Divergence-Minimizing Attack Against Challenge-Response Authentication with IRSs

Laura Crosara, Anna V. Guglielmi, Nicola Laurenti, and Stefano Tomasin
Dept. of Information Engineering (DEI), University of Padova, Italy
email: {laura.crosara.1@phd. , annavaleria.guglielmi@ , stefano.tomasin@, nicola.laurenti@ }unipd.it

2024 IEEE International Conference on Communications Workshops (ICC Workshops) | 979-8-3503-0405-3/24/$31.00 ©2024 IEEE | DOI: 10.1109/ICCWORKSHOPS59551.2024.10615572

***Abstract*—We propose a new attack against challenge response physical layer authentication (CR-PLA) with intelligent reflecting surfaces (IRSs). Drawing from prior work, we establish bounds on performance metrics, such as probabilities of false alarm and missed detection, using Kullback-Leibler (KL) divergence. Leveraging prior results in [1], we extend the analysis to the CR-PLA scenario with IRS. We derive the optimal attack strategy to minimize the divergence between authentic and forged signals when the attacker has either partial or no knowledge of the legitimate cascade channel. We evaluate the attack performance under different conditions, by varying the correlation between Eve's observation and the legitimate cascade channel, or the SNR at the legitimate receiver.**

***Index Terms*—Authentication, Challenge-response, Intelligent Reflecting Surfaces.**

I. Introduction

Source authentication is the problem of establishing if a received message truly comes from the declared sender or has been forged by an impersonating attacker. Risks in accepting unauthenticated messages go from denial of service to privacy to the loss of control of devices, e.g., in Internet of Things (IoT) contexts.

Several authentication mechanisms have been explored, beyond those operating at the application level and using cryptographic approaches. Here we focus on physical layer authentication (PLA), which leverages the propagation characteristics of the physical channel as signatures of the transmitting devices or the communication links. The basic approach, introduced by Simmons [2] includes two phases, the identification acquisition, and the identification verification. In the former, the receiver Bob (verifier) estimates the channel using signals transmitted by Alice (the authentic source) that are authenticated at higher layers (e.g., by cryptographic mechanisms). In the latter, whenever Bob receives a new message, he also estimates the channel over which the signal traveled and compares this estimate with that obtained in the first phase. If the two estimates are consistent (note that they are still affected by noise), then the received message is deemed authentic, otherwise it is considered fake. 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 tests [6] to machine learning approaches [7]. For an overview on PLA see also [8], [9].

This work has been funded in part by the European Commission through the Horizon Europe/JU SNS project ROBUST-6G (Grant Agreement no. 101139068).

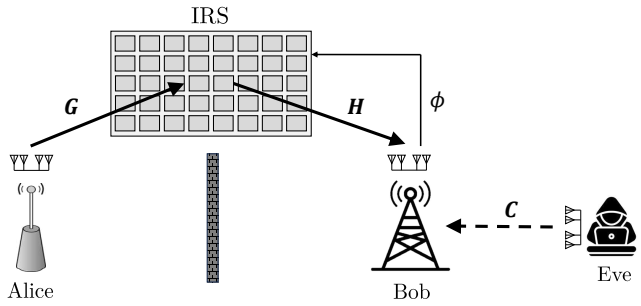

Fig. 1. Communication scenario.

Recently, a further evolution of PLA has been introduced by exploiting the controllable nature of wireless channels provided by new communication technologies. In particular, intelligent reflecting surfaces (IRSs) are controllable devices that can change the propagation of wireless signals by changing the phase shift introduced by their elements. When the IRS is under Bob's control, he can set a random configuration of the IRS which remains secret to the attacker, and verify that the estimated channel on a received message corresponds to the set configuration, [10]. This approach provides a *challenge response PLA (CR-PLA)* mechanism and can be applied also when other *controllable channels* are available (e.g., Bob is a drone that can change its position, [11]).

In this paper, we investigate novel attacks to be performed when the CR-PLA uses an IRS to perform the challenge-response approach. In particular, we leverage the results of [1, Th. 2] that has established bounds on the performance for a conventional (non-interactive) PLA mechanism, in terms of probabilities of false alarm (FA) and missed detection (MD). The bounds exploit the Kullback-Leibler (KL) divergence of observed channels at Alice, Bob, and Eve, and it turns out that when the observations in the legitimate case are jointly Gaussian distributed, the optimal attack strategy is also Gaussian distributed. Here we derive the bounds for the considered CR-PLA scenario with IRS and, under the assumption of a large number of IRS elements, we ensure that the assumption

in [1, Th. 2] are satisfied and derive the optimal attack by Eve. We assess the performance of the obtained attack by considering different correlations between Eve's observation and the legitimate cascade channel, and different signal-to-noise ratio (SNR) values for the legitimate channel.

The rest of the paper is organized as follows. Section II presents the system model. Section III describes the CR-PLA mechanism. Then, the KL divergence minimizing attack is derived in Section IV, considering no channel knowledge at Eve in Section IV-A, and partial channel knowledge in Section IV-B. Then, numerical results are discussed in Section V and, finally, conclusions are drawn in Section VI.[1]

II. System Model

We consider the scenario depicted in Fig. 1, where a legitimate receiver Bob authenticates messages from a legitimate transmitter Alice. We account for the presence of an attacker Eve, who aims to impersonate Alice by forging messages and transmitting them to Bob. We assume that Alice, Bob, and Eve are equipped with a uniform linear array (ULA) of N_{A}, N_{B}, and N_{E} antennas, respectively.

We further assume that the communication between Alice and Bob is supported by an IRS with N reflecting elements, each acting as a receive and transmit antenna. The n-th element, $n = 1, 2, \ldots, N$, of the IRS introduces a phase shift $\phi_n = e^{j\theta_n}$ on the equivalent baseband signal and has unitary gain. We define vector $\boldsymbol{\phi} = [\phi_1, \phi_2, \ldots, \phi_N]^{\mathrm{T}}$ as the *configuration* that collects the phase shifts introduced by the IRS on each element, and the $N \times N$ diagonal matrix

$$\boldsymbol{\Phi} = \mathrm{diag}\{\boldsymbol{\phi}\} = \mathrm{diag}\{\phi_1, \phi_2, \ldots, \phi_N\}\,. \tag{1}$$

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

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

We assume that Eve can directly transmit to Bob through the channel $\boldsymbol{C}$, whose channel matrix is known by Eve.

Note that all channels are time-invariant and reciprocal, while the IRS configuration that reflects onto matrix $\boldsymbol{\Phi}$ is under Bob's control and can be changed over time, making the cascade channels controllable.

