

A. Attack Strategy

We assume an adversary that aims to minimize the distance between her induced cascaded channel and the one expected by Bob from Alice. Specifically, Trudy seeks to minimize the mean squared error (MSE) given her observations

$$\boldsymbol{V}_0 = \arg\min_{\boldsymbol{V}'} \mathbb{E}[\|\boldsymbol{V}' - \boldsymbol{H}\boldsymbol{\Phi}'\boldsymbol{G}\|^2 \mid \boldsymbol{U}, \boldsymbol{H}]\,. \tag{15}$$

In the considered scenario, Trudy knows $\boldsymbol{H}$ and $\boldsymbol{U}$, which are correlated with $\boldsymbol{G}$, as well as the statistics of the RIS configuration, but not its actual realization $\boldsymbol{\Phi}'$. As a result, the design of the attack $\boldsymbol{V}_0$ cannot depend on the specific $\boldsymbol{\Phi}'$, but only on its statistics. Therefore, the average in (15) is also with respect to the RIS configuration. According to estimation theory, the solution of this problem is for Trudy to target the conditional expectation of the cascaded channel, given her available information [34], i.e.,

$$\boldsymbol{V}_0 = \mathbb{E}[\boldsymbol{H}\boldsymbol{\Phi}'\boldsymbol{G} \,|\, \boldsymbol{U}, \boldsymbol{H}]\,. \tag{16}$$

Now, the configuration statistics depend on $\boldsymbol{G}$, thus both the RIS configuration statistics and $\boldsymbol{G}$ depend on $\boldsymbol{U}$. This makes computing the expectation too complex. For a simpler solution, here we separate the averaging on the RIS configuration and $\boldsymbol{G}$, to obtain

$$\boldsymbol{V}_0 \approx \boldsymbol{H}\mathbb{E}[\boldsymbol{\Phi}' \,|\, \boldsymbol{U}, \boldsymbol{H}]\mathbb{E}[\boldsymbol{G} \,|\, \boldsymbol{U}, \boldsymbol{H}] = \boldsymbol{H}\mathbb{E}[\boldsymbol{\Phi}' \,|\, \boldsymbol{U}, \boldsymbol{H}]\overline{\boldsymbol{G}}\,. \tag{17}$$

The average $\overline{\boldsymbol{G}}$ can be obtained by averaging the right side of (3) with respect to $\boldsymbol{D}$, providing $\overline{\boldsymbol{G}} = \boldsymbol{U}/\rho$. It is worth noting that this strategy represents the mathematically optimal attack that Trudy can perform to minimize the distance metric under the uncertainty introduced by the secret challenge ϵ_n.

1) Attack Through the RIS: Trudy can exploit the current RIS configuration by transmitting through the RIS, but she does not know the instantaneous value of the RIS. Thus, she must determine a matrix $\boldsymbol{G}'$ to precode its transmitted signal, obtaining the cascaded channel $\boldsymbol{H}\boldsymbol{\Phi}'\boldsymbol{G}'$. However, the design of $\boldsymbol{G}'$ cannot depend on $\boldsymbol{\Phi}'$ but only on its statistics. As before, since $\boldsymbol{H}$ is fixed and $\boldsymbol{\Phi}'$ is random we consider

$$\overline{\boldsymbol{G}} = \arg\max_{\boldsymbol{G}'} \mathbb{P}[\boldsymbol{G} = \boldsymbol{G}'|\boldsymbol{U}], \tag{18}$$

which, considering the model (3), provides $\overline{\boldsymbol{G}} = \boldsymbol{U}/\rho$. In this case, the cascaded channel under attack through the RIS becomes

$$\boldsymbol{V} = \frac{1}{\rho}\boldsymbol{H}\boldsymbol{\Phi}'\boldsymbol{U} = \overline{\boldsymbol{Q}}_{AIB}(\boldsymbol{\Phi}') + \frac{\sqrt{1-\rho^2}}{\rho}\boldsymbol{H}\boldsymbol{\Phi}'\boldsymbol{D} + \boldsymbol{W}_B\,, \tag{19}$$

and we note that the resulting attack channel depends on the current configuration of the RIS, $\boldsymbol{\Phi}'$.

IV. Communication and Security Metrics

To perform the CR-PLA mechanism, Bob randomly selects the RIS configuration $\boldsymbol{\Phi}'$ to generate the challenge in the verification phase. However, such RIS configuration, while providing authentication capabilities, affects the data rate of the communication link between Alice and Bob. Therefore, we aim at properly designing the PDF of the RIS configuration p_ϵ to get a trade-off between the security and communication metrics. To this end, in this section, we first define the security performance of CR-PLA and then we optimize the PDF in the next section.

Let us define (recalling that $p_\epsilon(a)$ is even)

$$m = \mathbb{E}[e^{j\epsilon_n}] = \mathbb{E}[\cos\epsilon_n] = \int_{-pi}^{\pi} \cos(\alpha)p_\epsilon(\alpha)d\alpha\,, \tag{20}$$

$$s = \mathbb{E}[\cos 2\epsilon_n] = \int_{-\pi}^{\pi} \cos(2\alpha)p_\epsilon(\alpha)d\alpha\,. \tag{21}$$

From our derivation, we conclude that security and communication performance depend on the mean and variance of $\cos\epsilon$, rather than the perturbation ϵ. Indeed, m is the mean of $\cos\epsilon$, while from the relation $\cos^2 x = \frac{\cos(2x)+1}{2}$, we conclude s is a scaled and shifted version of the power of $\cos\epsilon$. Hence, values of m close to one correspond to small perturbations, preserving most of the beamforming gain, whereas smaller values indicate stronger deviations that increase channel randomness. On the other hand, parameter s captures the dispersion of the perturbation distribution around its mean value. Different phase-perturbation distributions with the same m but different s induce the same average misalignment but different levels of variability. In particular, a lower s corresponds to more dispersed phase perturbations, which further increases the unpredictability of the cascaded channel observed by Trudy.

In the remainder of this section, we write the communication and security metrics as a function of m and s.

A. Communication Metrics

For a given RIS configuration $\boldsymbol{\Phi}$, the resulting achievable rate of the Alice-RIS-Bob channel is defined as

$$C_{A,B}(\boldsymbol{\Phi}) = \log_2\left[\det\left(\boldsymbol{I}_{N_B} + \frac{1}{\sigma_B^2}\boldsymbol{Q}_{\text{AIB}}(\boldsymbol{\Phi})\boldsymbol{R}_A\boldsymbol{Q}_{\text{AIB}}^H(\boldsymbol{\Phi})\right)\right], \tag{22}$$

where $\boldsymbol{R}_A$ is the correlation matrix of the transmitted vector signal on the transmit antennas and $\boldsymbol{I}_{N_B}$ is the identity matrix of size N_B. The water-filling solution is adopted to determine $\boldsymbol{R}_A$ [35]. As the communication metric, we consider the average achievable rate for a given $\boldsymbol{H}$ and $\boldsymbol{G}$, where the average is done with respect to p_ϵ.

From the Jensen inequality, we have

$$\mathbb{E}[C_{A,B}(\boldsymbol{\Phi}')] \le \log_2\left[\det\left(\boldsymbol{I}_{N_B} + \frac{1}{\sigma_B^2}\mathbb{E}[\boldsymbol{Q}_{\text{AIB}}(\boldsymbol{\Phi}')\boldsymbol{R}_A\boldsymbol{Q}_{\text{AIB}}^H(\boldsymbol{\Phi}')]\right)\right] = \Omega\,, \tag{23}$$

since $\log\det$ is a concave function. Let us define $\boldsymbol{B} = \mathbb{E}[\boldsymbol{Q}_{\text{AIB}}(\boldsymbol{\Phi}')\boldsymbol{R}_A\boldsymbol{Q}_{\text{AIB}}^H(\boldsymbol{\Phi}')]$, then

$$\boldsymbol{B} = \mathbb{E}\left[\boldsymbol{H}\boldsymbol{\Phi}'\boldsymbol{G}\boldsymbol{R}_A\boldsymbol{G}^{\text{H}}\boldsymbol{\Phi}'^*\boldsymbol{H}^{\text{H}}\right] = \boldsymbol{H}\overline{\boldsymbol{\Phi}}\mathbb{E}[\boldsymbol{A}]\overline{\boldsymbol{\Phi}}^*\boldsymbol{H}^{\text{H}}\,, \tag{24}$$

where $\boldsymbol{A} = \boldsymbol{E}\boldsymbol{G}\boldsymbol{R}_A\boldsymbol{G}^{\text{H}}\boldsymbol{E}^*$ and $\boldsymbol{E} = \text{diag}\{e^{j\epsilon_1},\dots,e^{j\epsilon_N}\}$, with ϵ_i the random phases of the diagonal of $\boldsymbol{\Phi}'$. The (i,j)-th element of the inner expected value in (24) is

$$\begin{aligned}\mathbb{E}[\boldsymbol{A}_{i,j}] &= \mathbb{E}[e^{j(\epsilon_i-\epsilon_j)}]\sum_{k=1}^{N}\sum_{\ell=1}^{N}\boldsymbol{G}_{i,k}\boldsymbol{R}_{A,k,\ell}\boldsymbol{G}_{j,\ell}^*\\ &= \mathbb{E}[e^{j(\epsilon_i-\epsilon_j)}]\lambda_{i,j}\,,\end{aligned} \tag{25}$$

