Analysis of Challenge-Response Authentication With Reconfigurable Intelligent Surfaces

Stefano Tomasin, *Senior Member, IEEE*, Tarek N. M. M. Elwakeel, Anna Valeria Guglielmi, *Member, IEEE*, Robin Maes, Nele Noels, *Senior Member, IEEE*, and Marc Moeneclaey, *Life Fellow, IEEE*

***Abstract*— Physical-layer authentication (PLA) mechanisms exploit signals exchanged at the physical layer of communication systems to confirm the sender of a received message. In this paper, we propose a novel challenge-response PLA (CR-PLA) mechanism for a cellular system that leverages the reconfigurability property of a reconfigurable intelligent surface (RIS) (under the control of the verifier) in an authentication mechanism. In CR-PLA, the verifier base station (BS) sets a random RIS configuration, which remains secret to the intruder, and then checks that the resulting estimated channel is modified correspondingly. In fact, for a message sent by an attacker in a different location than the legitimate user equipment (UE), the BS will estimate a different channel and the message will be rejected as fake. Such a solution reduces the communication and computational overhead with respect to higher-layer cryptographic authentication. We derive the maximum a-posteriori attack when the attacker observes a correlated channel and the reconfigurable intelligent surface (RIS) has many elements, and the attacker transmits to Bob either directly or through the RIS. Using a generalized likelihood ratio test to test the authenticity at the base station (BS), we derive approximate expressions of the false alarm and misdetection probabilities when both the BS and the UE have a single antenna each, while the RIS has a large number of elements. We also evaluate the trade-off between security and communication performance, since choosing a random RIS configuration reduces the data rate. Moreover, we investigate the impact of various parameters (e.g., the RIS randomness, the number of RIS elements, and the operating signal-to-noise ratio) on security and communication performance.**

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

I. Introduction

DETERMINING if a received message is coming from its claimed sender, i.e., establishing its *authenticity* is a key security problem in communication systems. The current and future networks will include several interconnected devices with diversified energy and computational constraints, and physical-layer authentication (PLA) is an attractive security solution (see surveys [1] and [2]) since it requires simple signal processing capabilities and exploits existing signals without introducing communication overhead. In tag-based PLA the channel operates as a tag: the receiver authenticates newly received messages that appear to have traveled through the same channel as those (authentic) received in the past. When an attacker transmits from another location the resulting channel is different from that of the legitimate transmitter and is detected as fraudulent.

Since PLA is based on channel characteristics, devices that enable the manipulation of propagation properties should be considered to enhance security. To this end, a reconfigurable intelligent surface (RIS) is an interesting component, as it includes several reflective elements, each introducing a controllable phase shift in the equivalent baseband reflected signal [3], [4], [5]. RISs have also been considered for PLA to increase the signal-to-noise ratio (SNR) and improve the authentication process. However, the possibility of reconfiguring RISs also paves the way for a new PLA procedure, called challenge-response (CR) PLA, first introduced in [6]. In CR-PLA, the receiver first randomly modifies the propagation environment (which represents the challenge) and then estimates the channel through which the received signal has passed (which represents the response) to verify that it matches the modified environment. For an attacker that does not know the current challenge (i.e., the current RIS configuration) it will be harder to perform an effective authentication attack than in the PLA setting.

While RIS-based CR-PLA is a promising security solution, its performance has not been investigated in the literature. This study should include not only the security performance – in

Received 16 April 2024; revised 27 July 2024; accepted 21 September 2024. Date of publication 30 September 2024; date of current version 14 October 2024. This work was supported in part by the Project Innovative Security Paradigms for beyond 5G (ISP5G+), which is part of the Security and Rights in the Cyberspace (SERICS) Program under the National Recovery and Resilience Plan (NRRP) Ministero dell'Università e della Ricerca (MUR) Program funded by the European Union (EU)-NGEU under Grant CUP D33C22001300002 and Grant PE00000014; in part by European Commission through the Horizon Europe/Joint-Undertaking (JU) Smart Networks and Services (SNS) Project ROBUST-6G under Grant101139068; and in part by European Union's NextGenerationEU Instrument under Italian National Recovery and Resilience Plan (NRRP), Mission 4, Component 2, Investment 1.3, enlarged partnership "Telecommunications of the Future," Program "RESTART" under Grant PE0000001. The associate editor coordinating the review of this article and approving it for publication was Prof. Haijun Zhang. *(Corresponding author: Stefano Tomasin.)*

Stefano Tomasin is with the Department of Information Engineering and the Department of Mathematics, University of Padova, 35122 Padua, Italy, and also with Consorzio Interuniversitario delle Telecomunicazioni (CNIT), 43124 Parma, Italy (e-mail: stefano.tomasin@unipd.it).

Tarek N. M. M. Elwakeel was with the Department of Information Engineering, University of Padova, 35122 Padua, Italy, and also with HPE, 36050 Bolzano Vicentino, Italy.

Anna Valeria Guglielmi is with the Department of Information Engineering, University of Padova, 35122 Padua, Italy.

Robin Maes was with the Department of Telecommunications and Information Processing, Ghent University, 9000 Ghent, Belgium.

Nele Noels and Marc Moeneclaey are with the Department of Telecommunications and Information Processing, Ghent University, 9000 Ghent, Belgium.

Digital Object Identifier 10.1109/TIFS.2024.3471185

terms of false alarm (FA) and misdetection (MD) probabilities – but also communication performance, since the choice of the RIS configuration has an impact on the achieved data rates. Moreover, the ability to withstand advanced attacks that exploit the partial channel knowledge of the attacker is still to be investigated. By a thorough analysis of the RIS-based CR-PLA it will be possible not only to assess the merits of this security solution but also to design it properly, i.e., to select the RIS size and the randomness.

To address these issues, in this paper, we consider an RIS-supported CR-PLA mechanism for cellular networks, where a base station (BS) verifies the authenticity of messages received from a user equipment (UE). The CR-PLA procedure includes two stages. In a preliminary stage, the UE transmits a sequence of pilot samples (properly authenticated by a higher-layer procedure) to the BS via the RIS with several configurations, and the BS estimates the corresponding UE-RIS-BS cascaded channels. This will enable the BS to predict the cascaded channel under any other RIS configuration. In the second stage, aiming at authenticating a message potentially coming from the UE, the BS randomly chooses a RIS configuration while a new message is transmitted. The BS then compares the channel estimated from the received signal with that predicted for the selected RIS configuration, using the information obtained in the preliminary stage. From this comparison, the BS decides on the message authenticity.

We consider a generalized likelihood ratio test (GLRT) to decide about the authenticity of the message and analyze the performance of the CR-PLA scheme in terms of both FA and MD probabilities. Note that the random RIS configuration also affects the data rate of the communication link between the UE and the BS. To limit the rate loss, we restrict the random selection of each phase shift of the RIS to an angular sector centered around the phase shift that maximizes the data rate, and we investigate the security performance as a function of the size of the angular sector. We also derive the maximum a-posteriori (MAP) attack to be used when the attacker knows his channel to the legitimate UE and this channel is partially correlated with that from the legitimate UE to the RIS. The attacker can transmit the attack signal either directly to the BS or through the RIS. In the latter case, the attacker signal will also be determined by an instantaneous random RIS configuration. Approximate expressions for the FA and MD probabilities and the data rate are obtained when both the UE and the BS are equipped with a single antenna, and the RIS has a large number of elements.

The main contributions of this paper are:

1) The study of the MAP attack against the proposed authentication mechanism when the attacker transmits to Bob either directly or through the RIS: such attack exploits at best the partial information available to the attacker;
2) The derivation of FA probability for the RIS-based CR-PLA;
3) The derivation of MD probability for the RIS-based CR-PLA for the MAP attack; this provides an assessment of the security performance under an effective attack;
4) The derivation of the average SNR for communication purposes for the RIS-based CR-PLA; the two performance metrics of contributions 1. to 3. enable an assessment of the communication-security trade-off under legitimate conditions;
5) An investigation of the impact of various parameters (e.g., the RIS randomness, the number of RIS elements, the operating SNR) on the security and communication performance.

The proposed analysis offers an in-depth vision of the RIS-based CR-PLA and its security-communication trade-off.

By simulating both the CR-PLA and the tag-based PLA, we show that the proposed CR-PLA provides higher security (i.e., a lower MD probability) than tag-based PLA for a given data rate loss. We observe that the proposed CR-PLA solution is particularly effective when the attacker channel is highly correlated with the channel of the legitimate UE, which is a critical case for tag-based PLA.

The rest of the paper is organized as follows. In the next subsection, we consider related works in the literature and further highlight the contribution of our work. Section II introduces both the system model (including the assumptions on the attacker) and the communication performance metric. The CR-PLA mechanism is then described in Section III, where the GLRT is detailed. In Section IV the security performance metrics are defined and the MAP attack strategy is introduced. The scenario with single-antenna devices is studied in Section V. Section VI presents and discusses numerical results for the considered RIS-supported CR-PLA scenario. The main conclusions are drawn in Section VII.

A. Related Works

PLA has been investigated for several years, and the reader can refer to surveys (e.g., [1] and [2]) for an overview. Here we focus on CR-PLA (and similar solutions) and the use of RISs for authentication.

CR-PLA has been first introduced in [6], and its essential feature is that random modification of the propagation environment by the verifier plays the role of the challenge, while the corresponding estimated channel at the verifier plays the role of the response. Another scheme with the same name provides that instead of the channel itself, a confidential message (transmitted by the device to be verified) is used for authentication purposes, in [7]. This solution was further investigated in [8] (adding artificial noise) and in [9], where actuators continuously challenge the surrounding environment with random transmissions detected by other sensors. However, again, the challenge is provided by the transmitted message rather than the propagation environment itself. The analysis proposed in our paper differs from [7], [8], and [9] not only because we consider a RIS, but also because our challenge in CR-PLA the random modification of the propagation environment by the receiver, which sets random RIS configurations. Note that the CR-PLA technique has also been proposed to authenticate transmissions from

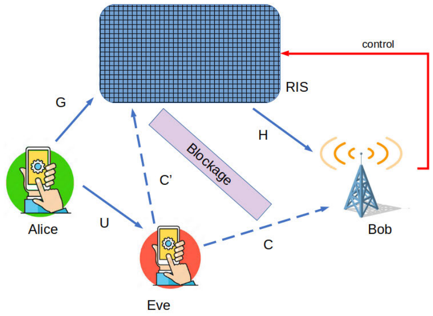

Fig. 1. Model of the RIS-supported communication system.

drones operating in swarms, where the challenge is defined by the movement of trusted drones acting as receivers [10].

CR-PLA has some commonalities with physically unclonable function (PUF)-based authentication. However, a PUF exploits hardware differences between individual chips for unique device identification [11], [12], [13] and the transmitter (rather than the propagation environment as in CR-PLA) is randomly modified (see [14] for a detailed comparison between PUFs and PLA).