[1]Notation: † denotes the Moore-Penrose inverse, $^{\mathrm{H}}$ represents complex conjugate and transpose, $|\boldsymbol{A}|$ stands for the determinant of $\boldsymbol{A}$, and log is the natural logarithm. If $\boldsymbol{a} \in \mathbb{C}^n$ and $\boldsymbol{b} \in \mathbb{C}^m$ are random vectors, $\boldsymbol{K}_{ab} = \mathbb{E}[\boldsymbol{a}\boldsymbol{b}^{\mathrm{H}}]$ is their $n \times m$ covariance matrix. The vectorization operator $[\mathrm{vec}\,(\boldsymbol{X})]_i = [\boldsymbol{X}]_{\lfloor \frac{i}{C} \rfloor, i \bmod C}$, $i = 1, \ldots, RC$, converts matrix $\boldsymbol{X}$ of size $(R \times C)$ to a column vector of length RC by indexing the matrix by rows.

III. Challenge-Response PLA with IRS

The CR-PLA mechanism introduced in [10] works as follows. In the *association* phase Alice transmits several authenticated pilot signals and Bob estimates the cascade channel $\boldsymbol{Q}_{\mathrm{AIB}}$ for various IRS configurations. We assume that this phase is long enough to let Bob obtain estimates $\bar{\boldsymbol{Q}}_{\mathrm{AIB}}(\boldsymbol{\Phi'})$ of the cascade channel for any IRS configuration $\boldsymbol{\phi}'$.

In the *verification* phase, Bob sets a random configuration $\boldsymbol{\Phi}$ of the IRS as follows. Bob modifies the IRS configuration around the communication-optimal[2] one $\bar{\theta}_n$, $n = 1, \ldots, N$, obtaining a new configuration as

$$\phi'_n = e^{j(\bar{\theta}_n + \epsilon_n)}\,, \quad n = 1, \ldots, N\,, \tag{3}$$

with $\epsilon_n \sim \mathcal{U}[-\gamma, \gamma]$, where γ denotes the maximum deviation from the communication-optimal IRS configuration. Thus, a larger γ introduces more randomness in the IRS configuration making it more difficult for Eve to build an attack that passes the authentication check. Moreover, ϵ_n are independent for each $n = 1, \ldots, N$, and regenerated independently at each new verification phase. Whenever Bob receives a message, he estimates the cascade channel $\boldsymbol{R}$ and checks if it is consistent with the expected channel $\bar{\boldsymbol{Q}}_{\mathrm{AIB}}(\boldsymbol{\Phi})$. Pilot signals are assumed to be known by all parties.

Under legitimate conditions, given the configuration $\boldsymbol{\phi}$, when Alice transmits, Bob obtains the channel estimate

$$\hat{\boldsymbol{Q}}_{\mathrm{AIB}} = \boldsymbol{Q}_{\mathrm{AIB}} + \boldsymbol{W}_{\mathrm{B}}\,, \tag{4}$$

where $\boldsymbol{W}_{\mathrm{B}}$ denotes the estimation error at Bob, modeled as additive white Gaussian noise (AWGN) with zero mean, independent entries and power σ^2_{B} per entry.

We assume that Bob has perfect knowledge of both channel matrices and can choose the IRS configuration $\boldsymbol{\Phi}$. Thus, Bob during the association phase obtains $\bar{\boldsymbol{Q}}_{\mathrm{AIB}}(\boldsymbol{\Phi}) = \boldsymbol{Q}_{\mathrm{AIB}}(\boldsymbol{\Phi})$.

Bob's goal is to figure out whether the estimated channel $\boldsymbol{R}$ is authentic ($\hat{\boldsymbol{Q}}_{\mathrm{AIB}}$), or forged ($\boldsymbol{V}$), by using his knowledge of $\boldsymbol{Q}_{\mathrm{AIB}}$. To detect the attack, Bob performs an authentication test, that, given $\boldsymbol{Q}_{\mathrm{AIB}}$ and $\boldsymbol{R}$, decides between the hypotheses:

$$\mathcal{H}_0: \text{ the message is from Alice,} \tag{5}$$

$$\mathcal{H}_1: \text{ the message is from Eve.} \tag{6}$$

The authentication procedure is summarized in block D of Fig. 2, which has $\boldsymbol{r}$ as input and outputs the Boolean value $\hat{b}$. Correct verification is achieved when $\hat{b} = b$.

A. Attack Model

The purpose of Eve is to impersonate Alice, i.e., to pass the authentication check at the legitimate receiver. We consider that some side information is available to Eve, represented by the matrix $\boldsymbol{Z}$. We assume that $\boldsymbol{C}$ is correlated with $\boldsymbol{H}$ and/or $\boldsymbol{G}$, and Eve can estimate these channel matrices and compute (an estimate of) the communication-optimal IRS configuration $\boldsymbol{\Phi}^\star = \mathrm{diag}\{e^{j\bar{\theta}_1}, \ldots, e^{j\bar{\theta}_N}\}$. A simple model of the resulting

[2]For further insights on optimal IRS configuration refer to [12].

estimate at Eve of the Alice-IRS-Bob cascade channel with communication-optimal IRS configuration is

$$\boldsymbol{Z} = \boldsymbol{H}\boldsymbol{\Phi}^{\star}\boldsymbol{G} + \boldsymbol{W}_{\boldsymbol{z}}\,, \tag{7}$$

where $\boldsymbol{W}_{\boldsymbol{z}}$ is an AWGN matrix modeling the resulting estimation error. In Section IV we will analyze the case where Eve's side information $\boldsymbol{Z}$ and the legitimate cascade channel $\boldsymbol{X}$ are either independent or correlated.

We assume that Eve knows: i) the actual realization of $\boldsymbol{Z}$, ii) the probability density function (pdf) of the channel matrices $\boldsymbol{H}$ and $\boldsymbol{G}$, and of the IRS configuration vector $\boldsymbol{\phi}$, and iii) the pdf of noise at both receivers. We assume that Eve transmits precoded pilot signals and can induce any channel estimate (apart from the presence of noise) to Bob. Then, the channel estimated by Bob when under attack is

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

where $\boldsymbol{V}_0$ is the channel selected by Eve for the attack. For the sake of generality, we consider that Eve adopts a probabilistic strategy, characterized by the conditional pdf $p_{\boldsymbol{V}_0|\boldsymbol{Z}}$. Moreover, since Eve knows the statistics of $\boldsymbol{W}_{\mathrm{B}}$, the attack strategy can be described by $p_{\boldsymbol{V}|\boldsymbol{Z}}$.

Finally, let the channel estimated at Bob be

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

where b indicates the legitimate/attack state. Therefore, Eve aims to prevent Bob from distinguishing between the attack $\boldsymbol{V}$ and the legitimate $\hat{\boldsymbol{Q}}_{\mathrm{AIB}}$. For ease of notation, let $\boldsymbol{v} = \mathrm{vec}(\boldsymbol{V})$ and $\boldsymbol{r} = \mathrm{vec}(\boldsymbol{R})$, as also shown in Fig. 2.