with $\lambda_{i,j} = \sum_{k=1}^{N}\sum_{\ell=1}^{N}\boldsymbol{G}_{i,k}\boldsymbol{R}_{A,k,\ell}\boldsymbol{G}_{j,\ell}^*$. Then, since

$$\mathbb{E}\left[e^{j(\epsilon_i-\epsilon_j)}\right] = \begin{cases} m^2, & i \neq j\,,\\ 1, & i = j\,,\end{cases} \tag{26}$$

the (i,j)-th element of matrix $\boldsymbol{B}$ is written as a function of m as

$$\boldsymbol{B}_{i,j}(m) = m^2\boldsymbol{M}_{i,j} + \boldsymbol{T}_{i,j}\,, \tag{27}$$

where

$$\boldsymbol{M}_{i,j} = \sum_{k=1}^{N}\boldsymbol{H}_{j,k}^*\sum_{\ell=1,\ell\neq k}^{N}\boldsymbol{H}_{i,\ell}e^{j(\bar{\theta}_\ell-\bar{\theta}_k)}\lambda_{\ell,k}\,, \tag{28}$$

$$\boldsymbol{T}_{i,j} = \sum_{k=1}^{N}\boldsymbol{H}_{i,k}\boldsymbol{H}_{j,k}^*\lambda_{k,k}\,. \tag{29}$$

Finally, the upper bound on average channel capacity is a function of m, i.e.,

$$\Omega(m) = \log_2\left[\det\left(\boldsymbol{I}_{N_B} + \frac{1}{\sigma_B^2}\boldsymbol{B}(m)\right)\right], \tag{30}$$

where we have indicated the dependency of $\boldsymbol{B}$ from m.

B. Security Metrics

Two possible error events might occur in the authentication mechanism: the FA, when Bob discards a message as forged by Trudy while it is coming from Alice, and the MD, when Bob accepts a message coming from Trudy as legitimate. Specifically, an FA occurs when, in state $b = 0$, $\Psi \ge \tau$, whereas, an MD occurs when, in state $b = 1$, $\Psi < \tau$. As security metrics, we then consider the probabilities of FA and MD. In formulas, for a given Alice-RIS and RIS-Bob channels $\boldsymbol{G}$ and $\boldsymbol{H}$ and any configuration $\boldsymbol{\Phi}'$, we define then the probability of FA and MD respectively as

$$P_{\text{FA}} = \mathbb{P}[\Psi \ge \tau|b=0],\ P_{\text{MD}} = \mathbb{E}\left[\mathbb{P}[\Psi < \tau|b=1,\boldsymbol{\Phi}']\right], \tag{31}$$

where the expectation is taken with respect to both the RIS configuration and $\boldsymbol{D}$. Note that the P_{MD} depends also on the attack $\boldsymbol{V}_0$, which here is assumed to be that described in Section III-A.

Under the legitimate state $b = 0$, from (14) we have that $\boldsymbol{\delta} = \boldsymbol{W}$ and Ψ becomes a central chi-square random variable with $2N_{\text{A}}N_{\text{B}}$ degrees of freedom and

$$P_{\text{FA}} = 1 - F_{\chi^2_{2N_{\text{A}}N_{\text{B}}},0}(\tau), \tag{32}$$

where $F_{\chi^2_{2N_{\text{A}}N_{\text{B}}},a}(\cdot)$ is the cumulative distribution function (CDF) of a non-central chi-square variable with $2N_{\text{A}}N_{\text{B}}$ degrees of freedom and non-centrality parameter a. It is worth noting that τ is typically set to reach a desired $\overline{P}_{\text{FA}}$, i.e.,

$$\tau^\star = F^{-1}_{\chi^2_{2N_{\text{A}}N_{\text{B}}},0}(1 - \overline{P}_{\text{FA}})\,. \tag{33}$$

In state $b = 1$ with attack $\boldsymbol{V}_0$ defined in Section III-A. Since the average optimal RIS configuration $\overline{\hat{\boldsymbol{\Phi}}}$ estimated by Trudy is obtained from $\overline{G}$, it is statistically dependent on $\boldsymbol{D}$, thus $\overline{\hat{\boldsymbol{\Phi}}}$ and $\boldsymbol{U}$ are not independent. To simplify the analysis, here we consider the worst case in which Trudy perfectly knows the *average* RIS configuration matrix, thus $\overline{\hat{\boldsymbol{\Phi}}} = \overline{\boldsymbol{\Phi}}$. Under this assumption, we can obtain an upper bound of the MD probability achieved by an attacker with imperfect knowledge of the average RIS configuration.

Under this assumption, for a given RIS configuration $\boldsymbol{\Phi}'$, we have

$$\boldsymbol{\delta} = \boldsymbol{V} - \overline{\boldsymbol{Q}}_{\text{AIB}} = \frac{m}{\rho}\boldsymbol{H}\overline{\boldsymbol{\Phi}}\boldsymbol{U} - \boldsymbol{H}\boldsymbol{\Phi}'\boldsymbol{G} + \boldsymbol{W}\,, \tag{34}$$

since $\mathbb{E}[\boldsymbol{\Phi}] = m\overline{\boldsymbol{\Phi}}$.

The upper bound on the MD probability, P_{MD}, is obtained as the average CDF of $\Psi = 2/\sigma^2\|\boldsymbol{\delta}\|^2$ evaluated at $\tau^\star$, and as derived in Appendix A, we have

$$P_{\text{MD}} = \mathbb{E}\big[\mathbb{P}[\Psi < \tau^\star | b = 1]\big] = \mathbb{P}\left[\frac{2}{\sigma^2}\|\boldsymbol{\delta}\|^2 < \tau^\star | b = 1\right], \tag{35}$$

where the expectation is still done with respect to the RIS configuration and $\boldsymbol{D}$ for given channel matrices. We now investigate the dependency of P_{MD} on p_ϵ, showing that it can be uniquely written as a function of m and s.

The generic entry of the matrix $\boldsymbol{\delta}$ is, from (34),

$$\delta_{r,c} = \sum_{n=1}^{N} H_{r,n}e^{j\bar{\theta}_n}\left(\frac{m}{\rho}U_{n,c} - e^{j\epsilon_n}G_{n,c}\right) + W_{r,c}. \tag{36}$$

Next, we define the vectorization of $\boldsymbol{\delta}$ as

$$\bar{\boldsymbol{\delta}} = [\text{vec}(\text{Re}\{\boldsymbol{\delta}\}), \text{vec}(\text{Im}\{\boldsymbol{\delta}\})] \in \mathbb{C}^{2N_{\text{A}}N_{\text{B}}}\,. \tag{37}$$

Therefore, for a given $\mathbf{H}$ and $\mathbf{G}$, while $\bar{\boldsymbol{\delta}}$ depends on both the Gaussian matrices $\mathbf{U}, \mathbf{W}$ and the random phase terms $e^{j\epsilon_n}$, it can be modeled as an approximately Gaussian vector for large N by invoking the Central Limit Theorem. This vector has zero mean entries (see Appendix B) and covariance matrix $\boldsymbol{\Sigma}_\delta(m, s)$ computed in Appendix C. There, we also show that $\boldsymbol{\Sigma}_\delta$ depends only on m and s and not on the specific p_ϵ.

The decision metric Ψ in (13) has the quadratic form

$$\Psi = \frac{2}{\sigma^2}\bar{\boldsymbol{\delta}}^*\bar{\boldsymbol{\delta}}\,. \tag{38}$$

By defining $\boldsymbol{Y} = \boldsymbol{\Sigma}_\delta^{-1/2}(m, s)\bar{\boldsymbol{\delta}}$, we can rewrite (38) as

$$\Psi = \boldsymbol{Y}^*\boldsymbol{\Sigma}_\delta^{1/2}(m, s)\frac{2}{\sigma^2}\boldsymbol{\Sigma}_\delta^{1/2}(m, s)\boldsymbol{Y}\,. \tag{39}$$

Note that $\boldsymbol{Y}$ has mean zero and identity covariance matrix. Using the spectral theorem, we obtain

$$\boldsymbol{\Sigma}_\delta^{1/2}(m, s)\frac{2}{\sigma^2}\boldsymbol{\Sigma}_\delta^{1/2}(m, s) = \boldsymbol{P}(m, s)^*\boldsymbol{\Lambda}(m, s)\boldsymbol{P}(m, s)\,, \tag{40}$$

with $\boldsymbol{P}(m, s)$ orthogonal and $\boldsymbol{\Lambda}(m, s)$ is diagonal with positive diagonal entries $\lambda_1, \ldots, \lambda_{N_{\text{B}}N_{\text{A}}}$. Let $\boldsymbol{L} = \boldsymbol{P}(m, s)\boldsymbol{Y}$, which turns to be a multivariate (real) normal with zero mean and identity covariance matrix for a given $\boldsymbol{H}$ and $\boldsymbol{G}$. Thus, we have

$$\Psi = \boldsymbol{L}^*\boldsymbol{\Lambda}(m, s)\boldsymbol{L} = \sum_{j=1}^{N_{\text{B}}N_{\text{A}}} \lambda_j(m, s)|L_j|^2\,, \tag{41}$$

where $\lambda_j(m, s)$ is the j-th eigenvalue of $2/\sigma^2\boldsymbol{\Sigma}_\delta(m, s)$. Also, Ψ is a linear combination of central chi-squared random variables with 1 degree of freedom each [36] and the MD probability is [36]

$$P_{\text{MD}}(m, s, \tau^\star) = \mathbb{P}\left[\sum_{j=1}^{N_{\text{A}}N_{\text{B}}} \lambda_j(m, s)|L_j|^2 < \tau^\star\frac{\sigma^2}{2}\right], \tag{42}$$

where we have made explicit the dependence of the P_{MD} by the parameters m and s of the PDF p_ϵ, and the threshold τ given by the desired P_{FA}. The derived $P_{\text{MD}}(m, s, \tau^\star)$ (42) can be expressed by a Laguerre polynomial as shown in Appendix E.