With particular reference to the integration of RISs and PLA, optimizing the RIS to collect signals from a specific location makes the resulting transmitter-RIS-receiver channel more peculiar, improving, for instance, the tag-based PLA [15]. A tag-based PLA scheme exploiting a RIS in non-line-of-sight conditions has been proposed in [16] for a vehicular ad-hoc network scenario showing that the RIS improves the authentication accuracy. In [17] a tag-based PLA mechanism in RIS-supported communication systems has been proposed: in a RIS communication system, the channel gain and background noise are first extracted; then, these features are applied to a random signal together with a private key to generate a cover tag signal. Since previously known secret bits rather than physical layer features are used for authentication, the investigations in [16] and [17] significantly differ from the analysis proposed in this paper. Moreover, beyond the security performance, we also investigate the communication, as the data rate is affected by the RIS configuration randomization in CR PLA. It is worth mentioning that advanced RIS with sensing capabilities can also be exploited to provide a more detailed knowledge of the propagation channel, yielding a better authentication [18], still based on the tag concept. However, we consider a passive RIS whose configuration only introduces phase shifts on the transmitted signals. The authors have also considered the same scenario to propose other attacks in [19] and the optimization of the RIS randomness in [20].

Finally, beyond authentication, RIS can also support other security mechanisms, e.g., confidentiality [21], [22], [23], [24], [25], [26] and anti-jamming [27]. However, they are out of the scope of this paper.

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

II. System Model

Consider the scenario shown in Fig. 1, where a BS (Bob) wants to authenticate a UE (Alice), and the signal transmitted by the UE reaches the BS via a RIS. Another device, Eve, attempts to impersonate Alice, i.e., to transmit messages that Bob will receive as coming from Alice. In turn, Bob will deploy authentication mechanisms to distinguish messages coming from Alice and Eve. A blockage prevents direct communication between Alice and Bob. Eve transmits signals either directly to Bob or through the RIS. The first setting (direct transmission from Eve to Bob) particularly highlights the role of the RIS in the CR-PLA mechanism, as further discussed in Section III-B.

Alice, Bob, and Eve are equipped with K, M, and V antennas, respectively, and transmit narrowband signals with the same carrier frequency; all channels are assumed to be reciprocal. The RIS has N reflecting elements, each acting as a receive and transmit antenna; in particular, element $n = 0, 1, \ldots, N-1$, introduces an additional phase shift $\phi_n = e^{j\theta_n}$ between the impinging and reflected baseband signals. We define the diagonal $N \times N$ matrix

$$\boldsymbol{\Phi} = \mathrm{diag}\{\boldsymbol{\phi}\} = \mathrm{diag}\{\phi_0, \ldots, \phi_{N-1}\}, \tag{1}$$

containing all the phase shifts introduced by the RIS. The RIS is under the control of Bob, who can set the values of θ_n, $n = 0, 1, \ldots, N-1$, using a secure dedicated channel (typically a wire or a fiber).

Let $\boldsymbol{G}$ be the $N \times K$ baseband equivalent narrowband channel matrix from Alice to the RIS, and $\boldsymbol{H}$ be the $M \times N$ matrix of the channel from the RIS to Bob. Thus, the resulting Alice-RIS-Bob channel matrix is

$$\boldsymbol{Q}^{(A,I,B)} = \boldsymbol{H}\boldsymbol{\Phi}\boldsymbol{G}. \tag{2}$$

Channels connecting Eve directly to Bob and the RIS are described by matrices $\boldsymbol{C}$ $\boldsymbol{C}'$), respectively. All channels ($\boldsymbol{H}$, $\boldsymbol{G}$, $\boldsymbol{C}$, and $\boldsymbol{C}'$) are time-invariant, while the RIS configuration can be changed over time so that the resulting cascaded channels are controllable.

Signals received by Bob are also affected by complex-valued additive white Gaussian noise (AWGN) with zero mean, and independent per antenna.

A. Communication-Optimal RIS Configuration

Since the RIS is also used for communication purposes between Alice and Bob, Bob should configure it with this functionality in mind.

We indicate the RIS configuration maximizing the spectral efficiency (or *communication-optimal RIS configuration*) as

$$\bar{\boldsymbol{\Phi}} = \mathrm{diag}\{e^{j\bar{\theta}_0}, \ldots, e^{j\bar{\theta}_{N-1}}\}, \tag{3}$$

where $\bar{\theta}_n$, $n = 0, \ldots, N-1$, are the communication-optimal phase shifts of the N elements of the RIS. Recently, various approaches to optimize the RIS configuration have been explored, including those based on electromagnetic and hardware features [28], [29], energy and efficiency [3], [30], multi-user and MIMO communication [29], [31], and communication-theoretic formulations [30]. We refer to these papers for the computation of $\bar{\theta}_n$, as the technique to optimize the RIS configuration is not in the scope of this work.

B. Assumptions on Eve

We assume that Eve:

- has at least the same number of antennas as Bob ($V \geq M$);
- perfectly knows channel matrices $\boldsymbol{C}$ and $\boldsymbol{C}'$;
- knows the pilot signals transmitted by Alice and used by Bob to estimate the channel;
- does not know the current RIS configuration $\boldsymbol{\Phi}$;
- has unlimited transmit power.

Moreover, channels $\boldsymbol{C}$ and $\boldsymbol{C}'$ are full-rank, which can easily be obtained when Eve has many antennas. In these conditions, Eve can precode the pilot signals transmitted to Bob with a suitable precoding matrix $\boldsymbol{B}$, so that the equivalent channel (the cascade of the precoder and $\boldsymbol{C}$) seen by Bob when Eve transmits directly to Bob is any $M \times K$ matrix $\boldsymbol{Z} = \boldsymbol{C}\boldsymbol{B}$. Similarly, when Eve transmits to the RIS, she can let Bob estimate any channel to the RIS, and the resulting cascaded channel will be $\boldsymbol{Z} = \boldsymbol{H}\boldsymbol{\Phi}\boldsymbol{C}'\boldsymbol{B}$.

In the following, we also assume that Alice and Bob neither know the instantaneous channels with Eve nor their statistics. In particular, Alice and Bob do not know where Eve is, so they cannot infer anything about the propagation of signals transmitted or received by Eve. This assumption is very conservative from a security point of view, while in practice Alice and Bob may know something about Eve, up to the extreme case where Eve is another UE served by the same BS.

C. Reference Scenario

In the rest of the paper, we will also consider a reference scenario, as an example. In the reference scenario, matrices $\boldsymbol{H}$ and $\boldsymbol{G}$ have independent random entries (within each matrix and between the two matrices), and each entry is assumed to be complex-valued Gaussian zero-mean with unit variance (Rayleigh fading model). We also assume that Eve has as many antennas as the RIS elements ($V = N$) and channels ($\boldsymbol{U}$ and $\boldsymbol{C}$ from Fig. 1) are correlated to the channels $\boldsymbol{G}$ and $\boldsymbol{H}$, respectively. Thus Eve can extract partial information on the two channels. In particular, we assume that Eve can either have *partial* or *perfect* channel knowledge.

1) Partial Channel Knowledge: Eve is assumed to perfectly know the RIS-Bob channel $\boldsymbol{H}$. Although generous, this can be motivated by the position of Eve. In fact, we assume that Eve is located in the vicinity of Bob, making it more reasonable for her to properly estimate the RIS-Bob channel. Furthermore, Eve is assumed to have partial information on $\boldsymbol{G}$ through $\boldsymbol{U}$, that is

$$\boldsymbol{U} = \rho \boldsymbol{G} + \sqrt{1-\rho^2}\boldsymbol{D}, \tag{4}$$

where $\rho \in [0, 1]$ is the correlation factor between channels $\boldsymbol{U}$ and $\boldsymbol{G}$; the entries of $\boldsymbol{D}$ are independent of those of both $\boldsymbol{G}$ and $\boldsymbol{H}$, independent among each other and complex-valued Gaussian distributed with zero mean and unit variance.

2) Perfect Channel Knowledge: Eve is assumed to perfectly know the Alice-RIS and RIS-Bob channel matrices $\boldsymbol{H}$ and $\boldsymbol{G}$. This assumption is very generous to Eve because Eve is neither co-located with Alice nor Bob, and these channels are only used in cascade via the RIS. Estimating the cascaded Alice-RIS-Bob channel is relatively easy for Alice and Bob, while it's harder for them (and even more so for Eve) to estimate the individual channels, i.e., $\boldsymbol{H}$ and $\boldsymbol{G}$. Considering this kind of knowledgeable attacker will therefore result in a conservative estimate of the security performance constituting a worst-case situation for the legitimate receiver. Note that the perfect channel knowledge scenario is obtained from (4) by assuming $\rho = 1$. On the opposite, when $\rho = 0$ Eve does not have any information at all on $\boldsymbol{G}$.

Note that this scenario is intended to capture the relation between the information by Eve on the channels to the RIS and the performance of CR-PLA in such a scenario. More realistic scenarios with different knowledge and different characteristics of Eve (e.g., the number of her antennas) are left for future studies. Note also that the knowledge of the channel to and from the RIS can be obtained more easily when assuming a geometric channel model with a small number of paths, which is not considered in this paper.

The assumption on $\boldsymbol{U}$ is quite generous to Eve, considering the distance between Alice and Eve (see Fig. 1). However, the probability of effective attacks will be even lower when this assumption is not verified and Eve does not know anything about $\boldsymbol{H}$ and $\boldsymbol{G}$. In this case, the proposed CR-PLA mechanism will be even more effective. Assuming some correlation will provide a more extensive security assessment.

3) Single Antenna Devices: As a particular case, we will also consider Alice and Bob with a single antenna each; while Eve will still be equipped with many antennas. In this scenario, a closed-form solution for the communication-optimal RIS configuration is available, enabling the security analysis of the authentication mechanism.

III. Challenge-Response Authentication

In this Section, we will first introduce the CR-PLA mechanism, describing in detail its steps. In the following we will mostly focus on the case in which Eve transmits directly to Bob (and not through the RIS) since in this case the CR-PLA mechanism is more effective. In Section IV-B we will also consider the attack through the RIS.

A. The CR-PLA Mechanism

The CR-PLA mechanism using the RIS comprises the following steps:

1) *Step 1, CSI measurements:* Bob first estimates the cascade Alice-RIS-Bob channel for a *finite set* of RIS configurations that will enable him to predict the Alice-RIS-Bob channel for *any* RIS configuration. Several techniques are available to this end in the literature (see [32] for an extensive survey). In this step, transmissions are authenticated at a higher layer to ensure

that Bob is estimating the channel from Alice (rather than a channel induced by Eve).

2) *Step 2, random configuration:* Bob chooses a random RIS configuration $\mathbf{\Phi}'$, as described in Section III-B. This constitutes the *challenge* posed to the device to be authenticated.

3) *Step 3, message transmission:* either Alice or Eve (impersonating Alice) transmits a message and Bob obtains the estimate $\hat{\boldsymbol{Q}}$ of the channel from the received signal: this represents the *response*. When Alice is transmitting, Bob obtains the channel estimate

$$\hat{\boldsymbol{Q}} = \boldsymbol{Q}^{(A,I,B)}(\mathbf{\Phi}') + \boldsymbol{W}'', \tag{5}$$

where $\boldsymbol{W}''$ is the estimation error matrix at Bob, modeled as AWGN with zero mean, independent entries, each with variance σ_{B}^2. When Eve is transmitting, Bob obtains instead the estimate

$$\hat{\boldsymbol{Q}} = \boldsymbol{Z} + \boldsymbol{W}''. \tag{6}$$

4) *Step 4, channel verification:* Bob predicts the cascaded channel from Alice with configuration $\mathbf{\Phi}'$ selected in Step 2, using the knowledge acquired in Step 1,

$$\overline{\boldsymbol{Q}}^{(A,I,B)}(\mathbf{\Phi}') = \boldsymbol{H}\mathbf{\Phi}'\boldsymbol{G} + \overline{\boldsymbol{W}}, \tag{7}$$

where $\overline{\boldsymbol{W}}$ is the prediction error, still a zero-mean AWGN matrix with independent entries, each with variance σ_{B}^2. Bob compares $\overline{\boldsymbol{Q}}^{(A,I,B)}(\mathbf{\Phi}')$ with $\hat{\boldsymbol{Q}}$, as better specified in Section III-C. If they match, the message is accepted as authentic, otherwise, it is rejected as fake.