From now on, to compact the notation, we let

$$\boldsymbol{x} = \mathrm{vec}(\boldsymbol{Q}_{\mathrm{AIB}})\,, \tag{10}$$

$$\boldsymbol{y} = \mathrm{vec}(\hat{\boldsymbol{Q}}_{\mathrm{AIB}}) = \boldsymbol{x} + \boldsymbol{\omega}\,, \tag{11}$$

$$\boldsymbol{z} = \mathrm{vec}(\boldsymbol{Z})\,, \tag{12}$$

where $\boldsymbol{\omega} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{K}_{\omega})$, $\boldsymbol{K}_{\omega} = \sigma_{\mathrm{B}}^2\boldsymbol{I}$. So, $\boldsymbol{x}$ is the information available to Bob during the verification phase, $\boldsymbol{y}$ is the legitimate channel estimation, and $\boldsymbol{z}$ contains the information Eve can leverage to build the attack. Moreover, we refer to the abstract representation of the authentication scenario in Fig. 2.

B. Performance Metrics

An authentication system's performance is typically evaluated through type-I (FA) error probability p_{FA}, i.e., the probability that Bob rejects a legitimate message from Alice, and type-II (MD) error probability p_{MD}, i.e., the probability that Bob accepts a message from Eve as authentic. The analytical derivation of such probabilities, depending on the verification decision rule, is often impractical. Thus, establishing bounds on the achievable error probability region, i.e., the set of feasible points in the $(p_{\mathrm{FA}}, p_{\mathrm{MD}})$ plane, is useful to prove the effectiveness of practical schemes.

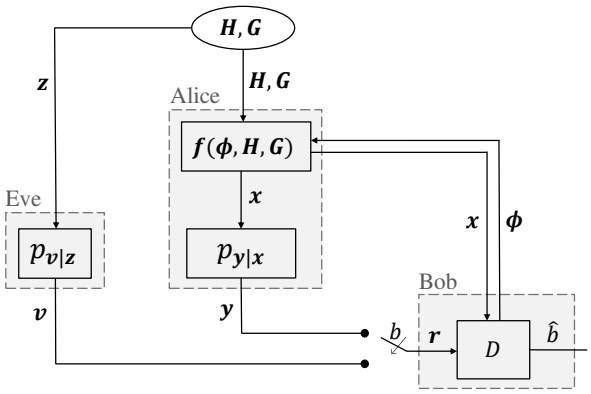

Fig. 2. Block scheme.

An outer bound is obtained by the KL divergence among the estimated channels at Alice, Bob, and Eve. From [6], as a consequence of the data processing inequality, we have

$$\mathbb{D}(p_{\hat{b}|\mathcal{H}_1} \| p_{\hat{b}|\mathcal{H}_0}) \leq \mathbb{D}(p_{\boldsymbol{r}\boldsymbol{x}|\mathcal{H}_1} \| p_{\boldsymbol{r}\boldsymbol{x}|\mathcal{H}_0}) = \mathbb{D}(p_{\boldsymbol{x}\boldsymbol{v}} \| p_{\boldsymbol{x}\boldsymbol{y}}), \tag{13}$$

where $\mathbb{D}(p\|q)$ is the KL divergence between pdfs p and q. It also holds the symmetric bound

$$\mathbb{D}(p_{\hat{b}|\mathcal{H}_0} \| p_{\hat{b}|\mathcal{H}_1}) \leq \mathbb{D}(p_{\boldsymbol{r}\boldsymbol{x}|\mathcal{H}_0} \| p_{\boldsymbol{r}\boldsymbol{x}|\mathcal{H}_1}) = \mathbb{D}(p_{\boldsymbol{x}\boldsymbol{y}} \| p_{\boldsymbol{x}\boldsymbol{v}}). \tag{14}$$

In (13) and (14) we have considered $p_{\boldsymbol{r}\boldsymbol{x}}$ based on the assumption that, at the time of verification, the legitimate receiver knows $\boldsymbol{x}$, and the decision $\hat{b}$ is made on both inputs. Therefore, observing that $p_{\hat{b}|\mathcal{H}_0}(1) = p_{\mathrm{FA}}$, $p_{\hat{b}|\mathcal{H}_1}(0) = p_{\mathrm{MD}}$, and defining the function

$$h(p_{\mathrm{MD}}, p_{\mathrm{FA}}) \triangleq p_{\mathrm{MD}} \log \frac{p_{\mathrm{MD}}}{1 - p_{\mathrm{FA}}} + (1 - p_{\mathrm{MD}}) \log \frac{1 - p_{\mathrm{MD}}}{p_{\mathrm{FA}}}\,, \tag{15}$$

(13) can be rewritten as

$$h(p_{\mathrm{MD}}, p_{\mathrm{FA}}) \leq \mathbb{D}(p_{\boldsymbol{x}\boldsymbol{v}} \| p_{\boldsymbol{x}\boldsymbol{y}}). \tag{16}$$

This limits the region of achievable $(p_{\mathrm{FA}}, p_{\mathrm{MD}})$ values, depending on $\mathbb{D}(p_{\boldsymbol{x}\boldsymbol{v}} \| p_{\boldsymbol{x}\boldsymbol{y}})$, i.e., the statistics of the observations at Bob and Eve, for any decision mechanism. The use of challenge-response authentication has an impact on $\boldsymbol{x}$ (and $\boldsymbol{y}$) through the additional randomness introduced by the choice of the IRS configuration in the verification phase. Since this randomness is not captured by $\boldsymbol{z}$, the challenge-response mechanism increases the KL divergence at Bob's benefit. Thus, a higher value of γ in the model (3) for the IRS randomness will yield a larger KL divergence.

The aim of the attacker Eve is to narrow the achievable region, by making the value of $\mathbb{D}(p_{\boldsymbol{x}\boldsymbol{v}} \| p_{\boldsymbol{x}\boldsymbol{y}})$ as small as possible, operating on the attack strategy $p_{\boldsymbol{v}|\boldsymbol{z}}$. The metric $\mathbb{D}(p_{\boldsymbol{x}\boldsymbol{v}} \| p_{\boldsymbol{x}\boldsymbol{y}})$ depends on the attack strategy as

$$\mathbb{D}(p_{\boldsymbol{x}\boldsymbol{v}} \| p_{\boldsymbol{x}\boldsymbol{y}}) \;=\; \mathbb{E}\left[\log \frac{\int p_{\boldsymbol{v}|\boldsymbol{z}}(\boldsymbol{v}|\boldsymbol{a}) p_{\boldsymbol{z}|\boldsymbol{x}}(\boldsymbol{a}|\boldsymbol{x}) \mathrm{d}\boldsymbol{a}}{p_{\boldsymbol{y}|\boldsymbol{x}}(\boldsymbol{v}|\boldsymbol{x})}\right]. \tag{17}$$

IV. Divergence Minimizing Attack

In this Section, we derive the attack strategy that minimizes the KL divergence, thus abstracting from the particular detection process, whose efficacy can be assessed beforehand. We consider two scenarios: 1) Eve has no information about the Alice-IRS-Bob channel, i.e., $\boldsymbol{z}$ and $\boldsymbol{x}$ are independent; 2) Eve has partial information about the Alice-IRS-Bob channel, thus $\boldsymbol{z}$ is correlated with $\boldsymbol{x}$.