1) Attack Through the RIS: When Trudy transmits through the RIS, we have that $\bar{\boldsymbol{\delta}}$ is zero-mean (see Appendix D) with covariance matrix $\boldsymbol{\Sigma}_\delta$ computed in Appendix D. Note that in this case, P_{MD} does not depend on m and s anymore, i.e., the randomness of the RIS for PLA. Thus, the performance of CR-PLA when Trudy transmits through the RIS is that of a TAG-based PLA, depending on the accuracy of the estimate of the Alice-RIS channel.

V. Design of p_ϵ

For a desired $\overline{P}_{\text{FA}}$, our objective is to find the optimal p_ϵ balancing the communication metrics and security requirements. As shown in Section IV, both the upper bound on the communication and security metrics depend only on the parameters m and s. Therefore, we aim at finding a feasible (m, s) pair such that upper bound on the average achievable rate $\Omega(m)$ is maximized, while ensuring that $P_{\text{MD}}(m, s, \tau^\star)$ remains below a certain threshold η. In formulas,

$$(m^\star, s^\star) = \arg\max_{(m,s)}\ \Omega(m) \tag{43a}$$

$$\text{s.t.}\quad P_{\text{MD}}(m, s, \tau^\star) \le \eta \tag{43b}$$

$$-1 \le m \le 1 \tag{43c}$$

$$(m, s) \in \mathcal{D}\,, \tag{43d}$$

where $\mathcal{D}$ is the set of couples (m, s) for which a PDF p_ϵ with these statistical parameters exists. We resort to a numerical approach (i.e., fmincon MATLAB function) to solve (43).

Note that we maximize an upper bound on capacity. Thus, in general, we cannot guarantee that we will achieve the actual objective function and that the resulting configuration parameters will be optimal.

A. Complexity

For each iteration, fmincon computes $\Omega(m)$ and the nonlinear constraint on the P_{MD}. According to (30), the former operation costs $O(N_B^3)$, while the cost of the latter is mainly due to the eigenvalue decomposition, which is $O((N_A N_B)^3)$ (see (42)). In fact, the numerical evaluation of the P_{MD} using Imhof integral method is negligible with respect to the eigenvalue decomposition. The overall cost of solving (43) is then $\text{O}(I(N_B^3 + (N_A N_B)^3))$, where I represents the number of iterations performed by fmincon. The procedure is then lightweight and suitable for online adaptation.

B. Attack Through the RIS

When Trudy transmits through the RIS, the optimization problem (43) becomes either trivially satisfied (with $m = 1$) or irrelevant, since introducing additional phase randomness does not provide further authentication benefits under this specific attack model.

We now provide a result characterizing the set $\mathcal{D}$. Note that the optimal (m, s) must be feasible according to (6), i.e., there should exist a PDF $p_\epsilon(a)$ (which is even in a) such that $m = \mathbb{E}[\cos \epsilon_n]$ and $s = \mathbb{E}[\cos 2\epsilon_n]$.

Theorem 1: There exists an even $p_\epsilon(a)$ such that

$$m = \int_{-\pi}^{\pi} \cos(\alpha) p_\epsilon(\alpha) d\alpha \,, \tag{44}$$

$$s = \int_{-\pi}^{\pi} \cos(2\alpha) p_\epsilon(\alpha) d\alpha \,, \tag{45}$$

if and only if (m, s) is in the set

$$\mathcal{D} = \{(m, s) : s \geq 2m^2 - 1 \quad \text{and} \quad m, s \in [-1, 1]\} \,. \tag{46}$$

Proof: We first prove that (46) is a necessary condition. Let $X = \cos \epsilon$, where ϵ is distributed according to a valid p_ϵ. Since $-1 \leq \cos \epsilon \leq 1$, the random variable X takes values in $[-1, 1]$. From

$$\cos(2\alpha) = 2\cos^2(\alpha) - 1 \,, \tag{47}$$

we have

$$m = \mathbb{E}(X) \,, \tag{48}$$

$$s = \mathbb{E}[2X^2 - 1] = 2\mathbb{E}[X^2] - 1 \,. \tag{49}$$

Let $\mathbb{E}[X^2] = \sigma^2 + m^2$, where $\sigma^2 = \text{Var}(X) \geq 0$. Then, since $\sigma^2 \geq 0$,

$$s = 2(\sigma^2 + m^2) - 1 \geq 2m^2 - 1 \,. \tag{50}$$

To prove that (46) is a sufficient condition, we show that for any pair (m, s) satisfying (46), there exists an even PDF p_ϵ such that (20) and (21) are satisfied. We consider the family of even PDFs

$$p_\epsilon(\alpha) = a\,\delta(\alpha - \theta) + a\,\delta(\alpha + \theta) + (1 - 2a)\,\delta(\alpha), \tag{51}$$

with $\theta \in (0, \pi]$ and $a \in \left[0, \frac{1}{2}\right]$. This is a valid even PDF as it is symmetric and integrates to 1. Then, we have

$$m = 2a\cos(\theta) + 1 - 2a, \tag{52}$$

$$s = 2a\cos(2\theta) + 1 - 2a. \tag{53}$$

By using (47), (52), and (53), by algebraic steps we have

$$s - 2m^2 + 1 = -4a(2a - 1)(\cos(\theta) - 1)^2 \,. \tag{54}$$

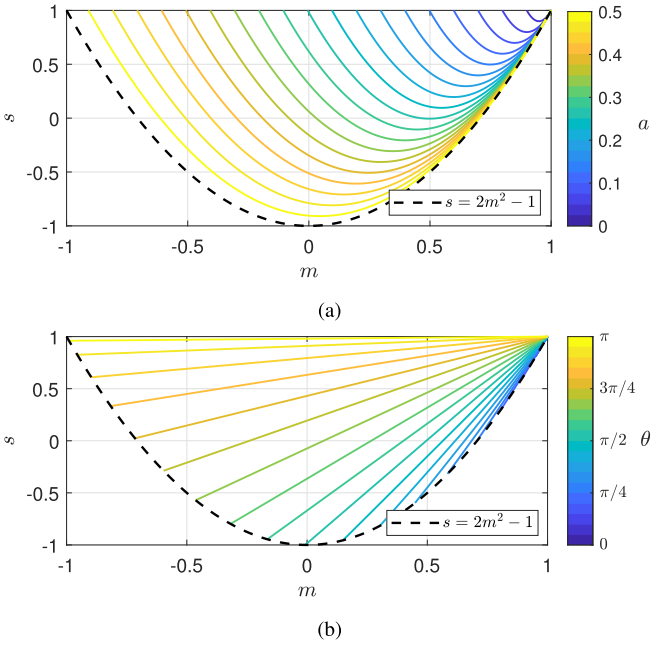

Fig. 2. m and s pairs that can be achieved varying a and θ in (51).

The right-hand side of (54) is non-negative for any for $\theta \in [0, \pi]$ and $a \in \left[0, \frac{1}{2}\right]$, thereby satisfying condition (50). Also, given the values of m and s such that $s \geq 2m^2 - 1$, with some algebraic computations we have

$$a = \frac{-(m-1)^2}{4m - s - 3} \,, \tag{55}$$

and

$$\theta = \arccos\left(1 + \frac{m-1}{2a}\right) \,. \tag{56}$$

Note that when $s = 2m^2 - 1$, we have $a = 1/2$. Thus, when m and s belong to the boundary, so (46) holds with equality, we have $a = \frac{1}{2}$ and $\theta = \arccos(m)$. The m and s pairs that can be achieved by varying a and θ in (51) are shown in Fig. 2. ■

It is important to note that the proof of Theorem 1 is constructive. Given a desired pair (m, s) satisfying (46), we recover the parameters (a, θ) such that the distribution p_ϵ in (51) achieves the target values. In this way, the proof provides an operative method to construct a valid p_ϵ satisfying the constraints explicitly.