Note that Steps 2-4 are repeated upon the transmission of each message, and the random RIS configuration is selected independently at each message transmission.

Some remarks are due on the considered mechanism:

Remark 1: Step 1 requires that the initial transmissions are authenticated with other mechanisms than CR-PLA. This is an assumption in common with other authentication schemes (including tag-based PLA) since the verifier must establish a correspondence between the observed signal and the authentic transmitter.

Remark 2: The presence of the RIS is essential in CR-PLA, since it allows the channel randomization for authentication in Step 2 that, together with the verification Step 4, enables the CR mechanism: Bob poses a challenge (the RIS configuration) to which Alice responds by transmitting over the expected channel as then verified by Bob in Step 4.

Remark 3: Note that if Eve can transmit signals directly to the RIS (though channel $\boldsymbol{C}'$, the random modification of the RIS configuration would be less useful for CR-PLA, since it would affect the signals received by Bob even when Eve is transmitting.

In this Section, we provide further details for the two key steps 2. and 4. of the CR-PLA mechanism.

B. RIS Configuration

RISs are typically used to enhance communication by providing better coverage and higher rates. Introducing randomness in its configuration in Step 2 will worsen the communication performance with respect to that obtained with the communication-optimal configuration (3). To control this performance loss, we consider a uniform random distribution[1] of the phase shifts in a sector of angular width 2γ around the communication-optimal values $\bar{\theta}_n$, $n = 0, \dots, N-1$, i.e.,

$$\theta_n = \bar{\theta}_n + \epsilon_n, \quad \epsilon_n \simeq \mathcal{U}[-\gamma, +\gamma]. \tag{8}$$

Indeed, a key point in CR-PLA is that a random perturbation on the channel is introduced by Bob, who then verifies that the received signal passed through the perturbed channel. By using a random RIS configuration at each transmission we make it harder for Eve to induce the same channel. However, while providing authentication capabilities, this randomness also affects the data rate of the communication link between Alice and Bob since we are not using the communication-optimal RIS configuration.

Hence, (8) determines a trade-off between the resulting achievable rate of the legitimate channel (obtained when the communication-optimal configuration is used) and the security of the authentication mechanism. Parameter γ controls the randomness of the RIS configuration determining at the same time the deviation from the communication-optimal RIS configuration. On the one hand, a larger γ pushes the RIS configuration further away from the communication-optimal configuration and results in a reduction of the average data rate. On the other hand, a larger γ introduces more randomness into the RIS configuration, making it more difficult for Eve to pass the authentication test. Indeed, the maximum randomness is achieved with $\gamma = \pi$, while the communication-optimal configuration is achieved with $\gamma = 0$. However, when $\gamma = 0$ the RIS is fixed, and once Eve knows one cascaded channel (with the unique RIS configuration), all future attacks will be successful. Indeed, in this case, CR-PLA boils down to the traditional PLA.

Remark 4: Note that in this paper we focus on the evaluation of the trade-off between security and communication performance as a function of γ, i.e., the tunable parameter of the probability density function of ϵ_n (p_{ϵ_n}). In [20] instead, we address the issue of designing the probability distribution of the random RIS configuration maximizing the average receiver SNR under an upper bound constraint on the MD and FA probabilities.

C. Channel Verification

We now focus on Step 4 for channel verification. This step is critical for authentication since now Bob decides on the response. We define two hypotheses

- $\mathcal{H}_0$: the response comes from Alice and (5) holds;
- $\mathcal{H}_1$: the response is not from Alice and (6) holds.

To decide between $\mathcal{H}_0$ and $\mathcal{H}_1$, Bob cannot perform a standard likelihood ratio test (LRT), because both $\boldsymbol{Q}$ and $\boldsymbol{Z}$ are unknown to Bob. Instead, a GLRT is performed, which involves replacing $\boldsymbol{Q}$ and $\boldsymbol{Z}$ in the standard LRT by their maximum likelihood estimates [33, Ch. 6]. The GLRT for

[1] Other options for the distribution have been considered in [20].

the considered problem is derived in Appendix A, and can be summarized as

$$\Psi < \tau : \hat{\mathcal{H}} = \mathcal{H}_0, \tag{9a}$$

$$\Psi \geq \tau : \hat{\mathcal{H}} = \mathcal{H}_1, \tag{9b}$$

where Ψ is defined as

$$\Psi = \frac{2}{\sigma^2} \sum_{m=0}^{KM-1} |\mathrm{vec}(\hat{\boldsymbol{Q}})_m - \mathrm{vec}(\overline{\boldsymbol{Q}}^{(A,I,B)}(\boldsymbol{\Phi}'))_m|^2, \tag{10}$$

with $\sigma^2 = 2\sigma_B^2$ denoting the variance of any element of matrix

$$\boldsymbol{W} = \boldsymbol{W}'' - \overline{\boldsymbol{W}}. \tag{11}$$

Remark 5: The CR-PLA (as the tag-based PLA) mechanism requires that channels $\boldsymbol{H}$ and $\boldsymbol{G}$ do not change significantly among different message transmissions. In fact, the GLRT uses the information on the channel estimated in Step 1 to infer the channel with the current RIS configuration. When channels $\boldsymbol{G}$ and $\boldsymbol{H}$ are time-varying, channel prediction techniques can be used to obtain an updated estimate of the channels to be used in the GLRT. The analysis of such a mechanism is left for future investigation.

IV. Security Performance Analysis

In this section we analyze the security performance of the RIS-based CR-PLA scheme. First, we derive the security metrics, namely the FA and MD probabilities. Then, we derive the MAP attack in the reference scenario.

A. Security Metrics

Authentication, as a binary hypothesis test, yields two possible error events, namely MD and FA. An FA occurs when deciding that the received message is fake while it is authentic; an MD occurs when a message transmitted by Eve is accepted as coming from Alice. In formulas, an FA occurs when, under hypothesis $\mathcal{H}_0$, $\Psi \geq \tau$. An MD occurs instead when, under hypothesis $\mathcal{H}_1$, $\Psi < \tau$. The FA probability P_{FA} and MD probability P_{MD} are then

$$P_{\mathrm{FA}} = P[\Psi \geq \tau | \mathcal{H}_0], \tag{12a}$$

$$P_{\mathrm{MD}} = P[\Psi < \tau | \mathcal{H}_1]. \tag{12b}$$

For messages transmitted by Alice, using (10) we have

$$\Psi = \frac{2}{\sigma^2} \sum_{m=0}^{KM-1} |\mathrm{vec}(\boldsymbol{W})_m|^2 , \tag{13}$$

thus Ψ is a central chi-square random variable with $2KM$ degrees of freedom and

$$P_{\mathrm{FA}} = 1 - F_{\chi^2,0}(\tau), \tag{14}$$

where $F_{\chi^2,a}(\cdot)$ is the cumulative distribution function (CDF) of a non-central chi-square random variable with $2KM$ degrees of freedom and non-centrality parameter a.

Under attack instead, by inserting (6) in (10), under the hypothesis $\mathcal{H}_1$ with attack channel $\boldsymbol{Z}$ and RIS configuration $\boldsymbol{\Phi}'$, we obtain

$$\begin{aligned}\Psi = \frac{2}{\sigma^2} \sum_{m=0}^{KM-1} &|\mathrm{vec}(\boldsymbol{Z})_m + \mathrm{vec}(\boldsymbol{W}'')_m \\ &- \mathrm{vec}(\overline{\boldsymbol{Q}}^{(A,I,B)}(\boldsymbol{\Phi}'))_m|^2 \\ = \frac{2}{\sigma^2} \sum_{m=0}^{KM-1} &|\mathrm{vec}(\boldsymbol{Z})_m + \mathrm{vec}(\boldsymbol{W})_m - \mathrm{vec}(\boldsymbol{H}\boldsymbol{\Phi}'\boldsymbol{G})_m|^2.\end{aligned} \tag{15}$$

In this case, Ψ is a non-central chi-square random variable with $2KM$ degrees of freedom and non-centrality parameter

$$\zeta = \frac{2}{\sigma^2} \left|\left|\boldsymbol{Z} - \boldsymbol{H}\boldsymbol{\Phi}'\boldsymbol{G}\right|\right|^2, \tag{16}$$

and the MD probability is the CDF of this variable computed at τ, i.e.,

$$P_{\mathrm{MD}}(\zeta) = F_{\chi^2,\zeta)}(\tau). \tag{17}$$

1) On The Threshold Choice: Threshold τ is typically set to obtain a desired FA probability, i.e.,

$$\tau = F_{\chi^2,0}^{-1}(1 - P_{\mathrm{FA}}), \tag{18}$$

whereas the MD probability (17) becomes

$$P_{\mathrm{MD}}(\zeta) = F_{\chi^2,\zeta}\left(F_{\chi^2,0}^{-1}(1 - P_{\mathrm{FA}})\right). \tag{19}$$

2) Average MD Probability: The average MD probability, averaging over ζ (i.e., over $\boldsymbol{Z}$, $\boldsymbol{H}$, $\boldsymbol{G}$, and $\boldsymbol{\Phi}'$) from (17) is

$$\bar{P}_{\mathrm{MD}} = \mathbb{E}[F_{\chi^2,\zeta}(\tau)]. \tag{20}$$

Unfortunately, a closed-form expression for $\bar{P}_{\mathrm{MD}}$ is not available and we should resort to numerical methods to compute it.

B. The MAP Attack in the Reference Scenario

In the reference scenario, under partial channel knowledge, Eve knows $\boldsymbol{H}$ and a noisy version of $\boldsymbol{G}$, as described by (4). Since she does not know the current RIS configuration, she will use the *average* RIS configuration

$$\mathbb{E}[\boldsymbol{\Phi}] = \mathbb{E}[\mathrm{diag}\{e^{j\epsilon_n}\}] = \frac{\sin\gamma}{\gamma}\bar{\boldsymbol{\Phi}}, \tag{21}$$

where we used the integral relation

$$(2\gamma)^{-1}\int_{-\gamma}^{\gamma} e^{j\epsilon} d\epsilon = \frac{\sin(\gamma)}{\gamma}. \tag{22}$$

Then, Eve computes

$$\boldsymbol{L} = \boldsymbol{H}\mathbb{E}[\boldsymbol{\Phi}]\boldsymbol{U} = \boldsymbol{H}\frac{\sin\gamma}{\gamma}\bar{\boldsymbol{\Phi}}\boldsymbol{U}, \tag{23}$$

and from $\boldsymbol{L}$ she obtains the MAP estimate of $\overline{\boldsymbol{Q}}^{(A,I,B)}(\boldsymbol{\Phi}')$.[2]

Assuming to know the conditional probability density function (PDF) $f_{\overline{\boldsymbol{Q}}^{(A,I,B)}|\boldsymbol{L}}(\overline{\boldsymbol{Q}}^{(A,I,B)}|\boldsymbol{L})$ (which expresses the uncertainty about the Alice-RIS-Bob channel for given $\boldsymbol{L}$), the MAP attack is

$$\bar{\boldsymbol{Z}} = \arg\max_{\boldsymbol{Z}} f_{\overline{\boldsymbol{Q}}^{(A,I,B)}|\boldsymbol{L}}(\boldsymbol{Z}|\boldsymbol{L}). \tag{24}$$

In the reference scenario, the (conditioned) cascade Alice-RIS-Bob channel is not Gaussian distributed since it is the

[2] In the following of this section we drop the reference to the RIS configuration and we use the notation $\overline{\boldsymbol{Q}}^{(A,I,B)}$.

product of (conditioned) Gaussian variables. However, assuming that the entries of channels $\boldsymbol{H}$ and $\boldsymbol{G}$ are independent and identically distributed and the number of RIS elements is large, we can invoke the large number theorem and model $\overline{\boldsymbol{Q}}^{(A,I,B)}$ as Gaussian distributed. Under this approximation, the MAP attack can be obtained following the steps of [34].