Now, assuming that the number of IRS elements N is large enough, since each element $x_{p,q}$ of the matrix $\boldsymbol{Q}_{\mathrm{AIB}}$ is given by $x_{p,q} = \sum_{i=1}^{N} \phi_i \boldsymbol{H}_{p,i} \boldsymbol{G}_{i,q}$, considering the model (3) for the IRS configuration and invoking the central limit theorem, we can model $\boldsymbol{x}$ as a Gaussian distributed random vector with vector mean $\boldsymbol{\mu}_x$ and covariance matrix $\boldsymbol{K}_x$. So, $\boldsymbol{x}$ and $\boldsymbol{y}$ are jointly Gaussian random vectors with covariance matrix

$$\boldsymbol{K}_{\left[\begin{smallmatrix}\boldsymbol{x}\\ \boldsymbol{y}\end{smallmatrix}\right]} = \begin{bmatrix} \boldsymbol{K}_{\boldsymbol{x}} & \boldsymbol{K}_{\boldsymbol{xy}} \\ \boldsymbol{K}_{\boldsymbol{yx}} & \boldsymbol{K}_{\boldsymbol{y}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{K}_{\boldsymbol{x}} & \boldsymbol{K}_{\boldsymbol{x}} \\ \boldsymbol{K}_{\boldsymbol{x}} & \boldsymbol{K}_{\boldsymbol{x}} + \boldsymbol{K}_{\boldsymbol{\omega}} \end{bmatrix}. \tag{18}$$

A. No Channel Knowledge

In the zero channel knowledge hypothesis, $\boldsymbol{x}$ and $\boldsymbol{v}$ are independent, thus, $p_{\boldsymbol{xv}} = p_{\boldsymbol{x}} p_{\boldsymbol{v}}$. Under these conditions, $\mathbb{D}(p_{\boldsymbol{xy}} \| p_{\boldsymbol{xv}})$ becomes tractable, hence we consider two cases for the design of the attack: *a)* Eve aims at minimizing $\mathbb{D}(p_{\boldsymbol{xv}} \| p_{\boldsymbol{xy}})$ or *b)* Eve aims at minimizing $\mathbb{D}(p_{\boldsymbol{xy}} \| p_{\boldsymbol{xv}})$.

Considering task *a)*, the optimal joint pdf $p_{\boldsymbol{xv}}$ that minimizes $\mathbb{D}(p_{\boldsymbol{xv}} \| p_{\boldsymbol{xy}})$ is also Gaussian [1, Th. 2]. When the attack strategy is $p_{\boldsymbol{v}} \sim \mathcal{N}(\boldsymbol{\mu}_{\boldsymbol{v}}, \boldsymbol{K}_v)$, we have

$$\begin{aligned}\mathbb{D}(p_{\boldsymbol{x}} p_{\boldsymbol{v}} \| p_{\boldsymbol{xy}}) &= \mathbb{E}\left[\log \frac{p_{\boldsymbol{v}}(\boldsymbol{v})}{p_{\boldsymbol{y}|\boldsymbol{x}}(\boldsymbol{v}|\boldsymbol{x})}\right] = \mathbb{E}\left[\log \frac{p_{\boldsymbol{v}}(\boldsymbol{v})}{p_{\boldsymbol{\omega}}(\boldsymbol{v}-\boldsymbol{x})}\right] \\ &= \mathbb{E}\left[\log \frac{p_{\boldsymbol{v}}(\boldsymbol{v})}{p_{\boldsymbol{\omega}}(\boldsymbol{v})}\right] + \mathbb{E}\left[\log \frac{p_{\boldsymbol{\omega}}(\boldsymbol{v})}{p_{\boldsymbol{\omega}}(\boldsymbol{v}-\boldsymbol{x})}\right] \end{aligned} \tag{19}$$

The first term in (19) is the KL divergence $\mathbb{D}(p_{\boldsymbol{v}} \| p_{\boldsymbol{\omega}})$, while the second term can be computed as

$$\begin{aligned}\mathbb{E}\left[\log \frac{p_{\boldsymbol{\omega}}(\boldsymbol{v})}{p_{\boldsymbol{\omega}}(\boldsymbol{v}-\boldsymbol{x})}\right] = \frac{1}{2}\Big[&\operatorname{tr}(\boldsymbol{K}_{\boldsymbol{x}} \boldsymbol{K}_{\boldsymbol{\omega}}^{-1}) \\ &- \boldsymbol{\mu}_{\boldsymbol{x}}^{\mathrm{H}} \boldsymbol{K}_{\boldsymbol{\omega}}^{-1} \boldsymbol{\mu}_{\boldsymbol{v}} - \boldsymbol{\mu}_{\boldsymbol{v}}^{\mathrm{H}} \boldsymbol{K}_{\boldsymbol{\omega}}^{-1} \boldsymbol{\mu}_{\boldsymbol{x}}\Big]. \end{aligned} \tag{20}$$

Replacing (20) in (19) we obtain

$$\begin{aligned}\mathbb{D}(p_{\boldsymbol{x}} p_{\boldsymbol{v}} \| p_{\boldsymbol{xy}}) = \frac{1}{2}\Big[&\log \frac{|\boldsymbol{K}_{\boldsymbol{\omega}}|}{|\boldsymbol{K}_{\boldsymbol{v}}|} + \operatorname{tr}(\boldsymbol{K}_{\boldsymbol{v}} \boldsymbol{K}_{\boldsymbol{\omega}}^{-1}) - N_{\mathrm{A}} N_{\mathrm{B}} \\ &+ \operatorname{tr}(\boldsymbol{K}_{\boldsymbol{x}} \boldsymbol{K}_{\boldsymbol{\omega}}^{-1}) + (\boldsymbol{\mu}_{\boldsymbol{v}} - \boldsymbol{\mu}_{\boldsymbol{x}})^{\mathrm{H}} \boldsymbol{K}_{\boldsymbol{\omega}}^{-1} (\boldsymbol{\mu}_{\boldsymbol{v}} - \boldsymbol{\mu}_{\boldsymbol{x}})\Big], \end{aligned} \tag{21}$$

which is minimized for $\boldsymbol{\mu}_{\boldsymbol{v}} = \boldsymbol{\mu}_{\boldsymbol{x}}$, $\boldsymbol{K}_{\boldsymbol{v}} = \boldsymbol{K}_{\boldsymbol{\omega}}$. So, when Eve has no clue about the cascade channel Alice-IRS-Bob the optimal probabilistic attack to minimize $\mathbb{D}(p_{\boldsymbol{xv}} \| p_{\boldsymbol{xy}})$ is