Remark: The result presented in Theorem 1 extends the analysis in [18], where the random variable ϵ is assumed to follow a uniform distribution, i.e., $\epsilon \sim \mathcal{U}(-\gamma, \gamma)$. From (20) and (21) we have

$$m = \frac{\sin\gamma}{\gamma}, \quad s = \frac{\sin\gamma\cos\gamma}{\gamma} = m\cos\gamma. \tag{57}$$

However, not all combinations of m and s can be attained by varying γ. As a counterexample, consider the pair $(m, s) = (0, 1)$, which from the second expression of s in (57) cannot be achieved for any γ. This demonstrates that the uniform assumption is not general enough to capture all possible combinations of m and s, while a more general expression for p_ϵ is given by (51).

Remark: We now consider the case in which the phases of RIS elements are quantized, i.e., for $\epsilon = \kappa\hat{\theta}$, with κ integer and

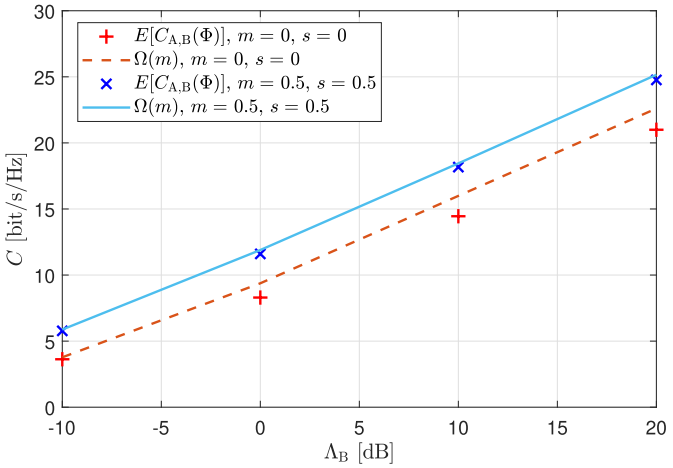

Fig. 3. Average capacity and upper bound vs the SNR on the Alice-Bob channel Λ_B, for $N_A = N_B = 2$, and $N = 36$.

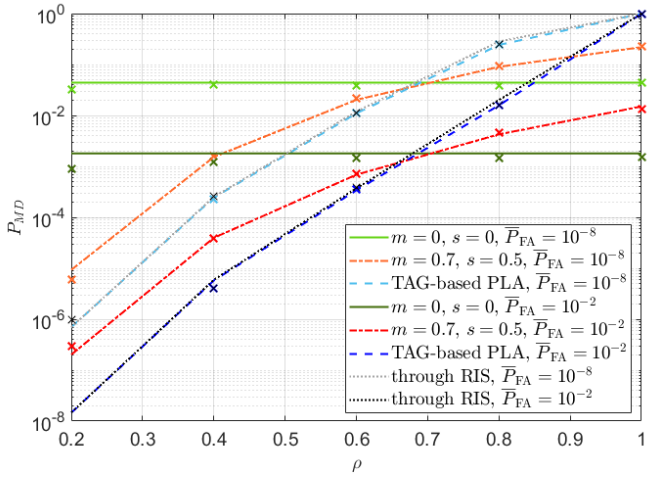

Fig. 4. P_{MD} as a function of the channel correlation factor ρ for $N_A = N_B = 2$, $N = 36$, $\sigma_B^2 = 0.5$, and $P_{FA} = 10^{-2}$ and 10^{-8}. Crosses show the estimate obtained via Monte Carlo simulations.

$\hat{\theta}$ the quantization level. For a given value of $\cos\hat{\theta} \le m \le 1$, we can still use the PDF (51) with

$$a = \frac{1-m}{2(1-\cos\hat{\theta})}. \tag{58}$$

With this PDF, using conventional trigonometric identities to relate $\cos(2\hat{\theta})$ and $\cos\hat{\theta}$, we obtain

$$s = 2m\cos\hat{\theta} + 2\cos\hat{\theta} - 1. \tag{59}$$

Therefore, for a fixed quantized phase $\hat{\theta}$, the achievable (m, s) pairs lie on a one-dimensional segment in the (m, s) plane. Phase quantization thus restricts the feasible set $\mathcal{D}$ to a finite union of such segments, while increasing the number of quantization levels progressively recovers the continuous feasible region characterized in Theorem 1. Consequently, RIS phase quantization restricts the set of perturbation distributions that can be physically implemented and may introduce an additional loss in communication rate. On the other hand, coarse quantization increases the inherent randomness of the cascaded channel, which may further enhance the discriminability between legitimate and forged responses in the CR-PLA mechanism.

VI. Numerical Results

In this Section, we validate the above analysis, providing numerical evidence of the balance between communication metrics and security requirements resulting from the optimization problem (43).

Fig. 3 compares our upper bound on the average channel capacity in (30) with the average capacity obtained through Monte Carlo simulations in (22), as a function of the Alice-Bob SNR defined as $\Lambda_B = 1/\sigma_B^2$. The upper bound holds and is generally tight.

Fig. 4 shows the upper bound of the MD probability P_{MD} versus the channel correlation factor ρ, for $N_A = N_B = 2$, $N = 36$, $P_{FA} = 10^{-2}$ and 10^{-8}, and $\sigma_B^2 = 0.5$. The curves show the values of P_{MD} computed using (42) and the Imhof integral method described in Appendix E, while the crosses indicate the upper bound of the MD probability estimated via Monte Carlo results. First, we observe that the analytical and the Monte Carlo simulations exhibit the same behavior, thereby validating the analytical derivations. The curves labeled by different values of the parameters m and s refer to the CR-PLA scheme under a direct impersonation attack. We also include the performance of the TAG-based PLA and CR-PLA with the attack through the RIS, labeled as TAG-based PLA and through RIS, respectively. In the TAG-based PLA, the RIS is fixed with the communication-optimal configuration. This scenario is equivalent to the special case CR-PLA with $m = s = 1$, where the CR-PLA mechanism degenerates into the conventional TAG-based authentication scheme. Moreover, when Trudy attacks through the RIS in the CR-PLA scheme, her forged transmission inherently incorporates the instantaneous RIS configuration, so that the random perturbations affect both channels similarly. As a result, the obtained P_{MD} coincides with the TAG-based PLA. As observed, for a low correlation ρ, the TAG-based PLA and the through- the-RIS attack are quite weak. However, as ρ increases, they become more effective, yielding a higher P_{MD}. This confirms that without phase randomizing the RIS configuration, Trudy can effectively hide her signal by mimicking Alice's channel or exploiting the effects of the RIS. On the other hand, CR-PLA with $m < 1$ and $s < 1$, even as $\rho \to 1$, is effective in detecting the attacks. This is because Trudy does not know the random phase challenge. Therefore, even if she knows the average channel perfectly, she cannot predict the instantaneous phase shift applied by the RIS for that specific message, forcing a high residual error δ at the verifier and making the attack much easier to detect. Finally, we observe that, when $m = 0$, the P_{MD} remains constant, since δ in (34) depends on the ratio m/ρ.

The optimization process and the resulting performance trade-offs are further analyzed in Figs. 5 and 6. Specifically, Fig. 5 shows the contour levels of the P_{MD} over the feasible region for the optimization parameters (m, s) varying $\eta \in \{0.07, 0.08, 0.09, 0.1\}$ for $N_A = N_B = 2$, $N = 36$, and $P_{FA} = 10^{-5}$, and $\rho = 0.9$. Since each contour curve corresponds to a different η value, any point (i.e., (m, s) pair) along a given curve has the same P_{MD}. However, due to the non-convexity and nonlinearity of the constraint surface

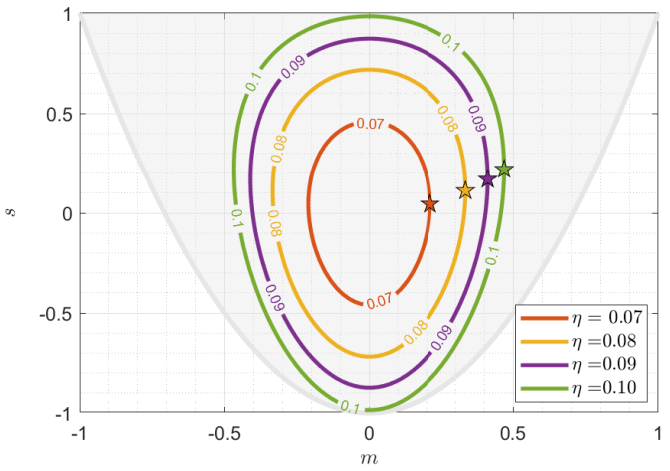

Fig. 5. Contour levels of the $P_{\mathrm{MD}}(m, s, \tau)$ over the (m, s) plane varying $\eta \in \{0.07, 0.08, 0.09, 0.1\}$ for $N_{\mathrm{A}} = N_{\mathrm{B}} = 2$, $N = 36$, and $P_{\mathrm{FA}} = 10^{-5}$. The shaded region represents the area satisfying (46). The star represents the resulting optimal m value, i.e., $m^{\star}$.