First, note that channels $\overline{\boldsymbol{Q}}^{(A,I,B)}$ and $\boldsymbol{L}$ do not have zero mean. Let $\boldsymbol{\mu}^{(A,I,B)} = \mathbb{E}[\mathrm{vec}(\overline{\boldsymbol{Q}}^{(A,I,B)})]$ and $\boldsymbol{\mu}^{(L)} = \mathbb{E}[\mathrm{vec}(\boldsymbol{L})]$. Then we define the covariance matrices of $\mathrm{vec}(\overline{\boldsymbol{Q}}^{(A,I,B)})$ and $\mathrm{vec}(\boldsymbol{L})$ as

$$\begin{aligned}\boldsymbol{R}^{(A,I,B)} &= \mathbb{E}[(\mathrm{vec}(\overline{\boldsymbol{Q}}^{(A,I,B)}) - \boldsymbol{\mu}^{(A,I,B)}) \\ &\quad \times (\mathrm{vec}(\overline{\boldsymbol{Q}}^{(A,I,B)}) - \boldsymbol{\mu}^{(A,I,B)})^{H}], \qquad (25a)\\ \boldsymbol{R}^{(L)} &= \mathbb{E}[(\mathrm{vec}(\boldsymbol{L}) - \boldsymbol{\mu}^{(L)}) \\ &\quad \times (\mathrm{vec}(\boldsymbol{L}) - \boldsymbol{\mu}^{(L)})^{H}], \qquad (25b)\\ \boldsymbol{R}^{((A,I,B),(L))} &= \mathbb{E}[(\mathrm{vec}(\overline{\boldsymbol{Q}}^{(A,I,B)}) - \boldsymbol{\mu}^{(A,I,B)}) \\ &\quad \times (\mathrm{vec}(\boldsymbol{L}) - \boldsymbol{\mu}^{(L)})^{H}]. \qquad (25c)\end{aligned}$$

Thus, for N large and $N \gg M$ and K, by invoking the law of large numbers, we can approximate the random vector $\left[\mathrm{vec}(\overline{\boldsymbol{Q}}^{(A,I,B)})^{T}, \mathrm{vec}(\boldsymbol{L})^{T}\right]^{T}$ as Gaussian distributed with mean $\boldsymbol{\mu} = [\boldsymbol{\mu}^{(A,I,B)T}, \boldsymbol{\mu}^{(L)T}]^{T}$ and covariance matrix

$$\boldsymbol{R}^{(I)} = \begin{bmatrix} \boldsymbol{R}^{(A,I,B)} & \boldsymbol{R}^{((A,I,B),(L))} \\ \boldsymbol{R}^{((A,I,B),(L))^{H}} & \boldsymbol{R}^{(L)} \end{bmatrix}. \qquad (26)$$

By partitioning matrix $\boldsymbol{S} = \boldsymbol{R}^{(I)-1}$ into blocks with sizes as in (26), i.e.,

$$\boldsymbol{S} = \begin{bmatrix} \boldsymbol{S}_{11} & \boldsymbol{S}_{12} \\ \boldsymbol{S}_{12}^{H} & \boldsymbol{S}_{22} \end{bmatrix}, \qquad (27)$$

the MAP attack by Eve is (in vectorial form) [34]

$$\mathrm{vec}(\bar{\boldsymbol{Z}}) = \mathrm{vec}(\boldsymbol{\mu}^{(A,I,B)}) - \boldsymbol{S}_{11}^{-1}\boldsymbol{S}_{12}[\mathrm{vec}(\boldsymbol{L}) - \mathrm{vec}(\boldsymbol{\mu}^{(L)})]. \qquad (28)$$

An alternative expression for (28) is the following

$$\begin{aligned}\mathrm{vec}(\bar{\boldsymbol{Z}}) &= \mathrm{vec}(\boldsymbol{\mu}^{(A,I,B)}) + \\ &\quad \boldsymbol{R}^{((A,I,B),(L))}\boldsymbol{R}^{(L)-1}[\mathrm{vec}(\boldsymbol{L}) - \mathrm{vec}(\boldsymbol{\mu}^{(L)})]. \qquad (29)\end{aligned}$$

In fact, as $\boldsymbol{S}\boldsymbol{R}^{(I)} = \boldsymbol{I}$, we have that

$$\boldsymbol{S}_{11}\boldsymbol{R}^{((A,I,B),(L))} + \boldsymbol{S}_{12}\boldsymbol{R}^{(L)} = \boldsymbol{0}. \qquad (30)$$

Left and right multiplication with $\boldsymbol{S}_{11}^{-1}$ and $\boldsymbol{R}^{(L)-1}$, respectively, yields

$$\boldsymbol{R}^{((A,I,B),(L))}\boldsymbol{R}^{(L)-1} + \boldsymbol{S}_{11}^{-1}\boldsymbol{S}_{12} = \boldsymbol{0}, \qquad (31)$$

from which $\boldsymbol{S}_{11}^{-1}\boldsymbol{S}_{12} = -\boldsymbol{R}^{((A,I,B),(L))}\boldsymbol{R}^{(L)-1}$. The advantage of (29) with respect to (28) is that in the former case only the inverse of $\boldsymbol{R}^{(L)}$ is required, while in the latter case the inverse of the entire matrix $\boldsymbol{R}^{(I)}$ is required.

1) Perfect Channel Knowledge: Under perfect channel knowledge, we assume that the channels among all devices are known, while the RIS configuration remains unknown. In this case, $\boldsymbol{\mu}^{(L)} = \boldsymbol{\mu}^{(A,I,B)} = \boldsymbol{L}$, and (29) boils down to

$$\bar{\boldsymbol{Z}} = \boldsymbol{L} = \boldsymbol{H}\mathbb{E}[\boldsymbol{\Phi}]\boldsymbol{G} = \frac{\sin\gamma}{\gamma}\boldsymbol{H}\overline{\boldsymbol{\Phi}}\boldsymbol{G}. \qquad (32)$$

2) Attack Through the RIS: We note that Eve can exploit the current RIS configuration by transmitting through the RIS, but she does not know the instantaneous value of the RIS. Thus, she chooses 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. The MAP estimate is then

$$\bar{\boldsymbol{G}} = \arg\max_{\boldsymbol{G}'} \mathbb{P}[\overline{\boldsymbol{Q}}^{(A,I,B)}(\boldsymbol{\Phi}') = \boldsymbol{H}\boldsymbol{\Phi}'\boldsymbol{G}'|\boldsymbol{U}], \qquad (33)$$

where $\boldsymbol{H}$ is fixed and $\boldsymbol{\Phi}'$ is random. Due to the complexity of solving problem (33), we consider instead a MAP estimate on $\boldsymbol{G}$, i.e.,

$$\bar{\boldsymbol{G}} = \arg\max_{\boldsymbol{G}'} \mathbb{P}[\boldsymbol{G} = \boldsymbol{G}'|\boldsymbol{U}], \qquad (34)$$

which, considering the model (4), provides the least squares (LS) estimate of $\boldsymbol{G}$ given $\boldsymbol{U}$, i.e., $\bar{\boldsymbol{G}} = \boldsymbol{U}/\rho$. In this case, the cascaded channel under attack through the RIS becomes

$$\bar{\boldsymbol{Z}} = \frac{1}{\rho}\boldsymbol{H}\boldsymbol{\Phi}'\boldsymbol{U} = \boldsymbol{Q}^{(A,I,B)}(\boldsymbol{\Phi}') + \frac{\sqrt{1-\rho^{2}}}{\rho}\boldsymbol{H}\boldsymbol{\Phi}'\boldsymbol{D}, \qquad (35)$$

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

V. Reference Scenario With Single-Antenna Devices

In the following, we consider the special case of a BS and a UE with a single antenna each, in the reference scenario. This case is easy to analyze because the communication-optimal RIS configuration has a closed-form expression. In fact, $\boldsymbol{H}$ and $\boldsymbol{G}$ become a row and a column vector, respectively, and the communication-optimal RIS configuration that maximizes the Alice-Bob achievable rate is

$$\bar{\theta}_{n} = -\angle H_{1,n} - \angle G_{n,1}, \quad n = 0, \ldots, N-1. \qquad (36)$$

The UE-RIS-BS cascaded channel and the attack channels become scalar, thus we will denote them in italics, Q, and Z, respectively.

1) Security Metric: Under attack Z, we have $\Psi = \frac{2}{\sigma^{2}}|\delta|^{2}$ with (from (15))

$$\delta = Z + W - \boldsymbol{H}\boldsymbol{\Phi}'\boldsymbol{G}, \qquad (37)$$

and the average MD probability becomes

$$\bar{P}_{\mathrm{MD}} = \mathbb{P}\left[\frac{2}{\sigma^{2}}|\delta|^{2} < \tau\right]. \qquad (38)$$

2) Communication Metric: The spectral efficiency of the Alice-Bob channel with the random RIS configuration $\boldsymbol{\Phi}$ is

$$C_{A,B}(\boldsymbol{\Phi}) = \log_{2}\left(1 + \frac{\left|\sum_{n=0}^{N-1} H_{1,n}G_{n,1}e^{j\theta_{n}}\right|^{2}}{\sigma_{B}^{\prime 2}}\right), \qquad (39)$$

where $\sigma_{B}^{\prime 2}$ is the noise variance on the received data symbols.