$$p_{\boldsymbol{v}}^{\star} \sim \mathcal{N}(\boldsymbol{\mu}_{\boldsymbol{x}}, \boldsymbol{K}_{\omega}) \tag{22}$$

where Eve mimics the estimation error distribution at Bob, centered at $\boldsymbol{\mu}_{\boldsymbol{x}}$. When Eve strategy is $p_{\boldsymbol{v}}^{\star}$, (19) becomes

$$\mathbb{D}(p_{\boldsymbol{x}} p_{\boldsymbol{v}}^{\star} \| p_{\boldsymbol{xy}}) = \frac{1}{2} \operatorname{tr}\left(\boldsymbol{K}_{\boldsymbol{x}} \boldsymbol{K}_{\boldsymbol{\omega}}^{-1}\right) = \frac{1}{2} \sigma_{\mathrm{B}}^{-2} \operatorname{tr}\left(\boldsymbol{K}_{\boldsymbol{x}}\right). \tag{23}$$

Moreover, with this attack strategy, we also get $\mathbb{D}(p_{\boldsymbol{x}} p_{\boldsymbol{v}}^{\star} \| p_{\boldsymbol{xy}}) = \mathbb{D}(p_{\boldsymbol{xy}} \| p_{\boldsymbol{x}} p_{\boldsymbol{v}}^{\star})$.

Considering now case *b)*, Eve aims to find the pdf $p_{\boldsymbol{v}}$ that minimizes $\mathbb{D}(p_{\boldsymbol{xy}} \| p_{\boldsymbol{xv}})$. For any joint pdf $p_{\boldsymbol{xy}}$, even non Gaussian, we have $\mathbb{D}(p_{\boldsymbol{xy}} \| p_{\boldsymbol{xv}}) = \mathbb{D}(p_{\boldsymbol{xy}} \| p_{\boldsymbol{x}} p_{\boldsymbol{v}})$, and

$$\begin{aligned}&\mathbb{D}(p_{\boldsymbol{xy}} \| p_{\boldsymbol{x}} p_{\boldsymbol{v}}) = \\ &= \mathbb{E}\left[\log \frac{p_{\boldsymbol{xy}}(\boldsymbol{x},\boldsymbol{y})}{p_{\boldsymbol{x}}(\boldsymbol{x}) p_{\boldsymbol{v}}(\boldsymbol{y})}\right] = \mathbb{E}\left[\log \frac{p_{\boldsymbol{xy}}(\boldsymbol{x},\boldsymbol{y})}{p_{\boldsymbol{x}}(\boldsymbol{x}) p_{\boldsymbol{y}}(\boldsymbol{y})} \frac{p_{\boldsymbol{y}}(\boldsymbol{y})}{p_{\boldsymbol{v}}(\boldsymbol{y})}\right] \\ &= \mathbb{I}(\boldsymbol{x};\boldsymbol{y}) + \mathbb{D}(p_{\boldsymbol{y}} \| p_{\boldsymbol{v}}) \geq \mathbb{I}(\boldsymbol{x};\boldsymbol{y}), \end{aligned} \tag{24}$$

where $\mathbb{I}(x;y)$ denotes the mutual information between random variables x and y. Therefore, the optimal probabilistic attack strategy is

$$p_{\boldsymbol{v}}^{\star} = \arg\min_{p_{\boldsymbol{v}}} \mathbb{D}(p_{\boldsymbol{xy}} \| p_{\boldsymbol{x}} p_{\boldsymbol{v}}) = \arg\min_{p_{\boldsymbol{v}}} \mathbb{D}(p_{\boldsymbol{y}} \| p_{\boldsymbol{v}}) = p_{\boldsymbol{y}} \tag{25}$$

and the KL divergence in the jointly Gaussian case (18) is

$$\mathbb{D}(p_{\boldsymbol{xy}} \| p_{\boldsymbol{x}} p_{\boldsymbol{v}}^{\star}) = \mathbb{I}(\boldsymbol{x};\boldsymbol{y}) = \frac{1}{2} \log\left(\frac{|\boldsymbol{K}_{\boldsymbol{x}} + \boldsymbol{K}_{\boldsymbol{\omega}}|}{|\boldsymbol{K}_{\boldsymbol{\omega}}|}\right). \tag{26}$$

B. Partial Channel Knowledge

When Eve's observations of her channel to Bob exhibit a correlation with the legitimate cascade channel, she can leverage this information in her attack. Since the observation $\boldsymbol{z}$ encloses all the information Eve can exploit to deceive Bob, the attack strategy $\boldsymbol{v}$ is conditionally independent of $\boldsymbol{x}$, given $\boldsymbol{z}$. We assume that $\boldsymbol{z}$ is jointly Gaussian distributed with $\boldsymbol{x}$ and $\boldsymbol{y}$, with joint covariance matrix

$$\boldsymbol{K}_{\left[\begin{smallmatrix}\boldsymbol{x}\\ \boldsymbol{y}\\ \boldsymbol{z}\end{smallmatrix}\right]} = \begin{bmatrix} \boldsymbol{K}_{\boldsymbol{x}} & \boldsymbol{K}_{\boldsymbol{xy}} & \boldsymbol{K}_{\boldsymbol{xz}} \\ \boldsymbol{K}_{\boldsymbol{yx}} & \boldsymbol{K}_{\boldsymbol{y}} & \boldsymbol{K}_{\boldsymbol{yz}} \\ \boldsymbol{K}_{\boldsymbol{zx}} & \boldsymbol{K}_{\boldsymbol{zy}} & \boldsymbol{K}_{\boldsymbol{z}} \end{bmatrix}. \tag{27}$$

To simplify the upcoming calculation, without loss of generality, we assume that $\boldsymbol{x}$, $\boldsymbol{y}$, and $\boldsymbol{z}$ are zero-mean[3]. In this scenario, we consider the bound in (13) since it is mathematically more tractable than (14).