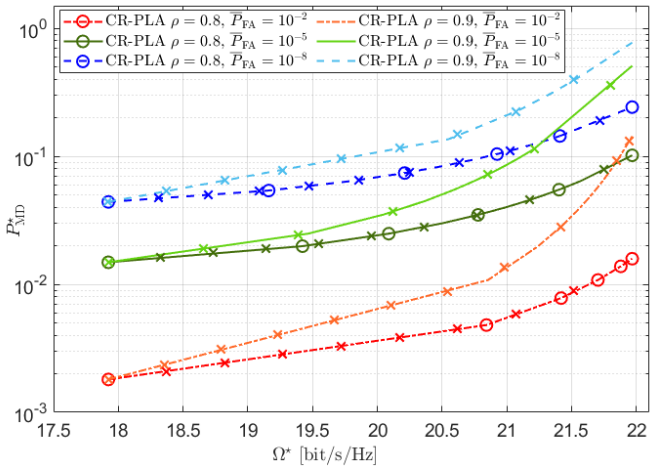

Fig. 6. P_{MD} as a function of the channel capacity Ω, for $N_{\mathrm{A}} = N_{\mathrm{B}} = 2$, $N = 36$, and different values of ρ and p_{FA}, with $\Lambda_{\mathrm{B}} = 8$ dB. Crosses represent the considered value of η.

defined by the constraint on P_{MD}, determining the optimal (m, s) pair that maximizes the objective function $\Omega(m)$ is not trivial. As a result, we resort to a numerical approach to solve (43), and the obtained optimal (m, s) pairs are indicated with the star marker. Fig. 6 shows the relationship between $P_{\mathrm{MD}}(m^{\star}, s^{\star}, \tau)$ and the upper bound on the communication rate $\Omega(m^{\star})$, providing an insight into the resulting trade-off between the communication and security metrics. The plot considers different target values for the FA probability $\overline{P}_{\mathrm{FA}} \in 10^{-8}, 10^{-5}, 10^{-2}$, with $N_{\mathrm{A}} = N_{\mathrm{B}} = 2$, $N = 36$, and $\Lambda_{\mathrm{B}} = 8$ dB. As shown, increasing $\Omega(m^{\star})$ results in a higher $P_{\mathrm{MD}}(m^{\star}, s^{\star}, \tau)$ for a fixed $\overline{P}_{\mathrm{FA}}$. Moreover, when aiming for a smaller $\overline{P}_{\mathrm{FA}}$, the required reduction in $\Omega(m^{\star})$ to achieve the minimum $P_{\mathrm{MD}}(m^{\star}, s^{\star}, \tau^{\star}) = P^{\star}_{\mathrm{MD}}$ becomes less pronounced. However, a lower $\overline{P}_{\mathrm{FA}}$ also leads to a higher floor in the achievable $P_{\mathrm{MD}}(m^{\star}, s^{\star}, \tau)$. Notably, the points at the far right of each curve, corresponding to the maximum achievable capacity $\Omega(m^{\star})$, represent the performance achieved by the conventional TAG-based PLA or, equivalently, by the CR-PLA under a through-RIS attack. In these cases, the system fails to provide additional security through phase randomness, and the P_{MD} reaches its upper floor. Additionally, for a fixed $\overline{P}_{\mathrm{FA}}$, a lower ρ (i.e., stronger channel correlation) results in a lower P_{MD} for the same $\Omega(m^{\star})$. Finally, note that for the optimal solution of problem (43), i.e., the $(m^{\star}, s^{\star})$ pair, we find that $P^{\star}_{\mathrm{MD}} = \eta$, i.e., the constraint on the P_{MD} holds with equality.

VII. Conclusion

In this paper, we considered a CR-PLA mechanism that leverages the presence of an RIS to perform the challenge-response protocol for devices equipped with multiple antennas. We designed the probability distribution of the randomly selected RIS phase shifts to maximize an upper bound on the average cascaded channel capacity while satisfying an upper bound on the MD probability for a desired FA probability. We derived approximate expressions for the FA and MD probabilities, and the average cascaded channel capacity. The simulation results show that a high average channel capacity would come at the expense of a reduction of the MD probability and vice versa.

Appendix A
P_{MD} Derivation

From (31) we have

$$P_{\mathrm{MD}}(\boldsymbol{\Phi}') = \int_{-\infty}^{\tau} p_{\Psi|\boldsymbol{\Phi},b=1}(\psi|\boldsymbol{\Phi}')\mathrm{d}\psi\,. \tag{60}$$

So, the P_{MD} is

$$\begin{aligned} P_{\mathrm{MD}} &= \int_{-\infty}^{\infty} p_{\mathrm{MD}}(\boldsymbol{\Phi}')p_{\boldsymbol{\Phi}}(\boldsymbol{\Phi}')\mathrm{d}\boldsymbol{\Phi}' \\ &= \int_{-\infty}^{\infty} p_{\boldsymbol{\Phi}}(\boldsymbol{\Phi}') \int_{-\infty}^{\tau} p_{\Psi|\boldsymbol{\Phi},b=1}(\psi|\boldsymbol{\Phi}')\mathrm{d}\psi\mathrm{d}\boldsymbol{\Phi}' \\ &= \int_{-\infty}^{\tau} p_{\Psi|b=1}(\psi)\mathrm{d}\psi = \mathbb{P}[\Psi < \tau | b = 1]. \end{aligned} \tag{61}$$

Appendix B
Expected Value of $\boldsymbol{\delta}$

The element in position (r, c) of $\boldsymbol{\delta}$ is given in (36). The expectation taken with respect to the noise and matrix $\boldsymbol{D}$ yields

$$\begin{aligned} \mathbb{E}[\mathrm{Re}\{\delta_{r,c}\}] &= \mathbb{E}\left[\mathrm{Re}\left\{\sum_n \alpha_n \left(m - e^{j\epsilon_n}\right)\right\}\right] \\ &= \mathbb{E}\left[\sum_n \alpha_n^{\mathrm{R}}(m - \cos\epsilon_n) + \alpha_n^{\mathrm{I}} \sin\epsilon_n\right], \end{aligned} \tag{62}$$

where $\alpha_n = H_{r,n} e^{j\bar{\theta}_n} G_{n,c}$, with $\alpha_n^{\mathrm{R}} = \mathrm{Re}\{\alpha_n\}$ and $\alpha_n^{\mathrm{I}} = \mathrm{Im}\{\alpha_n\}$. Under the assumption that the distribution $p_{\epsilon}(a)$ is an even function of a, we have $\mathbb{E}[\sin\epsilon_n] = 0$ and $\mathbb{E}[\cos\epsilon_n] = m$. As a result, it follows that $\mathbb{E}[\mathrm{Re}\{\delta_{r,c}\}] = 0$. By applying the same reasoning, we also obtain $\mathbb{E}[\mathrm{Im}\{\delta_{r,c}\}] = 0$.

Note that since $\overline{\boldsymbol{\delta}} = [\mathrm{vec}(\mathrm{Re}\{\boldsymbol{\delta}\}), \mathrm{vec}(\mathrm{Im}\{\boldsymbol{\delta}\})]$, it follows that each entry of $\overline{\boldsymbol{\delta}}$ has zero mean.

APPENDIX C
EVALUATION OF THE COVARIANCE MATRIX OF $\boldsymbol{\delta}$

The covariance matrix of $\bar{\boldsymbol{\delta}}$ can be expressed as

$$\boldsymbol{\Sigma}_{\delta} = \begin{bmatrix} \boldsymbol{\Sigma}_{\delta,\mathrm{R}} & \boldsymbol{\Sigma}_{\delta,\mathrm{RI}} \\ \boldsymbol{\Sigma}_{\delta,\mathrm{RI}}^{\mathrm{H}} & \boldsymbol{\Sigma}_{\delta,\mathrm{I}} \end{bmatrix}. \tag{63}$$

In the following, we will compute each component of $\boldsymbol{\Sigma}_{\delta}$ separately.

A. Cross-Covariance Between Real and Imaginary Part

We evaluate the cross-covariance matrix $\boldsymbol{\Sigma}_{\delta,\mathrm{RI}}$ between the real and the imaginary part of $\boldsymbol{\delta}$ in (34). Since $\mathbb{E}[\mathrm{Re}\{\delta_{r,c}\}] = \mathbb{E}[\mathrm{Im}\{\delta_{r,c}\}] = 0$ (see Appendix B), the element in position (i, j) of $\boldsymbol{\Sigma}_{\delta,\mathrm{RI}}$ is

$$\boldsymbol{\Sigma}_{\delta,\mathrm{RI}}(i,j) = \mathbb{E}\left[\mathrm{Re}\{\delta_{r_1,c_1}\}\mathrm{Im}\{\delta_{r_2,c_2}\}\right], \tag{64}$$

with $i, j \in \{0, 1, \ldots, N_{\mathrm{A}}N_{\mathrm{B}} - 1\}$, $\delta_{r,c}$ is as in (36), and

$$r_1 = \left\lfloor \frac{i}{N_{\mathrm{A}}} \right\rfloor, \quad c_1 = i \mod N_{\mathrm{A}}, \tag{65}$$

$$r_2 = \left\lfloor \frac{j}{N_{\mathrm{A}}} \right\rfloor, \quad c_2 = j \mod N_{\mathrm{A}}. \tag{66}$$

Defining

$$\alpha_{n,k} = H_{r_k,n} e^{j\bar{\theta}_n} G_{n,c_k}, \tag{67}$$

from (3), (34), and (36) we have

$$\mathbb{E}\Big[\mathrm{Re}\{\delta_{r_1,c_1}\}\mathrm{Im}\{\delta_{r_2,c_2}\}\Big] = \mathbb{E}\left[\sum_{n=1}^{N}\mathrm{Re}\{\alpha_{n,1}(m - e^{j\epsilon_n})\}\sum_{n=1}^{N}\mathrm{Im}\{\alpha_{n,2}(m - e^{j\epsilon_n})\}\right], \tag{68}$$

since $\boldsymbol{D}$ in (3) and $\boldsymbol{W}$ in (34) are CSCG, independent from all other variables.