In the considered reference scenario, as described in Appendix B, as $N \to \infty$ the channel SNR averaged with respect to the random RIS configuration can be approximated as

$$\Omega(\gamma) = \frac{N\sigma_{\rm sec}^2}{2\sigma_{\rm B}'^2}\left(2 + \frac{2N^2\mu_{\rm sec}^2}{N\sigma_{\rm sec}^2}\right) = \frac{N}{\sigma_{\rm B}'^2}(N\mu_{\rm sec}^2 + \sigma_{\rm sec}^2), \quad (40)$$

where $\mu_{\rm sec}$ and $\sigma_{\rm sec}^2$ are the mean and variance of each term of the sum in (39) (see (65a) and (65b)), and we can approximate the average spectral efficiency as

$$c(\gamma) = \mathbb{E}[C_{A,B}(\boldsymbol{\Phi})] \approx \tilde{c}(\gamma) = \log_2(1 + \Omega(\gamma)). \quad (41)$$

A. MAP Attack With Perfect Channel Knowledge

We first consider the scenario in which Eve transmits directly to Bob. With perfect channel knowledge at Eve, inserting (32) in (37) we obtain

$$\begin{aligned}\delta &= \frac{\sin\gamma}{\gamma}\boldsymbol{H}\bar{\boldsymbol{\Phi}}\boldsymbol{G} + W - \boldsymbol{H}\boldsymbol{\Phi}'\boldsymbol{G}\\ &= \sum_{n=0}^{N-1} H_{1,n}e^{j\bar{\theta}_n}\left[\frac{\sin\gamma}{\gamma} - e^{j\epsilon_n}\right]G_{n,1} + W. \quad (42)\end{aligned}$$

In the reference scenario, the terms in the summation of (42) are i.i.d. with mean

$$\mathbb{E}\left[H_{1,n}e^{j\bar{\theta}_n}\left[\frac{\sin\gamma}{\gamma} - e^{j\epsilon_n}\right]G_{n,1}\right] = 0, \quad (43)$$

the variances of the real and imaginary parts are

$$\begin{aligned}\sigma_{\rm R}^2 &= \mathbb{E}\left[\text{Re}\left\{H_{1,n}e^{j\bar{\theta}_n}\left[\frac{\sin\gamma}{\gamma} - e^{j\epsilon_n}\right]G_{n,1}\right\}^2\right] =\\ &= \mathbb{E}\left[\left(\frac{\sin\gamma}{\gamma} - \cos\epsilon_n\right)^2\right] \quad (44)\\ &= \frac{1}{2\gamma}\left[2\frac{\sin^2\gamma}{\gamma^2}\gamma + \sin\gamma\left(\cos\gamma - 4\frac{\sin\gamma}{\gamma}\right) + \gamma\right]\\ &= \frac{1}{2}\left[\frac{\sin\gamma}{\gamma}\cos\gamma + 1 - 2\frac{\sin^2\gamma}{\gamma^2}\right],\\ \sigma_{\rm I}^2 &= \mathbb{E}\left[\text{Im}\left\{H_{1,n}e^{j\bar{\theta}_n}\left[\frac{\sin\gamma}{\gamma} - e^{j\epsilon_n}\right]G_{n,1}\right\}^2\right]\\ \mathbb{E}\left[\sin^2\epsilon_n\right] &= \frac{1}{2\gamma}\left[\gamma - \sin(\gamma)\cos(\gamma)\right]\\ &= \frac{1}{2}\left[1 - \frac{\sin\gamma}{\gamma}\cos\gamma\right], \quad (45)\end{aligned}$$

and their cross-covariance is

$$\begin{aligned}&\mathbb{E}\left[\text{Re}\left\{H_{1,n}e^{j\bar{\theta}_n}\left[\frac{\sin\gamma}{\gamma} - e^{j\epsilon_n}\right]G_{n,1}\right\}\times\right.\\ &\left.\quad\text{Im}\left\{H_{1,n}e^{j\bar{\theta}_n}\left[\frac{\sin\gamma}{\gamma} - e^{j\epsilon_n}\right]G_{n,1}\right\}\right]\\ &= \mathbb{E}\left[\left(\frac{\sin\gamma}{\gamma} - \cos\epsilon_n\right)(-\sin\epsilon_n)\right] = 0. \quad (46)\end{aligned}$$

Thus, invoking the law of large numbers, we have for $N \to \infty$ that $\delta = \delta_{\rm R} + j\delta_{\rm I}$ is complex Gaussian distributed with average $\mu_\delta = 0$, independent real and imaginary components with variances $\sigma_{\delta,{\rm R}}^2 = N\sigma_{\rm R}^2 + \frac{\sigma^2}{2}$ and $\sigma_{\delta,{\rm I}}^2 = N\sigma_{\rm I}^2 + \frac{\sigma^2}{2}$, respectively.

Thus, the average MD probability can be approximated from (38) as

$$\bar{P}_{\rm MD} = \mathbb{P}\left[\delta_{\rm R}^2 + \delta_{\rm I}^2 \leq \frac{\sigma^2\tau}{2}\right] \approx \mathbb{P}\left[\sigma_{\delta,{\rm R}}^2 g_1^2 + \sigma_{\delta,{\rm I}}^2 g_2^2 \leq \frac{\sigma^2\tau}{2}\right], \quad (47)$$

where g_1 and g_2 are zero-mean unit-variance real Gaussian random variables. To compute (47) we must use the CDF of the linear combination of two independent central chi-squared random variables, each with one degree of freedom. No exact close-form expression for this CDF is known; however, a series expansion has been computed in [35].

For the special case of $\gamma = \pi$, we have $\sigma_{\delta,{\rm R}}^2 = \sigma_{\delta,{\rm I}}^2 = \frac{N}{2} + \frac{\sigma^2}{2}$, and (47) becomes the CDF of a central chi-square distribution, i.e.,

$$\bar{P}_{\rm MD} \approx \frac{1}{\Gamma(1)}\gamma_{\rm inc}\left(1, \frac{\sigma^2}{2(N+\sigma^2)}\tau\right), \quad (48)$$

where $\gamma_{\rm inc}(a, b) = \int_0^b t^{a-1}e^{-t}dt$ is the lower incomplete gamma function and $\Gamma(\cdot)$ is the gamma function.

For $\gamma = 0$, we have that the attacker is indistinguishable from Alice, and indeed we have $\sigma_{\rm R}^2 = \sigma_{\rm I}^2 = 0$ and $\sigma_{\delta,{\rm R}}^2 = \sigma_{\delta,{\rm I}}^2 = \sigma/2$, thus

$$\bar{P}_{\rm MD} = F_{\chi^2,0}(\tau), \quad (49)$$

which holds without approximations and coincides with the complement to the FA probability (see (12a)), i.e., $\bar{P}_{\rm MD} = 1 - P_{\rm FA}$.

1) Attack Through the RIS: In case of attack through the RIS and perfect channel knowledge, the resulting attack channel is $\boldsymbol{Z} = \boldsymbol{H}\boldsymbol{\Phi}'\boldsymbol{G}$ and the MD probability is equal to the FA probability, $\bar{P}_{\rm MD} = 1 - P_{\rm FA}$.

B. MAP Attack With Partial Channel Knowledge

We first consider the scenario in which Eve transmits directly to Bob. Under partial channel knowledge, defining

$$\omega = \frac{\pi^2\sin^2\gamma}{16\gamma^2}, \quad (50)$$

we have

$$\boldsymbol{R}^{(I)} = \begin{bmatrix} N(1-\omega) + \sigma_{\rm B}^2 & N\rho\left(\frac{\sin^2\gamma}{\gamma^2} - \omega\right)\\ N\rho\left(\frac{\sin^2\gamma}{\gamma^2} - \omega\right) & N\left(\frac{\sin^2\gamma}{\gamma^2} - \rho^2\omega\right)\end{bmatrix} \quad (51)$$

and

$$\boldsymbol{S} = \begin{bmatrix} \frac{R_{2,2}^{(I)}}{R_{1,1}^{(I)}R_{2,2}^{(I)} - R_{1,2}^{(I)2}} & -\frac{R_{1,2}^{(I)}}{R_{1,1}^{(I)}R_{2,2}^{(I)} - R_{1,2}^{(I)2}}\\ -\frac{R_{1,2}^{(I)2}}{R_{1,1}^{(I)}R_{2,2}^{(I)} - R_{1,2}^{(I)2}} & \frac{R_{1,1}^{(I)}}{R_{1,1}^{(I)}R_{2,2}^{(I)} - R_{1,2}^{(I)2}}\end{bmatrix}. \quad (52)$$

Now, as derived in Appendix C, we have

$$|\delta|^2 \simeq \sigma_{\delta,{\rm R}}^2\left(x_1 + \frac{\mu_\delta}{\sigma_{\delta,{\rm R}}}\right)^2 + \sigma_{\delta,{\rm I}}^2 x_2^2, \quad (53)$$

where x_1 and x_2 are independent zero-mean unit-variance real Gaussian random variables, and μ_δ, $\sigma_{\delta,{\rm R}}$, and $\sigma_{\delta,{\rm I}}$ are computed in Appendix C. Thus, $|\delta|^2$ is the linear combination of a non-central chi-squared variable (with 1 degree of freedom and non-centrality parameter $\mu_\delta^2/\sigma_{\delta,\mathrm{R}}^2$) and a central chi-square variable with 1 degree of freedom. Again, there is no exact close-form expression for the CDF of this combination, but we can resort to its series expansion derived in [35].

Note that for $\rho = 0$, $\boldsymbol{L}$ becomes uncorrelated to $\overline{\boldsymbol{Q}}^{(A,I,B)}$, i.e., Eve does not have any information on the target channel.

1) Attack Through the RIS: When Eve transmits through the RIS and she has a partial channel knowledge, we have (see Appendix C) that δ is zero-mean with variances

$$\mathrm{Var}(\mathrm{Re}\,\{\delta\}) = \mathrm{Var}(\mathrm{Im}\,\{\delta\}) = \frac{1-\rho^2}{\rho^2}\frac{N}{2} + \frac{\sigma^2}{2}. \tag{54}$$

When assuming N is large and invoking the central limit theorem, δ_R and δ_I follow a zero-mean normal distribution with variance σ_δ^2. Then $\bar{P}_{MD}$ can be approximated as (see (38) and (47))

$$\bar{P}_{MD} = \mathbb{P}\left[g_1^2 + g_2^2 < \frac{\tau\sigma^2}{\frac{1-\rho^2}{\rho^2}N + \sigma^2}\right] \tag{55}$$

$$\approx \frac{1}{\Gamma(1)}\gamma_{\mathrm{inc}}\left(1, \frac{\tau\sigma^2}{2\left(\frac{1-\rho^2}{\rho^2}N + \sigma^2\right)}\right). \tag{56}$$

As expected, the MD probability does not depend on γ, i.e., the randomness of the RIS for PLA. Thus, the performance of CR-PLA is that of a tag-based PLA, depending on the accuracy of the knowledge of the Alice-RIS and RIS-Bob channels.

VI. Numerical Results

Numerical results are obtained in the reference scenario of Section II-C with single-antenna Alice and Bob, and different numbers of RIS elements. We also set $\sigma_\mathrm{B}^2 = \sigma_\mathrm{B}'^2 = N/100$, to reduce the effects of the number of the RIS elements on the SNR at the receiver.

A. Analysis and Simulation Comparison

We first evaluate the effects of the Gaussian approximation of Section V, on the MD probability averaged over both the channel realizations and the RIS configurations, as a function of γ, for $\rho = 0.9$ and different values of RIS elements N. The threshold τ for the GLRT is set to obtain an FA probability $P_\mathrm{FA} = 10^{-3}$. To verify the theoretical analysis of Section V-B, we perform Monte Carlo simulations and compare the obtained results. For the scenario in which Eve transmits directly to Bob, Fig. 2 shows the performance obtained from both the Monte Carlo simulations (solid lines) and the analysis of Section V-B (crosses) for $N \to \infty$. Note that the analytical performance does not depend on the number of RIS elements N, due to the choice of the noise power at Bob (which scales with N). This choice gives a matrix $\boldsymbol{R}^{(I)}$ proportional to N, so the matrix $\boldsymbol{S}$ is inversely proportional to N; so δ does not depend on N since, from (72), its dependence on $\boldsymbol{S}$ is only through the ratio $S_{11}^{-1}S_{12}$, which does not depend on N. Thus, a single line is shown in the figure. Then we note that from about $N = 30$, the analysis of Section V-B provides a good approximation to the MD probability calculated by simulations. From Fig. 2 we also see that as γ increases, the MD probability significantly decreases, from 10^{-1} for $\gamma = 0$, which corresponds to the case of no CR PLA, to 10^{-3} for $\gamma = \pi$, for which the random phase of the RIS completely removes the phase information of the channels. Note also that a stronger decrease in the MD probability occurs for small values of γ.