From [1, Th. 2] and [13], the attack that minimizes $\mathbb{D}(p_{\boldsymbol{xv}} \| p_{\boldsymbol{xy}})$ when $\boldsymbol{x}$, $\boldsymbol{y}$, and $\boldsymbol{z}$ are zero mean jointly Gaussian random vectors is

$$p_{\boldsymbol{v}|\boldsymbol{z}}^{\star} \sim \mathcal{N}(\boldsymbol{G}\boldsymbol{z},\, \boldsymbol{C}\boldsymbol{C}^{\mathbf{H}} + \boldsymbol{K}_{\boldsymbol{\omega}}). \tag{28}$$

Thus, the optimal attack for KL divergence minimization is

$$\boldsymbol{v}^{\star} = \boldsymbol{G}\boldsymbol{z} + \boldsymbol{C}\boldsymbol{\omega}_{\mathbf{c}} + \boldsymbol{\omega}, \tag{29}$$

where $\boldsymbol{\omega}_{\mathbf{c}} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{I})$. Defining $\boldsymbol{P} = \boldsymbol{K}_{\boldsymbol{xz}}^{\mathrm{H}} \boldsymbol{K}_{\boldsymbol{x}}^{-1} \boldsymbol{K}_{\boldsymbol{xz}}$, we have

$$\boldsymbol{G} = \boldsymbol{K}_{\boldsymbol{xz}} \boldsymbol{P}^{\dagger}, \tag{30}$$

$$\boldsymbol{C}\boldsymbol{C}^{\mathrm{H}} = \boldsymbol{K}_{\boldsymbol{xz}} \boldsymbol{P}^{\dagger} (\boldsymbol{I} - \boldsymbol{K}_{\boldsymbol{z}} \boldsymbol{P}^{\dagger}) \boldsymbol{K}_{\boldsymbol{xz}}^{\mathrm{H}}. \tag{31}$$

We remark that $\boldsymbol{C}\boldsymbol{C}^{\mathrm{H}}$ is obtained with the closed form expression in (31) if the result is a positive-semidefinite matrix, otherwise $\boldsymbol{C}\boldsymbol{C}^{\mathrm{H}}$ is computed through an iterative process [1,

[3]Note that every agent is aware of the statistics of the random vectors, allowing the mean to be subtracted and re-added at a later stage.

Sec. IV]. We now compute $\mathbb{D}(p_{\boldsymbol{xv}}\|p_{\boldsymbol{xy}})$ under this divergence minimizing attack.

First, we observe that $\boldsymbol{z}$ can be expressed as a linear transformation of $\boldsymbol{x}$ plus an AWGN vector $\tilde{\boldsymbol{\omega}}$ with an appropriately tuned covariance matrix,

$$\boldsymbol{z} = \boldsymbol{F}\boldsymbol{x} + \tilde{\boldsymbol{\omega}}\,, \tag{32}$$

where $\boldsymbol{F} = \boldsymbol{K}_{\boldsymbol{xz}}^{\mathrm{H}}\boldsymbol{K}_{\boldsymbol{x}}^{-1}$ and $\boldsymbol{K}_{\tilde{\omega}} = \boldsymbol{K}_{\boldsymbol{z}} - \boldsymbol{P}$. Defining $\boldsymbol{B} \triangleq \boldsymbol{G}\boldsymbol{F}$ and $\boldsymbol{\eta} = \boldsymbol{G}\tilde{\boldsymbol{\omega}} + \boldsymbol{C}\boldsymbol{\omega}_{\mathbf{c}} + \boldsymbol{\omega}$, (29) is rewritten as

$$\boldsymbol{v}^{\star} = \boldsymbol{B}\boldsymbol{x} + \boldsymbol{\eta}\,, \tag{33}$$

with $\boldsymbol{\eta} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{K}_{\boldsymbol{\eta}})$. Now, considering $\boldsymbol{G}$ and $\boldsymbol{C}\boldsymbol{C}^{\mathrm{H}}$ defined as in (30) and (31) we have

$$\begin{aligned}\boldsymbol{K}_{\boldsymbol{\eta}} &= \boldsymbol{G}\boldsymbol{K}_{\tilde{\omega}}\boldsymbol{G}^{\mathrm{H}} + \boldsymbol{C}\boldsymbol{C}^{\mathrm{H}} + \boldsymbol{K}_{\boldsymbol{\omega}} = \\ &= \boldsymbol{K}_{\boldsymbol{xz}}\boldsymbol{P}^{\dagger}\boldsymbol{K}_{\boldsymbol{z}}\boldsymbol{P}^{\dagger}\boldsymbol{K}_{\boldsymbol{xz}}^{\mathrm{H}} - \boldsymbol{K}_{\boldsymbol{xz}}\boldsymbol{P}^{\dagger}\boldsymbol{P}\boldsymbol{P}^{\dagger}\boldsymbol{K}_{\boldsymbol{xz}}^{\mathrm{H}} \\ &+ \boldsymbol{K}_{\boldsymbol{xz}}\boldsymbol{P}^{\dagger}\boldsymbol{K}_{\boldsymbol{xz}}^{\mathrm{H}} - \boldsymbol{K}_{\boldsymbol{xz}}\boldsymbol{P}^{\dagger}\boldsymbol{K}_{\boldsymbol{z}}\boldsymbol{P}^{\dagger}\boldsymbol{K}_{\boldsymbol{xz}}^{\mathrm{H}} + \boldsymbol{K}_{\boldsymbol{\omega}} = \boldsymbol{K}_{\boldsymbol{\omega}}\,.\end{aligned} \tag{34}$$

When (31) is positive semidefinite, Eve gets $\boldsymbol{K}_{\boldsymbol{\eta}} = \boldsymbol{K}_{\boldsymbol{\omega}}$ and

$$\begin{aligned}D(p_{\boldsymbol{xv}^{\star}}\|p_{\boldsymbol{xy}}) &= \mathbb{E}\left[\log\frac{p_{\boldsymbol{v}^{\star}|\boldsymbol{x}}(\boldsymbol{v}|\boldsymbol{x})}{p_{\boldsymbol{y}|\boldsymbol{x}}(\boldsymbol{v}|\boldsymbol{x})}\right] = \mathbb{E}\left[\log\frac{p_{\boldsymbol{\eta}}(\boldsymbol{v}-\boldsymbol{B}\boldsymbol{x})}{p_{\boldsymbol{\omega}}(\boldsymbol{v}-\boldsymbol{x})}\right] \\ &= \mathbb{E}\left[\log\frac{p_{\boldsymbol{\eta}}(\boldsymbol{\eta})}{p_{\boldsymbol{\omega}}(\boldsymbol{\eta})}\right] + \mathbb{E}\left[\log\frac{p_{\boldsymbol{\omega}}(\boldsymbol{\eta})}{p_{\boldsymbol{\omega}}(\boldsymbol{\eta}-(\boldsymbol{I}-\boldsymbol{B})\boldsymbol{x})}\right] \\ &= \mathbb{D}(p_{\boldsymbol{\eta}}\|p_{\boldsymbol{\omega}}) + \mathbb{E}\left[((\boldsymbol{I}-\boldsymbol{B})\boldsymbol{x})^{\mathrm{H}}\boldsymbol{K}_{\boldsymbol{\omega}}^{-1}(\boldsymbol{I}-\boldsymbol{B})\boldsymbol{x}\right] \\ &= \frac{1}{2}\left[\mathrm{tr}\left((\boldsymbol{B}-\boldsymbol{I})\boldsymbol{K}_{\boldsymbol{x}}(\boldsymbol{B}-\boldsymbol{I})^{\mathrm{H}}\boldsymbol{K}_{\boldsymbol{\omega}}^{-1}\right)\right]\,.\end{aligned} \tag{35}$$