We now look at the product between two generic terms n_1 and n_2 of the two summations in (68), respectively,

$$\begin{aligned}&\mathbb{E}\Big[\mathrm{Re}\{\alpha_{n_1,1}(m - e^{j\epsilon_{n_1}})\}\mathrm{Im}\{\alpha_{n_2,2}(m - e^{j\epsilon_{n_2}})\}\Big] \\ &= \mathbb{E}[(\alpha_{n_1,1}^{\mathrm{R}}(m - \cos\epsilon_{n_1}) + \alpha_{n_1,1}^{\mathrm{I}}\sin\epsilon_{n_1}) \\ &\quad \cdot (\alpha_{n_2,2}^{\mathrm{I}}(m - \cos\epsilon_{n_2}) - \alpha_{n_2,2}^{\mathrm{R}}\sin\epsilon_{n_2})],\end{aligned} \tag{69}$$

where $\alpha_{n_i,i}^{\mathrm{R/I}}$ denote the real (R) and imaginary (I) part of $\alpha_{n_i,i}$. Since ϵ_n, $n = 0, \ldots, N-1$, are i.i.d. and we assumed p_ϵ even, for $n_1 \neq n_2$ we have $\mathbb{E}[\cos(\epsilon_{n_1})\cos(\epsilon_{n_2})] = m^2$, $\mathbb{E}[\sin(\epsilon_{n_1})\cos(\epsilon_{n_2})] = 0$, and $\mathbb{E}[\sin(\epsilon_{n_1})\sin(\epsilon_{n_2})] = 0$. So, with simple computations, we have that when $n_1 \neq n_2$ (69) is 0. Considering instead $n_1 = n_2 = n$ we have

$$\begin{aligned}&\mathbb{E}\Big[\mathrm{Re}\{\alpha_{n,1}(m - e^{j\epsilon_n})\}\mathrm{Im}\{\alpha_{n,2}(m - e^{j\epsilon_n})\}\Big] \\ &= -\alpha_{n,1}^{\mathrm{R}}\alpha_{n,2}^{\mathrm{I}}m^2 - \alpha_{n,1}^{\mathrm{R}}\alpha_{n,2}^{\mathrm{I}}\mathbb{E}[\cos^2\epsilon_n] - \alpha_{n,1}^{\mathrm{I}}\alpha_{n,2}^{\mathrm{R}}\mathbb{E}[\sin^2\epsilon_n] \\ &= \frac{1}{2}\alpha_{n,1}^{\mathrm{R}}\alpha_{n,2}^{\mathrm{I}}(-2m^2 + 1 + s) - \frac{1}{2}\alpha_{n,1}^{\mathrm{I}}\alpha_{n,2}^{\mathrm{R}}(1 - s),\end{aligned} \tag{70}$$

where $s = \mathbb{E}[\cos(2\epsilon_n)]$.

So, the element in position i, j of $\Sigma_{\delta,\mathrm{RI}}$ is

$$\begin{aligned}&\boldsymbol{\Sigma}_{\delta,\mathrm{RI}}(i,j) \\ &== \frac{1}{2}\sum_{n=1}^{N}\left(\alpha_{n,1}^{\mathrm{R}}\alpha_{n,2}^{\mathrm{I}}(1 + s - 2m^2) - \alpha_{n,1}^{\mathrm{I}}\alpha_{n,2}^{\mathrm{R}}(1 - s)\right).\end{aligned} \tag{71}$$

B. Covariance of the Real Part

We evaluate the covariance matrix $\boldsymbol{\Sigma}_{\delta,\mathrm{R}}$ of the real part of $\boldsymbol{\delta}$ in (34), whose element in position (i, j) is given by

$$\boldsymbol{\Sigma}_{\delta,\mathrm{R}}(i,j) = \mathbb{E}\left[\mathrm{Re}\{\delta_{r_1,c_1}\}\mathrm{Re}\{\delta_{r_2,c_2}\}\right], \tag{72}$$

with $i, j \in \{0, 1, \ldots, N_{\mathrm{A}}N_{\mathrm{B}} - 1\}$, $\delta_{r,c}$ is as in (36), and r_1, c_1, r_2, and c_2 as in (65)-(66). From (3), (34), (36), and (67) we have

$$\begin{aligned}&\mathbb{E}\Big[\mathrm{Re}\{\delta_{r_1,c_1}\}\mathrm{Re}\{\delta_{r_2,c_2}\}\Big] \\ &= \mathbb{E}\Bigg[\sum_{n=1}^{N}\mathrm{Re}\{\alpha_{n,1}(m - e^{j\epsilon_n})\} + \mathrm{Re}\left\{\alpha_{n,1}\frac{D_{n,c_1}}{G_{n,c_1}}\frac{m}{\rho}\sqrt{1-\rho^2}\right\} \\ &\quad \cdot \sum_{n=1}^{N}\mathrm{Re}\{\alpha_{n,2}(m - e^{j\epsilon_n})\} + \mathrm{Re}\left\{\alpha_{n,2}\frac{D_{n,c_2}}{G_{n,c_2}}\frac{m}{\rho}\sqrt{1-\rho^2}\right\}\Bigg] \\ &\quad + \mathbb{E}[\mathrm{Re}\{W_{r_1,c_1}\}\mathrm{Re}\{W_{r_2,c_2}\}],\end{aligned} \tag{73}$$

where $\boldsymbol{D}$ and $\boldsymbol{W}$ are CSCG, independent from all other variables. So, for $c_1 = c_2 = c$, we define

$$\beta = \frac{m^2}{2\rho^2}(1 - \rho^2)\sum_{n}\mathrm{Re}\left\{\frac{\alpha_{n,1}\alpha_{n,2}^*}{|G_{n,c}|^2}\right\}, \tag{74}$$

and we obtain

$$\begin{aligned}&\mathbb{E}\Big[\mathrm{Re}\{\delta_{r_1,c_1}\}\mathrm{Re}\{\delta_{r_2,c_2}\}\Big] \\ &= \mathbb{E}\left[\sum_{n=1}^{N}\mathrm{Re}\{\alpha_{n,1}(m - e^{j\epsilon_n})\}\sum_{n=1}^{N}\mathrm{Re}\{\alpha_{n,2}(m - e^{j\epsilon_n})\}\right] \\ &\quad + \beta\mathbb{1}_{c_1=c_2} + \frac{\sigma^2}{2}\mathbb{1}_{c_1=c_2, r_1=r_2}.\end{aligned} \tag{75}$$

We now look at the product between two generic terms n_1 and n_2 of the two summations in (75), that is

$$\begin{aligned}&\mathbb{E}\Big[\mathrm{Re}\{\alpha_{n_1,1}(m - e^{j\epsilon_{n_1}})\}\mathrm{Im}\{\alpha_{n_2,2}(m - e^{j\epsilon_{n_2}})\}\Big] \\ &= \mathbb{E}[(\alpha_{n_1,1}^{\mathrm{R}}(m - \cos\epsilon_{n_1}) + \alpha_{n_1,1}^{\mathrm{I}}\sin\epsilon_{n_1}) \\ &\quad \cdot (\alpha_{n_2,2}^{\mathrm{R}}(m - \cos\epsilon_{n_2}) + \alpha_{n_2,2}^{\mathrm{I}}\sin\epsilon_{n_2})].\end{aligned} \tag{76}$$

Following the same approach as for (69) we have

$$\begin{aligned}&\mathbb{E}\Big[\mathrm{Re}\{\alpha_{n,1}(m - e^{j\epsilon_n})\}\mathrm{Re}\{\alpha_{n,2}(m - e^{j\epsilon_n})\}\Big] \\ &= \frac{1}{2}\alpha_{n,1}^{\mathrm{R}}\alpha_{n,2}^{\mathrm{R}}(-2m^2 + 1 + s) + \frac{1}{2}\alpha_{n,1}^{\mathrm{I}}\alpha_{n,2}^{\mathrm{I}}(1 - s).\end{aligned} \tag{77}$$

Finally, the element in position (i, j) of $\boldsymbol{\Sigma}_{\delta,\mathrm{R}}$ is

$$\begin{aligned}&\boldsymbol{\Sigma}_{\delta,\mathrm{R}}(i,j) \\ &= \frac{1}{2}\sum_{n=1}^{N}\left(\alpha_{n,1}^{\mathrm{R}}\alpha_{n,2}^{\mathrm{R}}(1 + s - 2m^2) + \alpha_{n,1}^{\mathrm{I}}\alpha_{n,2}^{\mathrm{I}}(1 - s)\right) \\ &\quad + \beta\mathbb{1}_{c_1=c_2} + \frac{\sigma^2}{2}\mathbb{1}_{c_1=c_2, r_1=r_2}.\end{aligned} \tag{78}$$

C. Covariance of the Imaginary Part

We evaluate the covariance matrix $\mathbf{\Sigma}_{\delta,\mathrm{I}}$ of the imaginary part of $\boldsymbol{\delta}$ in (34), whose element in position (i, j) is given by

$$\mathbf{\Sigma}_{\delta,\mathrm{I}}(i,j) = \mathbb{E}\left[\mathrm{Im}\{\delta_{r_1,c_1}\}\mathrm{Im}\{\delta_{r_2,c_2}\}\right], \tag{79}$$

with $i, j \in \{0, 1, \ldots, N_\mathrm{A}N_\mathrm{B} - 1\}$, $\delta_{r,c}$ is as in (36), and r_1, c_1, r_2, and c_2 as in (65)-(66).