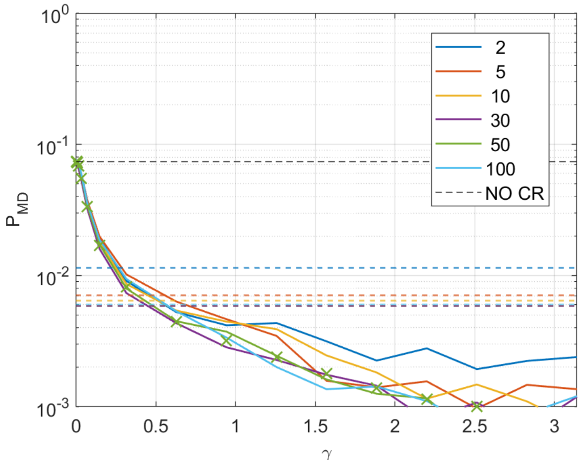

Fig. 2. Average MD probability as a function of γ for $\rho = 0.9$, an FA probability $P_\mathrm{FA} = 10^{-3}$, and $N = 2$, 5, 10, 30, 50, and 100 RIS elements. Solid lines denote Monte Carlo simulations and crosses denote analytical results when Eve transmits directly to Bob. Dashed lines are for the case of an attack through the RIS.

For the sake of completeness, Fig. 2 also shows the performance of an attack through the RIS. Note that the resulting MD probability does not depend on γ. The comparison between the two scenarios is problematic because when transmitting the attack signal directly to Bob, Eve has to find the cascaded channel (which is smaller but depends on the RIS configuration), while when sending through the RIS, she has to find Alice's RIS channel (which is larger but does not depend on the RIS configuration). In this case, we notice that the attack through the RIS is more effective, giving a higher PMD than the lowest obtained for a high value of γ and CR-PLA.

B. Spectral Efficiency

We now investigate the effect of the random phase of the RIS elements on the spectral efficiency of the resulting communication, as discussed in Section V. In particular, we consider as a metric the spectral efficiency loss

$$\eta = 1 - \frac{c(\gamma)}{c(0)}, \tag{57}$$

where $c(\cdot)$ was defined in (41) as the expected spectral efficiency. In the numerator of (57), the expectation is taken for both the channels and the random RIS configurations (for CR PLA), while in the denominator it is taken for the channels, with the configuration $\overline{\boldsymbol{\Phi}}$ optimized for each channel realization. Note that the denominator is obtained with the tag-based PLA. We verify the validity of the approximation by Monte Carlo simulations. Specifically, we compare the simulated value of η with its analytical approximation obtained

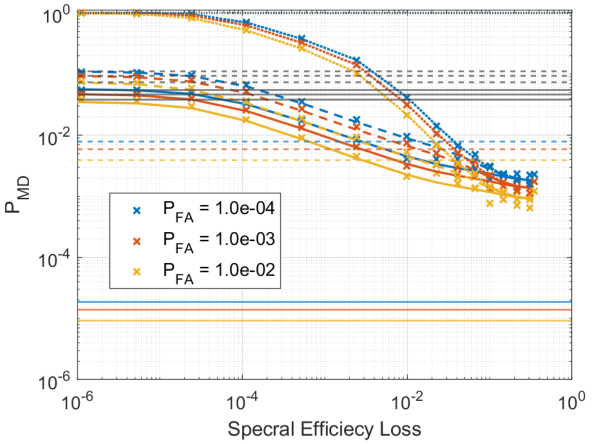

Fig. 3. Average MD probability as a function of the spectral efficiency loss for $N = 100$, $\rho = 0.1$ (solid lines), $\rho = 0.9$ (dashed lines), and $\rho = 1$ (dotted lines), and FA probabilities $P_{\mathrm{FA}} = 10^{-4}$, 10^{-3}, and 10^{-2}. Analytical results are reported as crosses.

using $\tilde{c}(\gamma)$ instead of $c(\gamma)$ (see Section V). Note that η varies from 0 to 1 and measures the amount of spectral efficiency reduction (normalized to the spectral efficiency of the tag-based PLA): a smaller η means a higher communication rate.

For the case of Eve transmitting directly to Bob, Fig. 3 shows the average MD probability as a function of spectral efficiency loss η, for $N = 100$, $\rho = 0$ (solid lines), $\rho = 0.9$ (dashed lines), and $\rho = 1$ (dotted lines), and FA probabilities $P_{\mathrm{FA}} = 10^{-4}$, 10^{-3}, and 10^{-2}. The analytical results are shown as crosses. Results are obtained by varying the value of γ and estimating the resulting spectral efficiency loss and average MD probability. We see a remarkable agreement between the simulated and analytical results. We find a trade-off between data rate and security, as lower MD probabilities are achieved for higher spectral efficiency losses. Nevertheless, we note that a loss of 1% (i.e, 99% of the optimal spectral efficiency) significantly reduces the MD probability for all values of ρ and FA probabilities considered. Moreover, even with $\gamma = \pi$ (corresponding to completely random phases), the loss of spectral efficiency stops at about 20%, while the MD probability reaches its minimum.

Fig. 3 also shows the MD probability for an attack through the RIS. As already seen in Fig. 2, the performance in this case does not depend on γ, so we have a constant P_{MD} for all spectral efficiency loss values. Compared to direct transmission, we see that for small correlation values, the RIS attack is less effective (lower MD probability) than the direct attack due to the difficulty in guessing the Alice-RIS channel. On the other hand, for high correlation values, the advantage of using the current RIS configuration is more relevant and the RIS attack is more effective than the direct attack.

C. Effect of Channel Correlation

Finally, we examine in detail the effects of the correlation between the Alice-RIS and Eve-RIS channels. Fig. 4 shows the analytical mean MD probability as a function of ρ for an FA probability $P_{\mathrm{FA}} = 10^{-3}$, $N = 50$, and for values of ρ that either give the minimum MD probability, i.e. $\gamma = \pi$ (independent of spectral efficiency), or give a spectral efficiency loss $\eta = 2\%$, 10%, or 50%. First, we consider the scenario in which Eve transmits directly to Bob. We also include the performance of the tag-based PLA for comparison. First, we observe that using the CR-PLA mechanism significantly reduces the MD probability compared to the case without CR. Regarding the behavior as a function of ρ, we observe that even for $\rho = 0$ (uncorrelated channels), the MD probability is high for the scheme without CR, since choosing the RIS configuration that maximizes the spectral efficiency makes the resulting channel non-zero mean, and this bias can be exploited by the attacker. Using the CR approach instead reduces the bias (which becomes zero for $\gamma = \pi$) and results in a lower MD probability. Indeed, for a spectral efficiency loss of only 2%, the MD probability drops to about 10^{-2} with $\rho = 1$, and even lower values are obtained for lower ρ or larger η. Also note that in the absence of CR, the MD mostly increases for $\rho > 0.8$, while the increase is smoother with CR PLA.

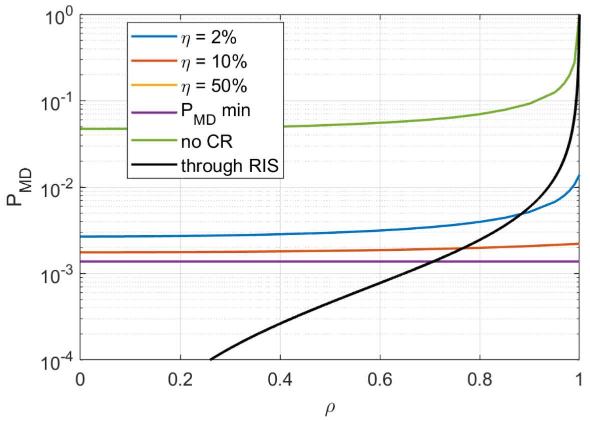

Fig. 4. Analytical average MD probability as a function of ρ for FA probability $P_{\mathrm{FA}} = 10^{-3}$, for values of γ providing the minimum MD probability (irrespective of the spectral efficiency), or giving values of $\eta = 2\%$, 10% or 50%, and in the absence of CR PLA mechanism.

For the RIS attack, we can see from Fig. 4 that the correlation factor has a higher impact on the MD probability. As observed before, a lower correlation makes the RIS attack less effective than the direct attack, while at high correlations the direct attack is more effective.

VII. Conclusion

We have proposed a novel CR-PLA technique where the phases of the RIS elements are randomly selected by the BS. This generates a challenge that induces a specific UE-RIS-BS cascaded channel (the response) to be verified from the channel estimate on the received message. We have also studied the MAP attack by an impersonation device that obtains the MAP attack from its knowledge of the channels. Finally, we analyzed the performance of the security mechanism and its impact on the communication performance when the BS and the UE each have a single antenna. The simulation results confirm the effectiveness of the authentication mechanism with a limited impact on the communication performance.

Appendix A
Derivation of the GLRT

Let us define $\widehat{\boldsymbol{q}} = \mathrm{vec}(\hat{\boldsymbol{Q}})$, $\boldsymbol{q} = \mathrm{vec}(\boldsymbol{Q}^{(A,I,B)}(\boldsymbol{\Phi}'))$, $\boldsymbol{w}'' = \mathrm{vec}(\boldsymbol{W}'')$, $\overline{\boldsymbol{q}} = \mathrm{vec}(\overline{\boldsymbol{Q}}^{(A,I,B)}(\boldsymbol{\Phi}'))$, and $\boldsymbol{z} = \mathrm{vec}(\boldsymbol{Z})$.

Under hypothesis $\mathcal{H}_0$, the joint PDF of $\widehat{\boldsymbol{q}}$ and $\overline{\boldsymbol{q}}$ is

$$f_{\widehat{\boldsymbol{q}},\overline{\boldsymbol{q}}|\mathcal{H}_0}(\widehat{\boldsymbol{q}},\overline{\boldsymbol{q}}) = \left(\pi\sigma_B^2\right)^{-2KM} \exp\left(-\frac{1}{\sigma_B^2}\left(||\widehat{\boldsymbol{q}}-\boldsymbol{q}||^2 + ||\overline{\boldsymbol{q}}-\boldsymbol{q}||^2\right)\right), \tag{58}$$

which contains the unknown vector $\boldsymbol{q}$.

Under hypothesis $\mathcal{H}_1$ the joint density of $\widehat{\boldsymbol{q}}$ and $\overline{\boldsymbol{q}}$ is

$$f_{\widehat{\boldsymbol{q}},\overline{\boldsymbol{q}}|\mathcal{H}_1}(\widehat{\boldsymbol{q}},\overline{\boldsymbol{q}}) = \left(\pi\sigma_{\mathrm{B}}^2\right)^{-2KM} \exp\left(-\frac{1}{\sigma_{\mathrm{B}}^2}\left(||\widehat{\boldsymbol{q}}-z||^2 + ||\overline{\boldsymbol{q}}-\boldsymbol{q}||^2\right)\right), \tag{59}$$

which contains the unknown vectors z and $\boldsymbol{q}$.