Under the same attack, we can also compute the divergence $\mathbb{D}(p_{\boldsymbol{xy}}\|p_{\boldsymbol{xv}})$, in the bound (14). Following similar computation steps as (35), we get

$$\begin{aligned}\mathbb{D}(p_{\boldsymbol{xy}}\|p_{\boldsymbol{xv}^{\star}}) &= \mathbb{D}(p_{\boldsymbol{\omega}}\|p_{\boldsymbol{\eta}}) + \mathbb{E}\left[\log\frac{p_{\boldsymbol{\eta}}(\boldsymbol{\omega})}{p_{\boldsymbol{\eta}}(\boldsymbol{\omega}-(\boldsymbol{B}-\boldsymbol{I})\boldsymbol{x})}\right] \\ &= \frac{1}{2}\left[\mathrm{tr}\left((\boldsymbol{B}-\boldsymbol{I})\boldsymbol{K}_{\boldsymbol{x}}(\boldsymbol{B}-\boldsymbol{I})^{\mathrm{H}}\boldsymbol{K}_{\boldsymbol{\eta}}^{-1}\right)\right]\,.\end{aligned} \tag{36}$$

Thus, in the partial channel knowledge scenario, when Eve uses the attack strategy (29), so that (34) holds, we obtain $\mathbb{D}(p_{\boldsymbol{xv}^{\star}}\|p_{\boldsymbol{xy}}) = \mathbb{D}(p_{\boldsymbol{xy}}\|p_{\boldsymbol{xv}^{\star}})$. It is important to note that $\boldsymbol{v}^{\star}$ defined in (29) is the optimal attack strategy that minimizes $\mathbb{D}(p_{\boldsymbol{xv}}\|p_{\boldsymbol{xy}})$, however, it is not the optimal attack strategy to minimize $\mathbb{D}(p_{\boldsymbol{xy}}\|p_{\boldsymbol{xv}^{\star}})$.

V. Numerical Results

In this Section, we validate the analytical derivations and provide numerical evidence of the effectiveness of the proposed KL divergence minimizing attack. As detection test, Bob uses the generalized likelihood ratio test (GLRT), which is done by classifying the observed channel $\boldsymbol{r}$ (see Fig. 2) legitimate, for a fixed threshold τ, if

$$\Lambda(\boldsymbol{r}) = \log\frac{\max_{v} p_{\boldsymbol{v}|\boldsymbol{x}}(\boldsymbol{v}|\boldsymbol{x})}{p_{\boldsymbol{y}|\boldsymbol{x}}(\boldsymbol{r}|\boldsymbol{x})} \leq \tau. \tag{37}$$

To simulate the scenario, we consider correlated Rayleigh fading channels. Specifically, we assume that the entries of $\boldsymbol{H}$ and $\boldsymbol{G}$ are random, independent within each matrix and between the two matrices, and Gaussian distributed with zero mean and unitary variance. Moreover, we assume that $N_{\mathrm{A}} = N_{\mathrm{E}}$ and the observation $\boldsymbol{Z}$ of Eve is correlated with the legitimate cascade channel as follows

$$\boldsymbol{Z} = \rho\boldsymbol{H}\boldsymbol{\Phi}^{\star}\boldsymbol{G} + \sqrt{1-\rho^2}\boldsymbol{D}, \tag{38}$$

with $\rho \in [0,1]$ representing the correlation factor. Furthermore, $\boldsymbol{D}$ entries are independent of each other, independent of those of both $\boldsymbol{G}$ and $\boldsymbol{H}$, and follow a Gaussian distribution with zero mean and unitary variance. We set $N = 100$ IRS elements so that the approximation of the distribution of $\boldsymbol{x}$, $\boldsymbol{y}$, and $\boldsymbol{z}$ as Gaussian is meaningful.

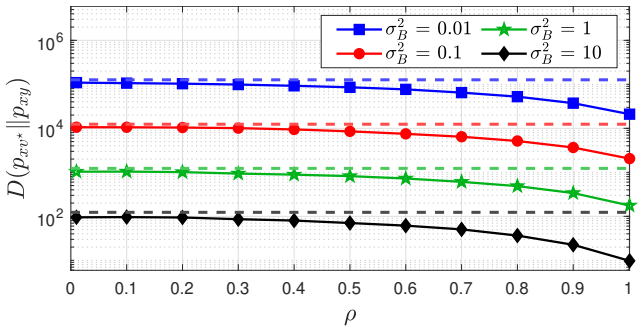

Fig. 3. KL divergence vs ρ, one curve for every value of σ_{B}^2, with $N_{\mathrm{A}} = N_{\mathrm{B}} = N_{\mathrm{E}} = 5$ antennas, and $\gamma = \pi/3$. Dashed lines represent (23) for each value of σ_{B}^2.

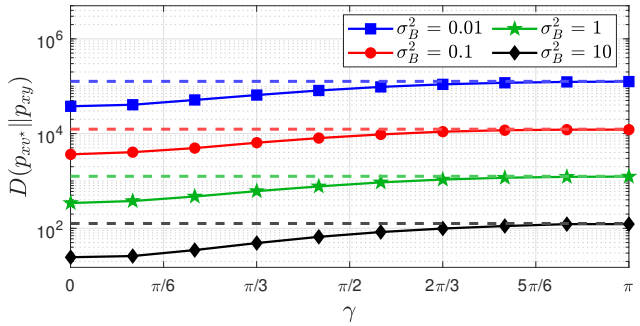

Fig. 4. KL divergence vs γ, one curve for every value of σ_{B}^2, with $N_{\mathrm{A}} = N_{\mathrm{B}} = N_{\mathrm{E}} = 5$ antennas, and $\rho = 0.8$. Dashed lines represent (23) for each value of σ_{B}^2.

Fig. 3 shows $\mathbb{D}(p_{\boldsymbol{xv}^{\star}}\|p_{\boldsymbol{xy}})$ vs ρ, when Eve uses the optimal attack strategy (29). The dashed lines represent the KL divergence (23), obtained with attack (22), which is optimal for $\rho = 0$. We show results for different values of σ_{B}^2, with $N_{\mathrm{A}} = N_{\mathrm{B}} = N_{\mathrm{E}} = 5$ antennas. We note that as $\rho \to 0$, the solid curves approached the dashed lines. Thus, the optimal attack (29) for the partial channel knowledge scenario converges, as ρ goes to zero, to the optimal attack (22) in the scenario without channel knowledge. Moreover, the KL divergence increases significantly with σ_{B}^2.