Following the same procedure used for computing $\mathbf{\Sigma}_{\delta,\mathrm{R}}$ in the previous paragraph, the element at position (i, j) of $\mathbf{\Sigma}_{\delta,\mathrm{I}}$ is

$$\begin{aligned}\mathbf{\Sigma}_{\delta,\mathrm{I}}(i,j) &= \frac{1}{2}\sum_{n=1}^{N}\left(\alpha^{\mathrm{I}}_{n,1}\alpha^{\mathrm{I}}_{n,2}(1+s-2m^2) + \alpha^{\mathrm{R}}_{n,1}\alpha^{\mathrm{R}}_{n,2}(1-s)\right)\\ &\quad + \beta\mathbb{1}_{c_1=c_2} + \frac{\sigma^2}{2}\mathbb{1}_{c_1=c_2,r_1=r_2}, \end{aligned}\tag{80}$$

where β is defined in (74).

Appendix D
Statistic of $\boldsymbol{\delta}$ for the Attack Through the RIS

In the case, Trudy transmits through the RIS, we have

$$\boldsymbol{\delta} = \boldsymbol{W} + \frac{\sqrt{1-\rho^2}}{\rho}\boldsymbol{H}\boldsymbol{\Phi}'\boldsymbol{D}. \tag{81}$$

The generic entry of the matrix $\boldsymbol{\delta}$ is then

$$\delta_{r,c} = \sum_{n=1}^{N} H_{r,n}e^{j\bar{\theta}_n}e^{j\epsilon_n}\left(\frac{\sqrt{1-\rho^2}}{\rho}D_{n,c}\right) + W_{r,c}. \tag{82}$$

Since the entries of D are CSCG and independent of both G and H, we have $\mathbb{E}[\mathrm{Re}\{\delta_{r,c}\}] = 0$ and $\mathbb{E}[\mathrm{Im}\{\delta_{r,c}\}] = 0$. We now derive the second-order statistics of $\boldsymbol{\delta}$ when the attacker transmits through the RIS. From the definition of $\boldsymbol{\delta}$ and recalling that the entries of $\boldsymbol{D}$ are CSCG and independent of both $\boldsymbol{H}$ and the RIS phase perturbations $\{\epsilon_n\}$, it follows that each $\delta_{r,c}$ is a CSCG random variable.

D. Covariance Matrix

We evaluate the complex covariance matrix $\mathbf{\Sigma}'$ of $\boldsymbol{\delta}$ in (81), whose element in position (i, j) is given by

$$\mathbf{\Sigma}'(i,j) = \mathbb{E}\left[\delta_{r_1,c_1}\delta^*_{r_2,c_2}\right], \tag{83}$$

with $i, j \in \{0, 1, \ldots, N_\mathrm{A}N_\mathrm{B} - 1\}$, $\delta_{r,c}$ is as in (82), and r_1, c_1, r_2, and c_2 as in (65)-(66). Specifically,

$$\begin{aligned}\mathbb{E}[\delta_{r_1,c_1}\delta^*_{r_2,c_2}] &= \frac{1-\rho^2}{\rho^2}\sum_{n=1}^{N}\sum_{n'=1}^{N}H_{r_1,n}H^*_{r_2,n'}e^{j(\bar{\theta}_n-\bar{\theta}_{n'})}\\ &\quad\cdot\mathbb{E}\left[e^{j(\epsilon_n-\epsilon_{n'})}D_{n,c_1}D^*_{n',c_2}\right] + \sigma_B^2\mathbb{1}_{r_1=r_2,c_1=c_2}. \end{aligned}\tag{84}$$

Since the entries of $\boldsymbol{D}$ are iid, only the terms with $n = n'$ contribute, yielding

$$\mathbb{E}[\delta_{r_1,c_1}\delta^*_{r_2,c_2}] = \frac{1-\rho^2}{\rho^2}\sigma_D^2\sum_{n=1}^{N}H_{r_1,n}H^*_{r_2,n} + \sigma_B^2\mathbb{1}_{r_1=r_2,c_1=c_2}. \tag{85}$$

The element in position (i, j) of the covariance of the real and imaginary parts of $\mathrm{vec}(\boldsymbol{\delta})$ is written as

$$\mathbf{\Sigma}'(i,j) = \frac{1-\rho^2}{\rho^2}\sigma_D^2[\boldsymbol{H}\boldsymbol{H}^H]_{r_1,r_2} + \sigma_B^2\mathbb{1}_{r_1=r_2,c_1=c_2}. \tag{86}$$

Note that since $\boldsymbol{\delta}$ is circularly symmetric complex Gaussian, the real and imaginary parts are independent and identically distributed. Therefore, the covariance matrix of $\overline{\boldsymbol{\delta}}$ is given by

$$\mathbf{\Sigma}_{\delta} = \frac{1}{2}\begin{bmatrix}\mathrm{Re}\{\mathbf{\Sigma}'\} & \mathbf{0}\\ \mathbf{0} & \mathrm{Re}\{\mathbf{\Sigma}'\}\end{bmatrix}, \tag{87}$$

with zero cross-covariance between the real and imaginary parts.

Appendix E
Laguerre-Polynomial Expression for P_{MD}

The P_{MD} in (42) can be expressed by a Laguerre polynomial, as derived in [37]. Let $\nu = N_\mathrm{A}N_\mathrm{B}$, $p = \frac{\nu}{2} + 1$, $\mu_0 = \frac{p}{3}$, $y = \frac{\sigma^2\tau}{2}$, $\beta = \frac{1}{n}\sum_{i=1}^{n}\lambda_i(m,s)$, and $x = \frac{(\nu+2)y}{4\beta\mu_0}$. We have

$$\begin{aligned}P_{\mathrm{MD}}(m,s,\tau) &= \frac{e^{-y/(2\beta)}}{(2\beta)^{\nu/2+1}}\frac{y^{\nu/2}}{\Gamma(\nu/2+1)}\\ &\quad\cdot\sum_{k=0}^{K=+\infty}\left(\frac{k!m_k}{\Gamma(\nu/2+1+k)/\Gamma(\nu/2+1)}L_k^{(\nu/2)}(x)\right), \end{aligned}\tag{88}$$

where

$$m_0 = \frac{2\left(\frac{\nu}{2}+1\right)^{\nu/2+1}\beta^{\nu/2+1}}{p-\mu_0}\prod_{i=1}^{n}\left(\beta\mu_0+\lambda_i(p-\mu_0)\right)^{-1/2},$$

$$m_k = \frac{1}{k}\sum_{j=1}^{k}m_{k-j}d_j,\ k \geq 1,$$

$$d_j = \left(-\frac{\mu_0}{p-\mu_0}\right)^j + \sum_{i=1}^{n}\frac{1}{2}\left(\frac{\mu_0(\beta-\lambda_i)}{\beta\mu_0+\lambda_i(p-\mu_0)}\right)^j,\ j \geq 1,$$

and $L_k^{(\nu/2)}(x)$ denotes the generalized Laguerre polynomial of order k with parameter $\nu/2$.

A more efficient way to evaluate the cumulative distribution function (CDF) of a weighted sum of independent chi-square random variables with 1 degree of freedom is Imhof's method [38]. This technique allows for the exact numerical computation of the distribution function by inverting the characteristic function of the sum via numerical integration.

References

[1] G. J. Simmons, "Authentication theory/coding theory," in *Proc. Annu. Int. Cryptol. Conf.* Berlin, Germany: Springer, 1985, pp. 411–431.

[2] S. Tomasin, H. Zhang, A. Chorti, and H. V. Poor, "Challenge-response physical layer authentication over partially controllable channels," *IEEE Commun. Mag.*, vol. 60, no. 12, pp. 138–144, Dec. 2022.

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

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

[5] R. Diamant, P. Casari, and S. Tomasin, "Cooperative authentication in underwater acoustic sensor networks," *IEEE Trans. Wireless Commun.*, vol. 18, no. 2, pp. 954–968, Feb. 2019.

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

[7] H. Fang, X. Wang, and L. Hanzo, "Learning-aided physical layer authentication as an intelligent process," *IEEE Trans. Commun.*, vol. 67, no. 3, pp. 2260–2273, Mar. 2019.