The presence of the unknown parameters in $f_{\widehat{\boldsymbol{q}},\overline{\boldsymbol{q}}|\mathcal{H}_0}(\widehat{\boldsymbol{q}},\overline{\boldsymbol{q}})$ and $f_{\widehat{q},\overline{q}|\mathcal{H}_1}(\widehat{\boldsymbol{q}},\overline{\boldsymbol{q}})$ prevents the application of a standard LRT. As an alternative, the GLRT can be applied instead: this involves replacing $f_{\widehat{\boldsymbol{q}},\overline{\boldsymbol{q}}|\mathcal{H}_0}(\widehat{\boldsymbol{q}},\overline{\boldsymbol{q}})$ and $f_{\widehat{q},\overline{q}|\mathcal{H}_1}(\widehat{\boldsymbol{q}},\overline{\boldsymbol{q}})$ by $\widetilde{f}_{\widehat{\boldsymbol{q}},\overline{\boldsymbol{q}}|\mathcal{H}_0}(\widehat{\boldsymbol{q}},\overline{\boldsymbol{q}})$ and $\widetilde{f}_{\widehat{\boldsymbol{q}},\overline{\boldsymbol{q}}|\mathcal{H}_1}(\widehat{\boldsymbol{q}},\overline{\boldsymbol{q}})$, where

$$\widetilde{f}_{\widehat{\boldsymbol{q}},\overline{\boldsymbol{q}}|\mathcal{H}_0}(\widehat{\boldsymbol{q}},\overline{\boldsymbol{q}}) = \max_{\boldsymbol{q}} f_{\widehat{\boldsymbol{q}},\overline{\boldsymbol{q}}|\mathcal{H}_0}(\widehat{\boldsymbol{q}},\overline{\boldsymbol{q}}), \tag{60}$$

$$\widetilde{f}_{\widehat{\boldsymbol{q}},\overline{\boldsymbol{q}}|\mathcal{H}_1}(\widehat{\boldsymbol{q}},\overline{\boldsymbol{q}}) = \max_{\boldsymbol{z},\boldsymbol{q}} f_{\widehat{\boldsymbol{q}},\overline{\boldsymbol{q}}|\mathcal{H}_1}(\widehat{\boldsymbol{q}},\overline{\boldsymbol{q}}). \tag{61}$$

The GLRT then becomes

$$\ln\left(\frac{\widetilde{f}_{\widehat{q},\overline{q}|\mathcal{H}_1}(\widehat{\boldsymbol{q}},\overline{\boldsymbol{q}})}{\widetilde{f}_{\widehat{q},\overline{q}|\mathcal{H}_0}(\widehat{\boldsymbol{q}},\overline{\boldsymbol{q}})}\right)\hat{\mathcal{H}} = \mathcal{H}_1 \\ \gtrless \\ \hat{\mathcal{H}} = \mathcal{H}_0\tau'. \tag{62}$$

It can be verified that from (58) and (59) we have the following results

- $f_{\widehat{q},\overline{q}|\mathcal{H}_0}(\widehat{\boldsymbol{q}},\overline{\boldsymbol{q}})$ achieves its maximum for $\boldsymbol{q} = \frac{1}{2}(\widehat{\boldsymbol{q}}+\overline{\boldsymbol{q}})$, yielding (from (13))

$$\widetilde{f}_{\widehat{\boldsymbol{q}},\overline{\boldsymbol{q}}|\mathcal{H}_0}(\widehat{\boldsymbol{q}},\overline{\boldsymbol{q}}) = \left(\pi\sigma_{\mathrm{B}}^2\right)^{-2KM}\exp\left(-\frac{1}{2\sigma_{\mathrm{B}}^2}||\widehat{\boldsymbol{q}}-\overline{\boldsymbol{q}}||^2\right) = \left(\pi\sigma_{\mathrm{B}}^2\right)^{-2KM}\exp(\Psi/2); \tag{63}$$

- $f_{\widehat{q},\overline{q}|\mathcal{H}_1}(\widehat{\boldsymbol{q}},\overline{\boldsymbol{q}})$ achieves its maximum for $z = \widehat{\boldsymbol{q}}$ and $\boldsymbol{q} = \overline{\boldsymbol{q}}$, yielding

$$\widetilde{f}_{\widehat{\boldsymbol{q}},\overline{\boldsymbol{q}}|\mathcal{H}_1}(\widehat{\boldsymbol{q}},\overline{\boldsymbol{q}}) = \left(\pi\sigma_B^2\right)^{-2KM}. \tag{64}$$

Then, we can rewrite GLRT (61) as (9), where $\tau = 2\tau'$.

Appendix B
Computation of the Capacity

We now analyze the impact of the random RIS configuration on the spectral efficiency in the asymptotic case of $N \to \infty$. The mean and variance of each term of the sum are

$$\mu_{\mathrm{sec}} = \mathbb{E}[H_{1,n}G_{n,1}e^{j\theta_n}] = \mathbb{E}[|H_{1,n}G_{n,1}|\cos\epsilon_n] = \frac{\pi\sin\gamma}{4\gamma}, \tag{65a}$$

$$\sigma_{\mathrm{sec}}^2 = \mathbb{E}\left[\left|H_{1,n}G_{n,1}e^{j\theta_n} - \frac{\pi\sin\gamma}{4\gamma}\right|^2\right] = 1-\mu_{\mathrm{sec}}^2, \tag{65b}$$

since $|H_{1,n}|$ and $|G_{n,1}|$ are independent Rayleigh variables with variance $1/2$.

Moreover, the cross-covariance between the real and imaginary parts is

$$\mathbb{E}\left[\mathrm{Re}\left\{H_{1,n}G_{n,1}e_n^{j\theta} - \frac{\pi\sin\gamma}{4\gamma}\right\}\times \mathrm{Im}\left\{H_{1,n}G_{n,1}e_n^{j\theta} - \frac{\pi\sin\gamma}{4\gamma}\right\}\right] = \mathbb{E}\left[\left(\cos\epsilon_n - \frac{\pi\sin\gamma}{4\gamma}\right)(-\sin\epsilon_n)\right] = 0. \tag{66}$$

Lastly, the terms of the sum in (39) are independent. Then, from the central limit theorem, approximating the sum in (39) as Gaussian distributed, with mean $N\mu_{\mathrm{sec}}$ and variance $N\sigma_{\mathrm{sec}}^2$, the average SNR goes to infinity as[3]

$$\mathbb{E}\left[\frac{\left|\sum_{n=0}^{N-1} H_{1,n}G_{n,1}e^{j\theta_n}\right|^2}{\sigma_B^2}\right] \approx \Omega(\gamma), \tag{68}$$

where $\Omega(\gamma)$ is given by (40). Note that $\Omega(\gamma)$ decreases as γ increases, thus, the choice of γ is a tradeoff between the spectral efficiency and the CR authentication effectiveness. Note, however, that even with $\gamma = \pi$, which corresponds to the highest variability of the RIS phases (highest security), we have $\Omega = \frac{N}{\sigma_{\mathrm{B}}'^2} > 0$, so we expect to have a non-zero spectral efficiency of the resulting system.

As $N \to \infty$ we have that the variance of the SNR is reduced with respect to its mean, so we can approximate the average spectral efficiency as

$$c(\gamma) = \mathbb{E}[C_{A,B}(\boldsymbol{\Phi})] \approx \tilde{c}(\gamma) = \log_2(1+\Omega(\gamma)), \tag{69}$$

where we have emphasized the dependence of the spectral efficiency on the random RIS configuration parameter γ. In Section VI we verify the validity of the approximation by simulations.

Appendix C
Derivation of Means and Variances

We have that $\bar{Q}^{(A,I,B)}$ has mean

$$\mu^{(A,I,B)} = N\mathbb{E}\left[|H_{1,n}||G_{n,1}|e^{j\epsilon_n}\right] = \frac{N\pi\sin\gamma}{4\gamma}, \tag{70}$$

and L has mean

$$\mu^{(L)} = N\mathbb{E}\left[|H_{1,n}||G_{n,1}|\rho e^{j\epsilon_n}\right] = \frac{N\pi\rho\sin\gamma}{4\gamma}. \tag{71}$$

[3] Recall that for a circularly symmetric complex random variable y with (complex) mean m and (real) variance σ^2, the average of $|y|^2$ is (assuming $w = w_R + jw_I$ circularly symmetric complex Gaussian variable with zero mean and unit variance)

$$\mathbb{E}[|m+\sigma w|^2] = \frac{\sigma^2}{2}\mathbb{E}\left[\left(m_R\frac{\sqrt{2}}{\sigma}+\sqrt{2}w_R\right)^2 + \left(m_I\frac{\sqrt{2}}{\sigma}+\sqrt{2}w_I\right)^2\right] = \frac{\sigma^2}{2}(2+\lambda), \tag{67}$$

with $\lambda = \frac{2|m|^2}{\sigma^2}$.

Inserting (28) into (37) we have

$$\begin{aligned}\delta &= \mu^{(A,I,B)} - S_{11}^{-1}S_{12}[L-\mu^{(L)}]+\\ &\quad + W'' - \boldsymbol{H\Phi}'\boldsymbol{G} - W'\\ &= \mu^{(A,I,B)} + S_{11}^{-1}S_{12}\mu^{(L)}+\\ &\sum_n H_{1,n}\left\{e^{j\bar{\theta}_n}\left[\nu_1 - e^{j\epsilon_n}\right]G_{n,1} + \nu_2 e^{j\bar{\theta}_n}D_{n,1}\right\}\\ &\quad + W'' - W',\end{aligned} \tag{72}$$

where the average is taken across the RIS configurations and

$$\nu_1 = -S_{11}^{-1}S_{12}\frac{\rho\sin\gamma}{\gamma}, \quad \nu_2 = S_{11}^{-1}S_{12}\frac{\sqrt{1-\rho^2}\sin\gamma}{\gamma}, \tag{73}$$

which are both real numbers.

Since $\mathbb{E}[e^{j\epsilon_n}] = \frac{\sin\gamma}{\gamma}$, term n in the summation in (72) has mean

$$\mu_1 = \mathbb{E}\left[|H_{1,n}||G_{n,1}|\left[\nu_1 - e^{j\epsilon_n}\right]\right] = \frac{\pi}{4}\left(\nu_1 - \frac{\sin\gamma}{\gamma}\right) \tag{74}$$

and thus

$$\mu_\delta = \mu^{(A,I,B)} + S_{11}^{-1}S_{12}\mu^{(L)} + N\mu_1. \tag{75}$$

The variance of the real part of term n in the summation in (72) is

$$\begin{aligned}\sigma^2_{1,\mathrm{R}} &= \mathbb{E}\left[\mathrm{Re}\left\{H_{1,n}\left\{e^{j\bar{\theta}_n}\left[\nu_1 - e^{j\epsilon_n}\right]G_{n,1} + \nu_2 e^{j\bar{\theta}_n}D_{n,1}\right\}\right.\right.\\ &\quad \left.-\mu_1\}^2\right]\\ &= \mathbb{E}\left[\left\{|H_{1,n}|\left\{[\nu_1 - \cos\epsilon_n]\,|G_{n,1}| + \nu_2\mathrm{Re}(e^{j\bar{\theta}_n}D_{n,1})\right\}\right\}^2\right]\\ &\quad - \mu_1^2\\ &= \nu_1^2 + \frac{\nu_2^2}{2} + \mathbb{E}[\cos^2\epsilon_n] - 2\nu_1\mathbb{E}[\cos\epsilon_n] - \mu_1^2\\ &= \nu_1^2 + \frac{\nu_2^2}{2} + \frac{1}{2} + \frac{\sin\gamma}{2\gamma}\cos\gamma - 2\nu_1\frac{\sin\gamma}{\gamma} - \mu_1^2,\end{aligned} \tag{76}$$

while for the imaginary part, we have

$$\begin{aligned}\sigma^2_{1,\mathrm{I}} &= \mathbb{E}\left[\mathrm{Im}\left\{H_{1,n}\left\{e^{j\bar{\theta}_n}\left[\nu_1 - e^{j\epsilon_n}\right]G_{n,1} + \nu_2 e^{j\bar{\theta}_n}D_{n,1}\right\}\right.\right.\\ &\quad \left.-\mu_1\}^2\right]\\ &= \mathbb{E}\left[\left\{|H_{1,n}|\left\{-\sin\epsilon_n|G_{n,1}| + \nu_2 e^{-j\bar{\theta}}D_{n,1}\right\}\right\}^2\right]\\ &= \mathbb{E}[\sin^2\epsilon_n] + \frac{\nu_2^2}{2} = \frac{1}{2} - \frac{1}{2}\cos\gamma\frac{\sin\gamma}{\gamma} + \frac{\nu_2^2}{2},\end{aligned} \tag{77}$$

and their cross variance is zero.