Fig. 4 shows $\mathbb{D}(p_{\boldsymbol{xv}^{\star}}\|p_{\boldsymbol{xy}})$ vs γ for $\rho = 0.8$, and $N_{\mathrm{A}} = N_{\mathrm{B}} = N_{\mathrm{E}} = 5$. We see that the KL divergence grows with γ, as expected, since the higher the γ, the more the IRS configurations chosen by Bob will deviate from the

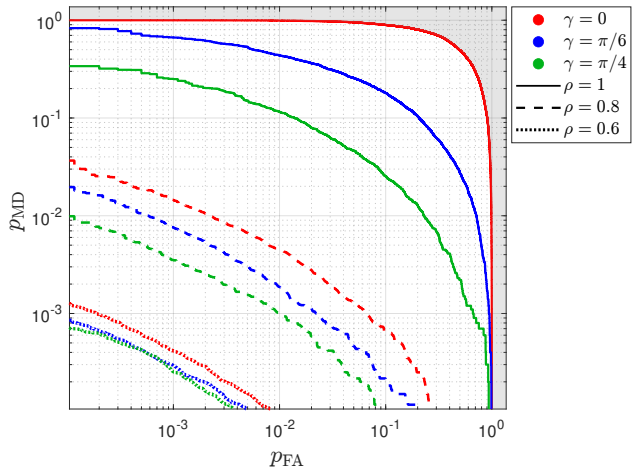

Fig. 5. DET curves for the GLRT for different value of ρ, with $N_{\mathrm{A}} = N_{\mathrm{E}} = 5$, $N_{\mathrm{B}} = 2$ antennas, and $\sigma_{\mathrm{B}}^2 = N/10$. The area corresponding to $p_{\mathrm{MD}} \geq 1 - p_{\mathrm{FA}}$ is shaded, and its edge represents the trivial limit case in which the decision is made tossing a biased coin.

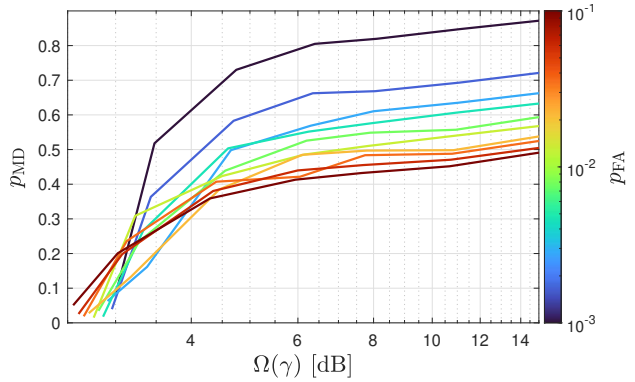

Fig. 6. p_{MD} as a function of $\Omega(\gamma)$ for various values of p_{FA} with $N_{\mathrm{A}} = N_{\mathrm{E}} = N_{\mathrm{B}} = 1$ antenna, $\rho = 0.6$, and $\sigma_{\mathrm{B}}^2 = N/2$.

communication-optimal one (3). Moreover, the KL divergence saturates to (23) when $\gamma \to \pi$.

Fig. 5 shows the detection error tradeoff (DET) curves for different value of ρ, with $N_{\mathrm{A}} = N_{\mathrm{B}} = N_{\mathrm{E}} = 5$ antennas, and $\sigma_{\mathrm{B}}^2 = N/10$. As ρ increases, the DET curves move towards the edge of the shaded area, which represents the trivial limit case when the decision is taken randomly, without using the signal. Thus, the FA probability for a given MD probability increases when the correlation between Eve's observation and the actual cascade Alice-IRS-Bob channel is higher.

The average SNR in the considered scenario is an indicator of communication performance and is defined as

$$\Omega(\gamma) = (1 - p_{\mathrm{FA}})\mathbb{E}\left[\frac{1}{\sigma_B^2}\left|\sum_{n=1}^{N} H_n G_n e^{j\theta_n}\right|^2\right] \approx (1 - p_{\mathrm{FA}})\frac{N}{\sigma_B^2}\left((N-1)\frac{\pi^2 \sin(\gamma)^2}{16\gamma^2} + 1\right), \quad (39)$$

where the approximation holds asymptotically ($N \to \infty$ by the central limit theorem). Fig. 6 shows the p_{MD} as a function of the average SNR when Eve uses the optimal attack strategy (29), for $p_{\mathrm{FA}} \in [10^{-3}, 10^{-1}]$ and considering a single antenna scenario, i.e., $N_{\mathrm{A}} = N_{\mathrm{E}} = N_{\mathrm{B}} = 1$. From Fig. 6 we see the trade-off between p_{MD} and Ω. Specifically, Ω grows with γ. However, achieving a significantly higher Ω comes at the cost of a higher p_{MD}, while choosing a lower γ setting corresponds to a high p_{MD} and thus greater vulnerability to attack.

VI. Conclusions

In this paper, we have investigated new attacks within the context of a CR-PLA, building upon the findings of [1], i.e., on performance bounds (for FA and MD probabilities) for conventional PLA. We have provided an analysis of CR-PLA with IRS, which falls under the assumptions of [1] when a large number of IRS elements is considered. We have derived the optimal attack strategy when the attacker has either partial or no knowledge of the cascade Alice-IRS-Bob channel. We have assessed the performance of the proposed attack under various correlations between Eve's observation and the legitimate cascade channel, and have presented results on the p_{MD} vs p_{FA} tradeoff for GLRT, showing the effectiveness of the proposed attack.

References

[1] A. Ferrante, N. Laurenti, C. Masiero, M. Pavon, and S. Tomasin, "On the error region for channel estimation-based physical layer authentication over Rayleigh fading," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 5, pp. 941–952, 2015.

[2] G. J. Simmons, "Authentication theory/coding theory," in *Advances in Cryptology*, G. R. Blakley and D. Chaum, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 1985, pp. 411–431.

[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, 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 Transactions on Communications*, vol. 62, no. 5, pp. 1658–1667, 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, 2019.

[6] U. M. Maurer, "Authentication theory and hypothesis testing," *IEEE Transactions on Information Theory*, vol. 46, no. 4, p. 1350–1356, 2000.

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

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

[9] N. Xie, Z. Li, and H. Tan, "A survey of physical-layer authentication in wireless communications," *IEEE Communications Surveys & Tutorials*, vol. 23, no. 1, pp. 282–310, 2021.

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

[11] F. Mazzo, S. Tomasin, H. Zhang, A. Chorti, and H. V. Poor, "Physical-layer challenge-response authentication for drone networks," in *Proc. IEEE Global Commun. Conference (GLOBECOM)*, 2023.

[12] Ö. Özdogan, E. Björnson, and E. G. Larsson, "Using intelligent reflecting surfaces for rank improvement in MIMO communications," in *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2020, pp. 9160–9164.

[13] L. Crosara, F. Ardizzon, S. Tomasin, and N. Laurenti, "Worst-case spoofing attack and robust countermeasure in satellite navigation systems," *IEEE Transactions on Information Forensics and Security*, vol. 19, pp. 2039–2050, 2024.