[8] E. Jorswieck, S. Tomasin, and A. Sezgin, "Broadcasting into the uncertainty: Authentication and confidentiality by physical-layer processing," *Proc. IEEE*, vol. 103, no. 10, pp. 1702–1724, Oct. 2015.

[9] N. Xie, Z. Li, and H. Tan, "A survey of physical-layer authentication in wireless communications," *IEEE Commun. Surveys Tuts.*, vol. 23, no. 1, pp. 282–310, 1st Quart., 2021.
[10] R. Meng et al., "Efficient Gaussian process classification-based physical-layer authentication with configurable fingerprints for 6G-enabled IoT," 2023, *arXiv:2307.12263*.
[11] M. A. Shawky et al., "Reconfigurable intelligent surface-assisted cross-layer authentication for secure and efficient vehicular communications," 2023, *arXiv:2303.08911*.
[12] P. Zhang, Y. Teng, Y. Shen, X. Jiang, and F. Xiao, "Tag-based PHY-layer authentication for RIS-assisted communication systems," *IEEE Trans. Dependable Secure Comput.*, vol. 20, no. 6, pp. 4778–4792, Nov. 2023.
[13] M. M. Selim and S. Tomasin, "Physical layer authentication with simultaneous reflecting and sensing RIS," in *Proc. IEEE 97th Veh. Technol. Conf. (VTC-Spring)*, Jun. 2023, pp. 1–5.
[14] D. Shan, K. Zeng, W. Xiang, P. Richardson, and Y. Dong, "PHY-CRAM: Physical layer challenge-response authentication mechanism for wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 9, pp. 1817–1827, Sep. 2013.
[15] X. Wu, Z. Yang, C. Ling, and X.-G. Xia, "Artificial-noise-aided physical layer phase challenge-response authentication for practical OFDM transmission," *IEEE Trans. Wireless Commun.*, vol. 15, no. 10, pp. 6611–6625, Oct. 2016.
[16] Y. Shoukry, P. Martin, Y. Yona, S. Diggavi, and M. Srivastava, "PyCRA: Physical challenge-response authentication for active sensors under spoofing attacks," in *Proc. 22nd ACM SIGSAC Conf. Comput. Commun. Secur.*, Denver, CO, USA, Oct. 2015, pp. 1004–1015.
[17] F. Mazzo, S. Tomasin, H. Zhang, A. Chorti, and H. V. Poor, "Physical-layer challenge-response authentication for drone networks," in *Proc. GLOBECOM - IEEE Global Commun. Conf.*, Dec. 2023, pp. 3282–3287.
[18] S. Tomasin, T. N. M. M. Elwakeel, A. V. Guglielmi, R. Maes, N. Noels, and M. Moeneclaey, "Analysis of challenge-response authentication with reconfigurable intelligent surfaces," *IEEE Trans. Inf. Forensics Security*, vol. 19, pp. 9494–9507, 2024.
[19] L. Crosara, A. V. Guglielmi, N. Laurenti, and S. Tomasin, "Divergence-minimizing attack against challenge-response authentication with IRSs," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops)*, Jun. 2024, pp. 1986–1991.
[20] A. V. Guglielmi, L. Crosara, S. Tomasin, and N. Laurenti, "Physical-layer challenge-response authentication with IRS and single-antenna devices," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops)*, Jun. 2024, pp. 560–565.
[21] A. Almohamad et al., "Smart and secure wireless communications via reflecting intelligent surfaces: A short survey," *IEEE Open J. Commun. Soc.*, vol. 1, pp. 1442–1456, 2020.
[22] L. Dong and H.-M. Wang, "Secure MIMO transmission via intelligent reflecting surface," *IEEE Wireless Commun. Lett.*, vol. 9, no. 6, pp. 787–790, Jun. 2020.
[23] H. Niu, Z. Chu, F. Zhou, and Z. Zhu, "Simultaneous transmission and reflection reconfigurable intelligent surface assisted secrecy MISO networks," *IEEE Commun. Lett.*, vol. 25, no. 11, pp. 3498–3502, Nov. 2021.
[24] H. Niu et al., "Joint beamforming design for secure RIS-assisted IoT networks," *IEEE Internet Things J.*, vol. 10, no. 2, pp. 1628–1641, Jan. 2023.
[25] Z. Zhang, J. Chen, Y. Liu, Q. Wu, B. He, and L. Yang, "On the secrecy design of STAR-RIS assisted uplink NOMA networks," *IEEE Trans. Wireless Commun.*, vol. 21, no. 12, pp. 11207–11221, Dec. 2022.
[26] Y. Han, N. Li, Y. Liu, T. Zhang, and X. Tao, "Artificial noise aided secure NOMA communications in STAR-RIS networks," *IEEE Wireless Commun. Lett.*, vol. 11, no. 6, pp. 1191–1195, Jun. 2022.
[27] C. Lipps et al., "Reconfigurable intelligent surfaces: A physical layer security perspective," in *Proc. Int. Conf. Data Intell. Security (ICDIS)*, 2022, pp. 174–181.
[28] P. Jin, X. Yang, Z. Cheng, and J. Li, "Multi-functional RIS-assisted semantic anti-jamming communication and computing in integrated aerial-ground networks," *IEEE J. Sel. Areas Commun.*, vol. 42, no. 10, pp. 2785–2802, Oct. 2024.
[29] H. Zhang, W. Guo, Y. Liu, and L. Song, "Dual-polarized stacked metasurface transceiver design with rate splitting for next generation wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 42, no. 9, pp. 2420–2435, Sep. 2024.
[30] H. Zhou, M. Erol-Kantarci, Y. Liu, and H. V. Poor, "A survey on model-based, heuristic, and machine learning optimization approaches in RIS-aided wireless networks," *IEEE Commun. Surveys Tuts.*, vol. 26, no. 2, pp. 781–823, 2nd Quart., 2024.
[31] A. V. Guglielmi and S. Tomasin, "Joint RIS optimization and channel estimation with decision tree-based adaptive reconfiguration," *IEEE Trans. Commun.*, vol. 73, no. 6, pp. 4049–4060, Jun. 2025.
[32] D. Wijekoon, A. Mezghani, and E. Hossain, "Phase shifter optimization in RIS-aided MIMO systems under multiple reflections," *IEEE Trans. Wireless Commun.*, vol. 23, no. 8, pp. 8969–8983, Aug. 2024.
[33] A. V. Guglielmi and S. Tomasin, "Fast iterative configuration of reconfigurable intelligent surfaces in mmWave systems," in *Proc. GLOBECOM - IEEE Global Commun. Conf.*, Dec. 2023, pp. 631–636.
[34] S. M. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory*. Upper Saddle River, NJ, USA: Prentice-Hall, 1993.
[35] N. Benvenuto, G. Cherubini, and S. Tomasin, *Algorithms for Communications Systems and Their Applications*. Hoboken, NJ, USA: Wiley, 2020.
[36] S. J. Press, "Linear combinations of non-central chi-square variates," *Ann. Math. Statist.*, vol. 37, no. 2, pp. 480–487, Apr. 1966, doi: 10.1214/aoms/1177699531.
[37] A. Castaño-Martínez and F. López-Blázquez, "Distribution of a sum of weighted noncentral chi-square variables," *TEST*, vol. 14, no. 2, pp. 397–415, Dec. 2005.
[38] J. P. Imhof, "Computing the distribution of quadratic forms in normal variables," *Biometrika*, vol. 48, nos. 3–4, pp. 419–426, 1961.

Anna Valeria Guglielmi (Member, IEEE) received the B.Sc., M.Sc., and Ph.D. degrees in information engineering from the University of Padova, Italy, in 2012, 2014, and 2018, respectively. In 2017, she was a Visiting Scientist with the BIOTEC Technische Universität Dresden (TUD), Germany, and the University of California, Irvine. She is currently an Assistant Professor with the University of Padova. Her current research interests include machine-learning architectures and signal processing for wireless communication systems and physical layer security.

Laura Crosara (Member, IEEE) received the B.Sc. degree in information engineering, the M.Sc. degree in telecommunications engineering, and the Ph.D. degree in information engineering from the University of Padova, Padua, Italy, in 2019, 2021, and 2024, respectively. She has been a Visiting Scientist with the ESA European Space Research and Technology Centre, Noordwijk, The Netherlands. She is currently an Adjunct Professor with the University of Padova. Her research interests include authentication techniques for global navigation satellite systems, physical layer security, and wireless communications.

Stefano Tomasin (Senior Member, IEEE) received the Ph.D. degree from the University of Padova, Italy, in 2003, where he is currently a Full Professor. During his career, he has visited IBM Research, Switzerland, Philips Research, The Netherlands, Qualcomm, California, the Polytechnic University in Brooklyn, New York, and Huawei, France. His current research interests include physical layer security, security of global navigation satellite systems, signal processing for wireless communications, synchronization, and scheduling of communication resources. He is a member of EURASIP. He is or has been an Editor of IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGIES, IEEE TRANSACTIONS ON SIGNAL PROCESSING from 2017 to 2020, the *EURASIP Journal of Wireless Communications and Networking*, and IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY.

Open Access funding provided by 'Universita degli Studi di Padova' within the CRUI CARE Agreement