Thus, invoking the law of large numbers, for $N \to \infty$, we have that δ is complex Gaussian distributed with independent real and imaginary parts, having average

$$\mu_\delta = \mu^{(A,I,B)} + S_{11}^{-1}S_{12}\mu^{(L)} + N\mu_1, \tag{78}$$

while the real-part and imaginary-part variances are

$$\sigma^2_{\delta,\mathrm{R}} = N\sigma^2_{1R} + \sigma^2/2, \quad \sigma^2_{\delta,\mathrm{I}} = N\sigma^2_{1I} + \sigma^2/2. \tag{79}$$

A. Attack Through the RIS

In the case, Eve transmits through the RIS and she has partial channel knowledge we have

$$\begin{aligned}\delta &= W + \frac{\sqrt{1-\rho^2}}{\rho}\boldsymbol{H\Phi}'\boldsymbol{D} &(80)\\ &= W + \frac{\sqrt{1-\rho^2}}{\rho}\sum_{n=0}^{N-1}H_nD_ne^{j(\bar{\theta}_n+\epsilon_n)} &(81)\\ &= W + \frac{\sqrt{1-\rho^2}}{\rho}\sum_{n=0}^{N-1}\delta_n &(82)\end{aligned}$$

Observe that the mean of δ_n is

$$\mathbb{E}[\delta_n] = \mathbb{E}[D_n]\,\mathbb{E}\left[H_ne^{j(\bar{\theta}_n+\epsilon_n)}\right] = 0 \tag{83}$$

and the variances of its real and imaginary parts are

$$\begin{aligned}\sigma_R^2 &= \mathbb{E}\left[\mathrm{Re}\{\delta_n - \mathbb{E}[\delta_n]\}^2\right] &(84)\\ &= \mathbb{E}\left[|H_n|^2|D_n|^2\right]\cdot\frac{1}{8\pi^2\gamma}\kappa = \frac{1}{2} &(85)\\ \sigma_I^2 &= \mathbb{E}\left[\mathrm{Im}\{\delta_n - \mathbb{E}[\delta_n]\}^2\right] &(86)\\ &= \mathbb{E}\left[|H_n|^2|D_n|^2\right]\cdot\frac{1}{8\pi^2\gamma}\kappa = \frac{1}{2}, &(87)\end{aligned}$$

where

$$\begin{aligned}\kappa &= \int_0^{2\pi}\int_0^{2\pi}d(\angle G_n)d(\angle D_n)\int_{-\gamma}^{\gamma}\cos^2(\angle D_n - \angle G_n + \epsilon_n)d\epsilon_n\\ &= \int_0^{2\pi}\int_0^{2\pi}d(\angle G_n)d(\angle D_n)\int_{-\gamma}^{\gamma}\sin^2(\angle D_n - \angle G_n + \epsilon_n)d\epsilon_n\\ &= 4\pi^2\gamma.\end{aligned} \tag{88}$$

This yields (54).

References

[1] 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.

[2] 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.

[3] C. Huang, A. Zappone, G. C. Alexandropoulos, M. Debbah, and C. Yuen, "Reconfigurable intelligent surfaces for energy efficiency in wireless communication," *IEEE Trans. Wireless Commun.*, vol. 18, no. 8, pp. 4157–4170, Aug. 2019.

[4] Z. Lin et al., "Refracting RIS-aided hybrid satellite-terrestrial relay networks: Joint beamforming design and optimization," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 58, no. 4, pp. 3717–3724, Aug. 2022.

[5] X. Zhang, H. Zhang, K. Sun, K. Long, and Y. Li, "Human-centric irregular RIS-assisted multi-UAV networks with resource allocation and reflecting design for metaverse," *IEEE J. Sel. Areas Commun.*, vol. 42, no. 3, pp. 603–615, Mar. 2024.

[6] 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.

[7] 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.

[8] 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.

[9] 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.

[10] 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.

[11] C. Pu, A. Wall, K. R. Choo, I. Ahmed, and S. Lim, "A lightweight and privacy-preserving mutual authentication and key agreement protocol for Internet of Drones environment," *IEEE Internet Things J.*, vol. 9, no. 12, pp. 9918–9933, Jun. 2022.

[12] G. Bansal and B. Sikdar, "Fault resilient authentication architecture for drone networks," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops)*, May 2022, pp. 866–871.

[13] M. Mitev, M. Shakiba-Herfeh, A. Chorti, M. Reed, and S. Baghaee, "A physical layer, zero-round-trip-time, multifactor authentication protocol," *IEEE Access*, vol. 10, pp. 74555–74571, 2022.

[14] K. Lounis, S. H. H. Ding, and M. Zulkernine, "D2D-MAP: A drone to drone authentication protocol using physical unclonable functions," *IEEE Trans. Veh. Technol.*, vol. 72, no. 4, pp. 5079–5093, Apr. 2023.

[15] R. Meng, F. Zhu, X. Xu, B. Wang, B. Xu, and P. Zhang, "Efficient Gaussian process classification-based physical-layer authentication with configurable fingerprints for 6G-enabled IoT," 2023, *arXiv:2307.12263*.

[16] M. A. Shawky et al., "Reconfigurable intelligent surface-assisted cross-layer authentication for secure and efficient vehicular communications," 2023, *arXiv:2303.08911*.

[17] 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./Dec. 2023.

[18] 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.

[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. 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. 4th Int. Conf. Data Intell. Security (ICDIS)*, 2022, pp. 174–181.

[28] M. Di Renzo et al., "Smart radio environments empowered by reconfigurable intelligent surfaces: How it works, state of research, and road ahead," 2020, *arXiv:2004.09352*.

[29] W. Tang et al., "MIMO transmission through reconfigurable intelligent surface: System design, analysis, and implementation," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 11, pp. 2683–2699, Nov. 2020.

[30] M. Di Renzo et al., "Smart radio environments empowered by AI reconfigurable meta-surfaces: An idea whose time has come," 2019, *arXiv:1903.08925*.

[31] S. Abeywickrama, R. Zhang, and C. Yuen, "Intelligent reflecting surface: Practical phase shift model and beamforming optimization," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2020, pp. 1–6.

[32] A. L. Swindlehurst, G. Zhou, R. Liu, C. Pan, and M. Li, "Channel estimation with reconfigurable intelligent surfaces—A general framework," *Proc. IEEE*, vol. 110, no. 9, pp. 1312–1338, Sep. 2022.

[33] S. M. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory*. Upper Saddle River, NJ, USA: Prentice-Hall, 1993.

[34] 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.

[35] A. Casta no-Martínez and F. López-Blázquez, "Distribution of a sum of weighted noncentral chi-square variables," *Test*, vol. 14, no. 2, pp. 397–415, 2005.

Stefano Tomasin (Senior Member, IEEE) received the Ph.D. degree from the University of Padova, Italy, in 2003. He is currently a Full Professor with the University of Padova. During his career, he has visited IBM Research, Switzerland, Philips Research, The Netherlands, Qualcomm, CA, USA, Polytechnic University, Brooklyn, NY, 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. He was an Editor of IEEE TRANSACTIONS ON SIGNAL PROCESSING from 2017 to 2020 and *EURASIP Journal of Wireless Communications and Networking* and IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY.

Tarek N. M. M. Elwakeel received the double B.Sc. degree in electrical, electronics, and communications engineering from Modern Sciences and Arts University, Egypt, and Greenwich University, U.K., in 2015, and the M.Sc. degree in ICT for internet and multimedia from the University of Padova, Italy, in 2022. His thesis focused on enhancing security in 5G/6G networks. He is currently a Telecommunications Engineer. He has held key roles with HPE, Orange, and Etisalat, where he contributed to mobile network enhancement.

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. Since 2017, she has been a Visiting Scientist with BIOTEC, Technische Universität Dresden (TUD), Germany, and the University of California at 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.

Robin Maes received the B.S. and M.S. degrees in electrical engineering, with a specialization in communication and information theory from Ghent University. He wrote his dissertation under supervision during his Erasmus visit to the University of Padova, Italy.

Nele Noels (Senior Member, IEEE) received the Diploma degree in electrical engineering and the Ph.D. degree in electrical engineering from Ghent University, Ghent, Belgium, in 2001 and 2009, respectively. She is currently a Professor with the Department of Telecommunications and Information Processing (TELIN), Ghent University. She is the (co-)author of over 60 academic papers in international journals and conference proceedings and a recipient of several scientific awards. Her main research interests include statistical communication theory, carrier and symbol synchronization, bandwidth-efficient modulation and coding, and satellite and mobile communication. In 2010, she received the Scientific Award Alcatel Lucent Bell for the best Belgian thesis concerning an original study of information and communication technology, concepts, and/or applications. She is an Editor of Communication Theory and Systems (Division I) for *Journal of Communications and Networks*.

Marc Moeneclaey (Life Fellow, IEEE) received the Diploma and Ph.D. degrees in electrical engineering from Ghent University, Ghent, Belgium, in 1978 and 1983, respectively. He is currently a Professor Emeritus with the Department of Telecommunications and Information Processing (TELIN), Ghent University. He is the author of more than 500 scientific papers in international journals and conference proceedings. Together with Prof. H. Meyr (RWTH Aachen) and Dr. S. Fechtel (Siemens AG), he co-authored the book *Digital Communication Receivers Synchronization, Channel Estimation, and Signal Processing* (Wiley, 1998). His main research interests include statistical communication theory, carrier and symbol synchronization, channel estimation, bandwidth-efficient modulation and coding, and multi-antenna wireless communication. He was a co-recipient of the Mannesmann Innovations Prize 2000. From 1992 to 1994, he was an Editor of *Synchronization* and IEEE TRANSACTIONS ON COMMUNICATIONS. He has served as the Co-Guest Editor for Special Issues on *Wireless Personal Communications* (Equalization and Synchronization in Wireless Communications) in 1998 and IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS (Signal Synchronization in Digital Transmission Systems) in 2001.

Open Access funding provided by ‘Università degli Studi di Padova’ within the CRUI CARE Agreement